An evaluation of eight computer models of mammalian inner hair-cell function

Michael J. Hewitt and Ray Meddis
Department of Human Sciences, University of Technology, Loughborough LE11 3TU, United Kingdom

(Received 10 July 1990; revised 5 December 1990; accepted 29 March 1991)

Eight computer models of auditory inner hair cells have been evaluated. From an extensive literature on mammalian species, a subset of well-reported auditory-nerve properties in response to tone-burst stimuli were selected and tested for in the models. This subset included tests for: (a) rate-level functions for onset and steady-state responses; (b) two-component adaptation; (c) recovery of spontaneous activity; (d) physiological forward masking; (e) additivity; and (f) frequency-limited phase locking. As models of hair-cell functioning are increasingly used as the front end of speech-recognition devices, the computational efficiency of each model was also considered. The evaluation shows that no single model completely replicates the subset of tests. Reasons are given for our favoring the Meddis model [R. Meddis, J. Acoust. Soc. Am. **83**, 1056–1063 (1988)] both in terms of its good agreement with physiological data and its computational efficiency. It is concluded that this model is well suited to provide the primary input to speech recognition devices and models of central auditory processing.

PACS numbers: 43.64.Ld, 43.64.Pg, 43.64.Bt

INTRODUCTION

During recent decades, the input/output characteristics of the mammalian hair cell/primary-fiber complex have been the focus of extensive research. In response many computational models of the various properties of the junction have been proposed. The models provide a convenient environment for evaluating and developing theories of spike generation. Furthermore, they are an integral part of many peripheral auditory models and speech recognition systems. The problem facing the designer of such systems lies in the choice of *which* synapse model to incorporate into his/her composite model. A review of the literature finds many possible candidates (Siebert, 1965; Weiss, 1966; Eggermont, 1973, 1985; Schroeder and Hall, 1974; Oono and Sujaku, 1974, 1975; Geisler *et al.*, 1979; Ross, 1982; Schwid and Geisler, 1982; Smith and Brachman, 1982; Allen, 1983; Cooke, 1986; Meddis, 1986b, 1988; Westerman and Smith, 1988; Meddis *et al.*, 1990).

The models have appeared over a substantial period of time and each lays claim to simulate a different subset of data. Moreover, many important empirical findings have been published since the appearance of some of the models. As a result, it is difficult to make meaningful comparisons between the various proposals and it is timely to review their performance in the light of recent developments.

The choice of an appropriate model of synaptic-primary auditory-nerve (AN) models does not only affect our theoretical perspective. It could be crucial to other future modeling efforts where the model outputs will serve as inputs to simulations of more central processes. Modeling the central auditory system can become massively complex; in such systems inaccuracies at the input stage will rapidly propagate throughout the system.

This paper reports a comparative investigation of eight computational models of hair-cell function. The model results are compared to data from electrophysiological studies for the following properties: (a) onset and steady-state rate-intensity functions; (b) two-component adaptation; (c) exponential recovery of spontaneous activity after stimulus offset; (d) recovery of function to respond to a second stimulus after offset of a first; (e) additivity; and (f) low-frequency phase locking. This is the same subset of tests used by Meddis (1988) to evaluate his model; however, two additional methods of analysis are introduced. Johnson's (1980) synchronization index is introduced to analyze changes in maximum synchrony across frequency. Second, the computational efficiency of each model is considered. For this purpose, comparative times to process a 1-s tone burst are presented.

Hair-cell models can be thought of as variations on a basic proto-model consisting of reservoirs of a transmitter substance that is released across the cell membrane into the synaptic cleft. Figure 1 shows a generalized representation of many features included in hair-cell simulations. No model contains all of the features as shown in the diagram, but the transmitter flow paths of almost all the models can be represented by a reduced version of the proto-model. Early attempts modeled a single-reservoir system with loss and replenishment of transmitter quanta. Later models added extra reservoirs or complicated the principles of transmitter flow control.

A reservoir represents a store of transmitter most probably located near to the base of the hair cell. This feature is consistent with known physiological and anatomical structures which are ubiquitous among receptor–synaptic systems. Release of transmitter in quanta or discrete packets is another well-reported synaptic mechanism.

The introduction of multiple-reservoir and multiple-site models was stimulated by the empirical findings of Furukawa and his colleagues (Furukawa *et al.*, 1978; Furukawa

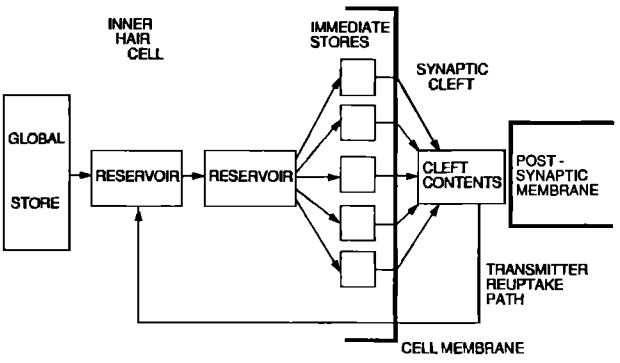


FIG. 1. Generalized representation of hair-cell models. Redrawn from Meddis (1986a).

and Matsuura, 1978; Furukawa *et al.*, 1982) working on the AN fibers of goldfish. They recorded the strengths of successive excitatory post-synaptic potentials (EPSPs) following onset of tone bursts imposed on pedestal tones. The strengths of the EPSPs (m) were, according to Furukawa, determined by the product of two independent parameters (n and p). After a stimulus increment, parameter n increased and parameter p remained relatively unchanged. Furukawa and Matsuura designated n to be the amount of transmitter available for release and p, the probability that an available quantum would be released.

These conclusions have been represented in more recent models by increasing the number of reservoirs and making their availability dependent upon stimulus intensity. In such models, the transmitter lies in reservoirs or sites close to the presynaptic membrane. The sites are ordered by increasing threshold. If the stimulus amplitude is sufficient to activate a given site, a fixed proportion per time unit of that reservoirs transmitter is released.

Meddis (1986b), however, has offered a reinterpretation of Furukawa's data. Instead of p representing the probability that a quantum of transmitter is released and n representing the amount of transmitter available for release, Meddis designated p as the probability that a given quantum will successfully traverse the cleft (and be invariant with stimulus level), and n as the number of quanta in the cleft (which would vary with stimulus level). Empirical data collected to date and known concepts of synaptic physiology cannot resolve the differences between these interpretations, and the matter remains controversial.

I. THE MODELS

Below, we study eight models of hair-cell functioning. They were chosen for study because they were capable of accepting an arbitrary[1] input signal and delivering, as output, AN event probabilities. We restricted our attention to those models that involved intermediate processing which resembled known physiological mechanisms. We also used only those models that had been developed for use on digital computers because this is the most likely medium for modelers in the immediate future. The main features of each model are outlined below.

Schroeder and Hall (1974) proposed one of the first hair-cell models to meet our criteria. A single reservoir released quanta of transmitter in amounts proportional to a stimulus-related permeability function and the number of quanta available for release. Transmitter released from the cell stimulated the post-synaptic afferent fiber. Once released, transmitter quanta were irretreviably lost from the system. A fixed-rate replenishment scheme was used to supply the reservoir with new transmitter.

Oono and Sujaku's (1974, 1975) model showed certain structural similarities to the Schroeder and Hall proposal. A single reservoir of transmitter was used together with a stimulus-related permeability function. Unlike the Schroeder and Hall model, however, only a fraction of the total transmitter was released per time unit when the cell was activated. Replenishment was effected according to the formula

$$dq(t)/dt = q(t)[1 - q(t)]/\tau, \tag{1}$$

where $dq(t)/dt$ is the rate of transmitter flow into the reservoir, $q(t)$ is the amount of transmitter in the reservoir at time t, and τ is a time constant.

Schwid and Geisler (1982) represented the conclusions of Furukawa and Matsuura (1978) with a model that consisted of six independent reservoirs of synaptic material. The reservoirs were ordered by increasing threshold such that transmitter was released from only one reservoir at low stimulus levels and from all six at very high stimulus levels. Transmitter release from an individual reservoir was governed by a permeability function related to the stimulus amplitude. Each reservoir released a constant fraction of its contents per time unit when activated. Replenishment of transmitter was a sequential process from the least- to the most-sensitive reservoir. The flow of new transmitter cascaded down the reservoir chain, each higher threshold reservoir being filled until the available replacement store was exhausted.

Ross (1982) in an attempt to replicate short- and long-term adaptation phenomena describes a model constructed of physical and mathematical elements. Four reservoirs are chained together in series. Metabolic energy flows unidirectionally through the series determined by concentration differences and fixed permeabilities. Energy from the final reservoir was fed into a Poisson generator at a rate governed by its concentration and two permeabilities, one of which is fixed, the other related to the instantaneous signal amplitude. The Poisson generator emitted unitary charges at an average rate proportional to the energy flow into the generator. These charges were fed to a leaky integrator which fired when the summed input reached a threshold. After firing the integrator was discharged and a refractory period introduced whereby inputs to the unit were inhibited for 0.5 ms. Ross modeled eight units each with a different spontaneous firing rate. For this study, we selected unit 297-43 (Ross 1982, Table I), which had a spontaneous rate of 60.6 spikes per second.

Brachman (1980) and Smith and Brachman (1982) developed a synapse model specifically to replicate two-component adaptation and additivity. Half-wave-rectified stimulus envelopes were used as inputs to the model to simulate

high-frequency tone bursts. Transduction of the input to an electrical signal was simulated by applying a modified version of Zwislocki's (1973) generalized rate-intensity function for sensory receptors. The final shape of the function was similar to the unadapted rate-intensity curve determined in empirical studies. The receptor potential was then low-pass filtered. This stage was ascribed to the functioning of the hair-cell membrane on the basis that all membranes have intrinsic resistance and capacitance that enable them to act as low-pass filters. The filtered receptor potential evoked release of transmitter.

In Brachman's model, transmitter available for release was stored in 512 independent immediate sites. During any one time frame, the number of sites contributing to the response (total transmitter released into the synaptic gap) was determined by the corresponding filtered receptor potential. Each site released a constant fraction of its contents per time unit when activated. These immediate sites were replenished by a local store of transmitter, and this, in turn, was replenished at a slower rate by a global reservoir of transmitter. Flow of transmitter was driven by diffusion gradients between the stores.

The implementation used in this study is Payton's (1988) modified version of Brachman's original model. Essentially, Payton rendered the model suitable for input of arbitrary stimuli. These modifications enhanced the performance of the model, and so it will be referred to as the Brachman–Payton model.

Cooke's (1986) peripheral auditory model includes a two-stage hair-cell simulation that transformed the mechanical input signal into AN event-rate output. The first stage simulated transduction of sound pressure to the hair-cell receptor potential using an asymmetric square-root function. His state-partition model (SPM) simulated transmitter release from the hair cell in response to the receptor potential. Following the hypothesized model proposed by Furukawa and Matsuura (1978), Cooke's conceptual model consists of many release sites spatially ordered by increasing threshold. Moreover, each site releases a constant fraction of its available transmitter per time unit when activated. However, in order to achieve computational efficiency, the working model grouped many release sites into three separate compartments. As a result only two variables (total volume and number of sites active) were required to describe each group of sites.

Sites in the first or "immediate" state have thresholds below that of the instantaneous signal amplitude and release transmitter. The second or "relax" state consists of sites that have been depleted in the recent past and are being replenished. The final state labeled "reserve" contains fully replenished sites. Each site can be instantaneously transferred between states independently of the number of sites in a particular state. This differs from other *n*-reservoir models where the transfer of sites occurs as a result of concentration gradients between sites. The replenishment rate of any one site was inversely proportional to the physical distance between the site and the replacement transmitter pool, modeled as a point source. The biological correlate of the replacement transmitter store was postulated to be the dense synaptic body seen in electron micrographs of mammalian hair cells.

Meddis (1986b, 1988) proposed a three-reservoir model incorporating a novel transmitter reuptake and resynthesis process loop. The hair cell manufactured a chemical transmitter that was delivered to a site adjacent to the cell's membrane (the free transmitter pool). The permeability of the membrane fluctuated as a function of the instantaneous amplitude of the acoustic stimulus. A small fraction of transmitter released into the synaptic gap was lost through diffusion. A further fraction was actively transferred back into the cell from the cleft. The remaining transmitter is left in the cleft to stimulate the post-synaptic afferent fiber. The transmitter taken back into the cell is held in a reprocessing reservoir for a short time before being transferred back into the free transmitter pool for later release.

Allen (1983) combined a variable resistance model of the generation of the receptor potential with an active linear transform to produce the neural response. The variable-resistance model describes the receptor potential as a function of cilia displacement. In Allen's model the receptor potential was calculated from the instantaneous signal amplitude. A linear transform was applied to the receptor potential to compute the current. The final stage was to low-pass filter the current. The low-pass filter was ascribed to the diffusion of calcium ions from the cell membrane to the site of vesicle release.

II. MODEL EVALUATION

All models were programmed in FORTRAN 77 on a Masscomp 5450 computer. To ensure, as far as possible, correct implementation of the models, we first replicated all of the figures in the eight original publications.

For the purpose of evaluation, simulations were run using a sample step of 0.05 ms (sample rate = 20 kHz). Stimuli were generated with a 2.5-ms rise time and were 1-kHz sine-phase tone bursts except where stated. All models were programmed as deterministic systems, individual spikes using pseudorandom numbers were not generated. The output of each model is expressed in terms of a firing rate and represents the average synaptic drive during a test period. The synaptic drive or "excitation function" (Gaumond *et al.*, 1983) was used by Westerman (1985) for reporting animal data and by Meddis (1988) reporting model outputs. This practice disregarded any post-synaptic mechanisms (e.g., refractory effects) included in certain models.

A. Rate-intensity functions

A steady-state rate-intensity function (RIF) derived from typical AN data is shown in Fig. 2. The spontaneous rate is a measure of the fiber's activity in the absence of sound stimulation. Spontaneous activity in AN fibers has a trimodal distribution (Liberman, 1978). The majority of fibers (>60%) fire with a spontaneous rate of greater than 18 spikes per second (known as high spontaneous-rate fibers). All of the models implemented in this study were proposed to replicate the response properties characteristic of this group of fibers. Models of low and medium spontaneous-rate

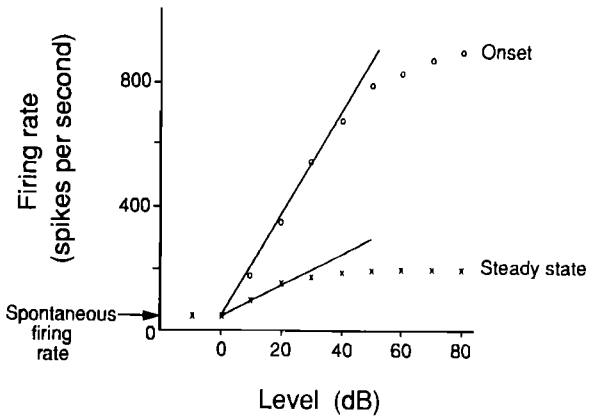


FIG. 2. Onset and steady-state rate-level functions for an auditory-nerve fiber stimulated by a tone at the fiber's characteristic frequency.

fibers have been proposed (e.g., Geisler, 1981, 1990; Meddis *et al.*, 1990) but are not evaluated here.

Electrophysiological studies distinguish a second rate-intensity function. The onset RIF is a rate measured over a short period of time following the onset of a tone burst. The function increases monotonically with tone level and shows little or no sign of saturation at high stimulus levels (Fig. 2).

The exact shape of the onset-RIF varies with the rise-time of the signal and the duration of the recording interval (Smith, 1988). In all cases, however, the refractory properties of spike generation provide an upper limit on the onset firing rate. The upper limit is about 1000 spikes/s, which corresponds to an absolute refractory period of 1 ms.

Smith *et al.* (1983) have shown that even when the onset firing rate is limited by refractory effects, increases in stimulus intensity continue to influence the onset response. As intensity increases the synchrony of onset response within the first bin also increases. Smith and colleagues conclude that the underlying synaptic drive to the fiber at stimulus onset continues to grow with increases in stimulus intensity, even when the firing rate is limited by refractory effects.

Methods to uncover the probability of spike occurrence disregarding refractory effects have been developed (e.g., Gray, 1967; Gaumond *et al.*, 1983; Westerman, 1985). When applied to post-stimulus time-histogram responses, the effects of neural refractoriness are shown to be particularly strong at onset where firing rates are highest; further evidence that the synaptic drive at onset continues to grow with increasing intensity, even when the maximum firing rate has been reached.

For each model we have determined a reference sound pressure level corresponding to the rate-intensity threshold. Using the method of Smith and Zwislocki (1975), we define zero dB as the level at which straight lines pass through the steepest portions of the onset and steady-state RIFs intersect a horizontal line drawn at an ordinate corresponding to the rate of spontaneous activity (Fig. 2).

Following the methods of Westerman and Smith (1984), onset and steady-state RIFs were obtained for each model in response to 300-ms tone bursts. The onset function represents the firing rate averaged over the first (or highest, if not the first) ms of the response and the steady-state function represents the firing rate averaged over the last 20 ms of response.

The results are presented in Fig. 3. The empirical data obtained from gerbil (Westerman, 1985) are shown in the top-left panel. This format is common to all figures where modeled outputs are compared to animal data. All models

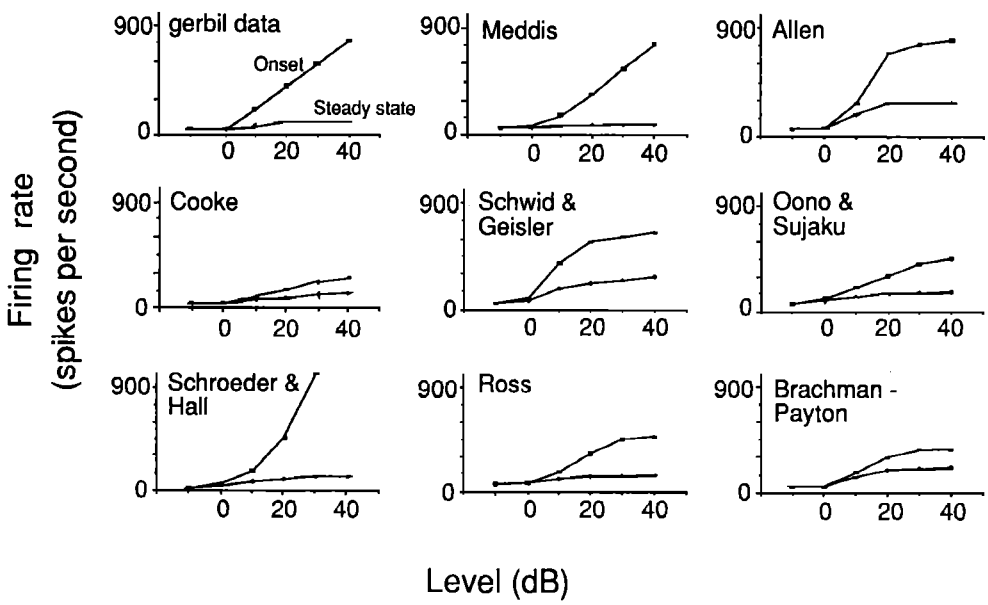


FIG. 3. Comparison of rate-level functions between the models and Westerman's (1985) gerbil data. The onset function represents the firing rate during the first (or highest) ms after tone onset. The steady-state function represents the average firing rate over the last 20 ms of a 300-ms tone burst.

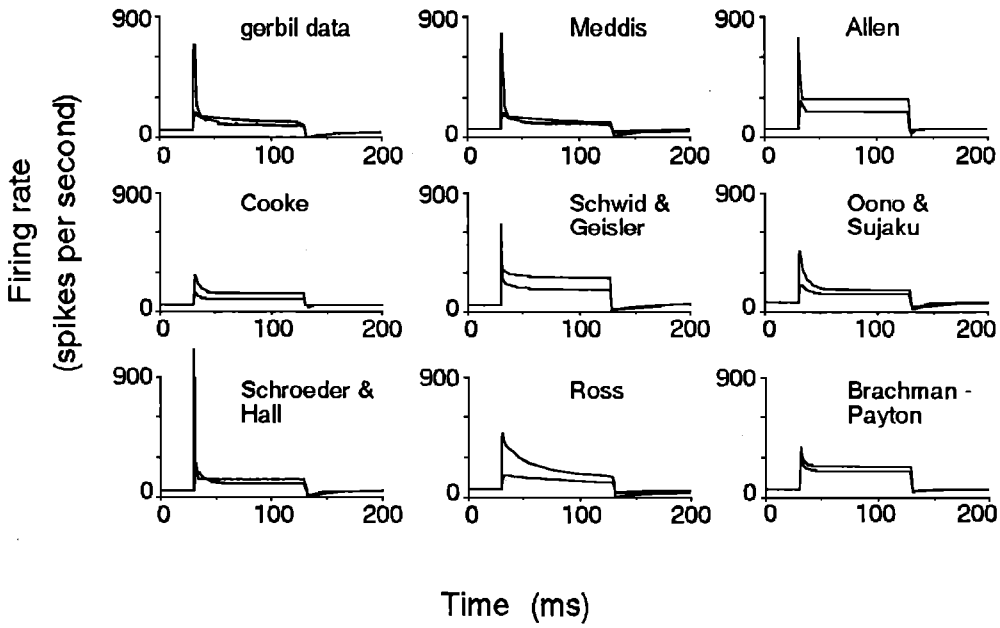


FIG. 4. Post-stimulus time excitation histograms for the models compared to Westerman's (1985) gerbil data for stimulus presentation levels of 10 and 40 dB.

produce the predicted dichotomy between onset and steady state RIFs although the precise form of the functions varies between models. However, it should be noted that the onset function of the Schroeder and Hall model is very unrealistic. The rate grows far too rapidly (spikes/s per dB) at medium and high stimulus levels.

B. Two-component adaptation

Following a period of silence, a tone burst will produce an initial peak of activity followed by a decline that is initially rapid then slower. Westerman and Smith (1984) propose that the adaptation response consists of two exponentially decaying components:

$$y(t) = A_R e^{-t/\tau_R} + A_{ST} e^{-t/\tau_{ST}} + A_{SS}, \quad (2)$$

where A_R and A_{ST} are the magnitudes and τ_R and τ_{ST} the decay time constants of the rapid and short-term adaptation components; A_{SS} is a constant representing the steady-state response.

The first component, known as rapid adaptation, has a time constant (τ_R) of less than 10 ms and decreases with increasing tone level to as low as 1-ms for high-amplitude stimuli (Westerman and Smith, 1984; Yates and Robertson, 1980; Yates *et al.* 1985). The second time constant (τ_{ST}) is about 70-ms (range 20–100 ms) and is independent of tone level (Smith and Zwislocki, 1975; Westerman and Smith, 1984).

Post-stimulus time histograms (PSTHs) were constructed from model responses to 100-ms tone bursts at various intensities. Figure 4 shows the result for each model for stimulus levels of 10 and 40 dB above threshold. Time constants of adaptation derived from the responses and typical AN data (Westerman, 1985) are shown in Fig. 5. The method used for fitting the exponentials to the model responses is given in Meddis (1988).

The response of the Schroeder and Hall model shows an onset peak that is far too large at medium and high signal levels. This gives an adaptation response that fails to match the two-component function reported in the literature. Allen's model also fails to produce a two-component adaptation response.

The remaining models show two-component adaptation; however, only the Meddis model demonstrates approximately level-*independent* short-term time constants as well as level-*dependent* rapid time constants.

The models proposed by Ross, and Oono and Sujaku show two-component adaptation only at levels above 10 dB. Ross highlighted the insufficient onset peaking at medium stimulus levels as a major fault of the model. Further analysis has shown that this defect was due, in part, to the model neuron that Ross included in the original simulation. Removal of this component has shown that insufficient onset peaking is only a problem at low (below 10 dB) stimulus levels. In contrast, Westerman's gerbil data (Westerman and Smith, 1984; Westerman, 1985) show a prominent onset peak at a stimulus level of 8 dB.

C. Recovery of spontaneous activity

At the offset of a tone burst, AN firing briefly ceases before recovering back to spontaneous rate. The recovery function is described by a single exponential, with a time constant between 40 and 100 ms (Harris and Dallos, 1979; Smith, 1977; Westerman, 1985). Yates *et al.* (1985) noted that the amount of depression of firing at tone offset, and the exact nature of recovery, depended on the stimulus level, and not on the absolute value of spontaneous rate. They report a "typical" recovery time constant of 20 ms, accompanied by a small amount of slower recovery following the offset of more intense stimuli.

TABLE I. Time constants of recovery to spontaneous rate after the offset of a 40-dB tone burst.

Model	Time constant (ms)
Schroeder and Hall	9.2
Oono and Sujaku	7.9
Allen	2.3
Schwid and Geisler	9.4
Ross	57.6
Brachman–Payton	4.0
Cooke	8.1
Meddis	30.4
Chinchilla (+ 20 dB) Harris and Dallos (1979)	37.0
Gerbil (+ 40 dB) Westerman (1985)	47.0
Guinea pig Yates *et al.* (1985)	20.0

The time constants of recovery to spontaneous-firing rates were derived from the model PSTHs (Fig. 4) and are shown in Table I along with data from three species. The models of Meddis and Ross show the most realistic recovery of spontaneous activity. However, the complete cessation of firing at tone offset is not shown by any model. Moreover, the remaining models show very unrealistic recovery properties. Quantitatively, the process occurs in the models at a much faster rate than that derived empirically.

D. Recovery of adaptation functions

Physiological forward masking paradigms have further helped our understanding of AN recovery processes. The work of Smith (1977) and Harris and Dallos (1979) investigated the recovery from adaptation of single AN fibers in gerbil and chinchilla, respectively. In brief, it was established that AN recovery from masking stimuli followed a single exponential curve. Both studies noted, however, that the response at stimulus onset recovered at a faster rate than did the total response.

This initial, faster component of recovery was studied in detail by Westerman (1985) using the paradigm shown in Fig. 6. The response to the probe was measured as a decrement when compared to the response in the absence of the preceding masking tone. Using analysis windows of 1 and 20 ms (onset and short-term measures), Westerman fitted two recovery time constants to the data; the fast component with a time constant of between 20 and 50 ms and the second, slower component with a time constant of greater than 150 ms.

An examination of Table I suggests that all but two models (Ross, and Meddis) would fail to mimic the data of Westerman. In most cases, the recovery of spontaneous activity following tone offset is far too rapid. Figure 7 confirms the poor modeling of recovery processes. The models proposed by Ross and Meddis quantitatively predict the slower short-term recovery rate but also incorrectly predict an onset recovery rate of equal slope. It is worth noting, however, that the data points between 0 and 20 ms of Westerman's results do seem to depart from a single exponential improvement, although, he chose to draw a single best-fit line

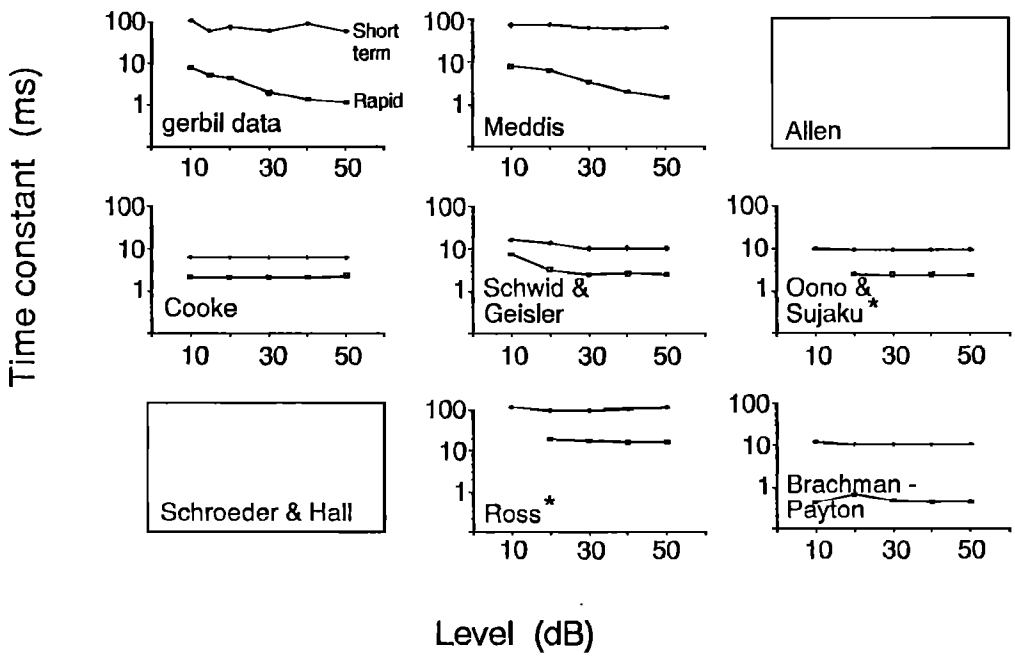


FIG. 5. Time constants of adaptation fitted to model responses compared to Westerman's (1985) data fitted to derived excitation functions. The two-component adaptation response equation (2) could not be fitted to the responses from the Schroeder and Hall model, or to Allen's model. Asterisks indicate single-component adaptation at 10 dB.

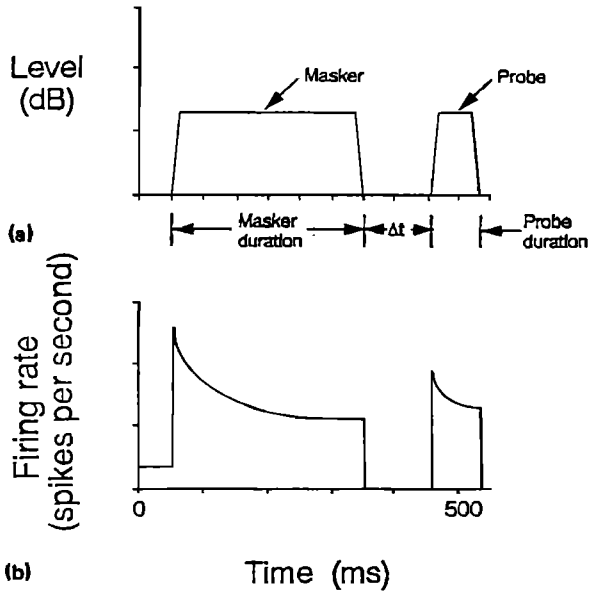


FIG. 6. (a) Forward masking paradigm used by Westerman (1985) to study recovery of response after adaptation. Masker duration: 300 ms; probe duration: 30 ms; masker and probe at 43 dB, Δt varied between 0 and 200 ms. (b) Schematic of corresponding AN fiber responses.

through the points. The same deviation can be seen in the onset responses of the Meddis model. The issue is not clear-cut, however, as a more detailed examination of Westerman's data show large variations in the recovery times measured from his fiber population. It seems that additional data are required to quantify the exact nature of the recovery functions.

E. Additivity

Smith and Zwislocki (1975) have established that the process of short-term adaptation is additive in nature. They showed that the increase in firing rate elicited by a stimulus amplitude increment is independent of the state of adaptation of the fiber. Similar additivity was found for responses to decrements in intensity applied after adapting tones (Smith, 1977; Abbas, 1979).

Later, Smith *et al.* (1985) studied additivity with analysis windows designed to emphasize the properties of rapid adaptation. The stimulus paradigms used by Smith and his colleagues are shown in Fig. 8. Large-window analysis (20-ms) confirmed the earlier findings that adaptation was additive for stimulus increments and decrements. The effect was also shown to be valid for small-window (1-ms) responses to stimulus increments. In contrast, the small-window decremental response decreased with increasing time delay, and in proportion to the decrease in firing rate produced by the pedestal. Normalized model responses to the additivity paradigms are shown in Fig. 9, together with the empirical data of Smith *et al.* (1985).

All of the computed responses showed spike-rate changes that were quantitatively very different from the empirical data of Smith *et al.*; however, it was evident that some of the models showed the correct trends. The multiple-reservoir models proposed by Brachman-Payton, and Schwid and Geisler qualitatively reproduce the data of Smith *et al.* (1985). The small- and large-window increment response

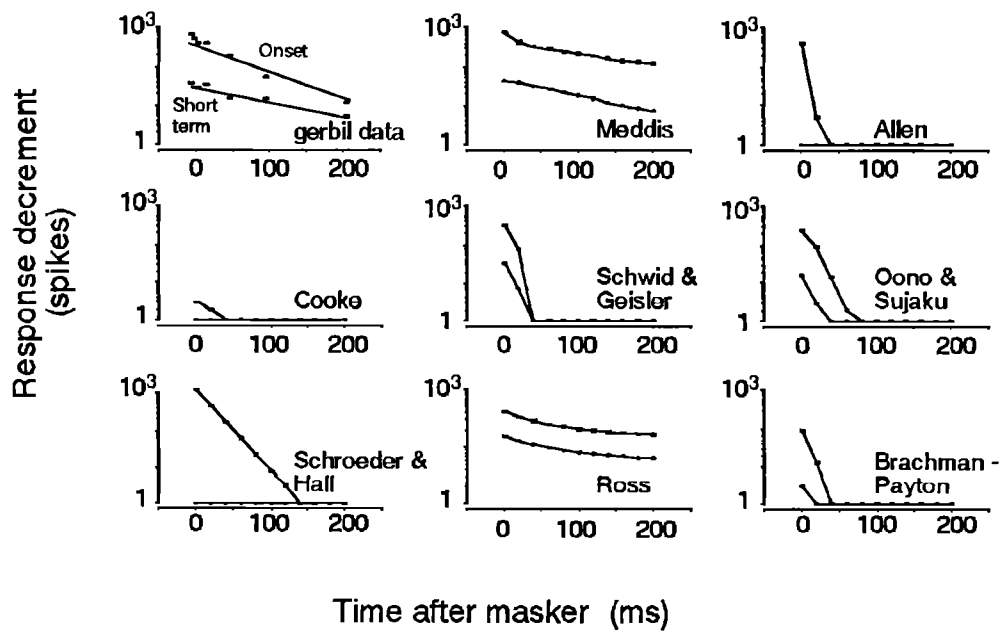


FIG. 7. Recovery of response components after adaptation. Comparison of models and Westerman's (1985) data (including his "best-fit" lines). The decrement is the difference between the response of an unadapted fiber to a 43-dB tone and the response of a fiber to the probe after the offset of a 300-ms, 43-dB masking tone of the same frequency. The onset decrements are based on the maximal 1-ms firing rate after probe onset. The short-term decrements are based on the rate between 10 and 30 ms after the probe onset.

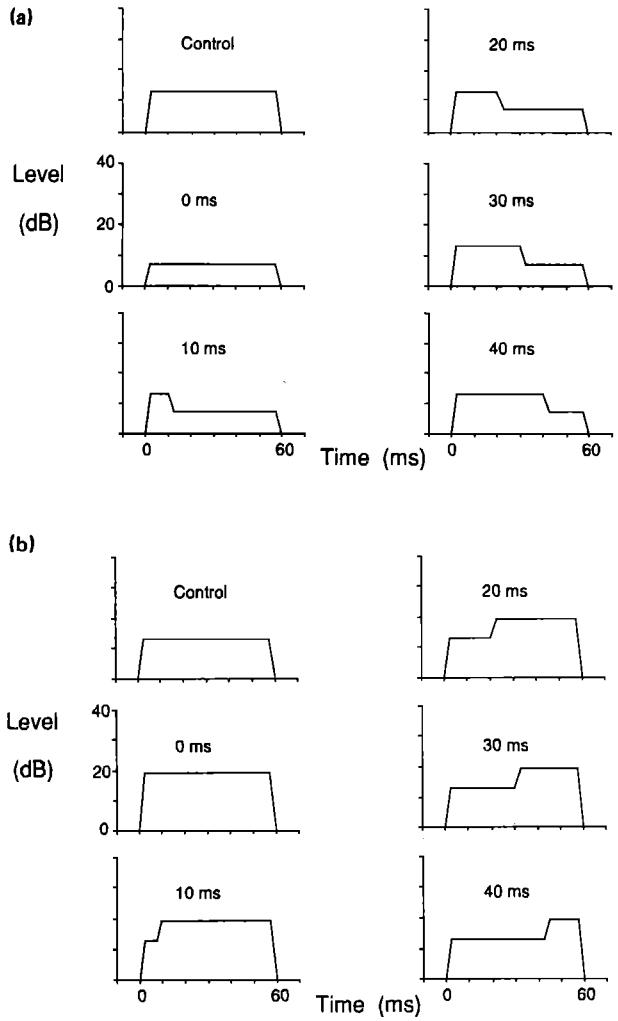


FIG. 8. Schematic representation of the paradigms for studying additivity. (a) Decrements of 6 dB applied at various time delays after the onset of a pedestal. (b) Increments of 6 dB applied to a pedestal. In both (a) and (b), the pedestal was of 60-ms duration and 13 dB above threshold. The control stimulus was the pedestal alone.

functions are horizontal (i.e., additive), as is the large-window decrement response function; finally, the small-window decrement function decreases with increasing delay.

Cooke's model is additive for stimulus increments but fails to produce a large-window time-independent decremental response. The models proposed by Oono and Sujaku, Allen, and Meddis successfully replicate three functions but the small window incremental response is nonadditive and discrepant with the empirical data at very short time delays (< 10 ms).

The models proposed by Schroeder and Hall, and Ross predict time-dependent incremental responses for both large and small windows and are clearly nonadditive in nature.

F. Low-frequency phase locking

This section deals with the fine-time structure of AN responses to single-frequency tones at different levels and frequencies.

1. Changes with frequency

Period histograms (PHs) show that AN-fiber responses phase-lock to the positive half-cycle of low-frequency tones. At high signal frequencies, however, the period histogram shows no relationship to the signal's phase characteristics. Rose *et al.* (1967) quantified the loss of synchrony with the synchronization coefficient (the density of the most populous half of the period histogram divided by total density).

Johnson (1980) studied first-order synchrony effects over a range of frequencies using the synchronization index (SI). Using the Fourier transform of the period histogram, Johnson calculated the SI by taking the response magnitude at the stimulus frequency divided by the magnitude at dc. Below 1 kHz, the SI is independent of stimulus frequency. Above 1-kHz, the data are modeled by $\mathrm{SI} = 1 - f/6$, where $f =$ frequency in kHz.

Johnson's method assumes that the period histogram is sinusoidal, an assumption that is not always justified especially at low frequencies and high signal levels. Accordingly, it is helpful to use both measures.

To measure model synchronization coefficients and indices, stimuli of 1, 2, 3, 4, and 5 kHz were used and the sample rate increased to 100 kHz. The results are presented in Figs. 10 and 11. Only three models show frequency-dependent synchronization measures. The models proposed by Meddis and Allen simulate the falloff of synchronization with increasing frequency accurately. One small deviation from the empirical data occurs in the 5-kHz case where almost complete loss of synchronization should be evident. The measures of synchronization for the Brachman–Payton model are similar, but the rate of falloff at 3 kHz and above is too shallow.

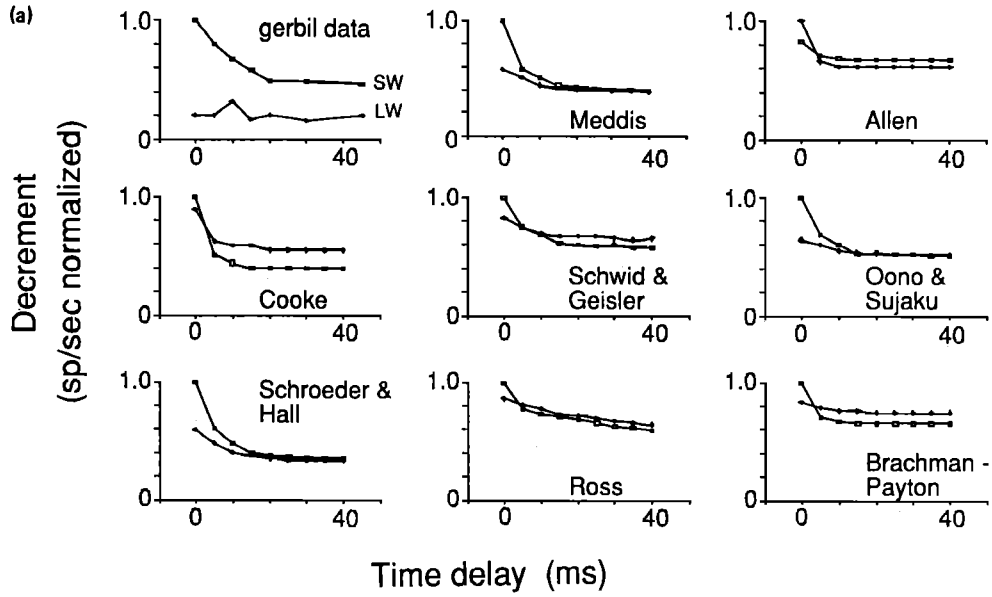
2. Changes with level

Johnson (1980) also used the SI to monitor synchronization changes with level. Double ordinate plots of SI and firing rate changes over a range of stimulus intensities show the modulation of spontaneous activity by an input stimulus occurs at intensities below rate threshold. From such data, the modulation threshold first noted by Rose *et al.* (1971) can be derived. Figure 12 shows that all models successfully replicate this effect.

G. Computational efficiency

As stated in the Introduction, hair-cell models are incorporated into models of speech recognition and central auditory processing.[2] Modelers of such systems have to consider the trade-off between speed and accuracy of the hair-cell model they select. To complete this study, we have measured the relative computational efficiency of each model.

Comparative times taken for each model to process a 1 s tone burst (including preceding and following periods of silence) presented at 40 dB above threshold are shown in Table II. No attempt was made to optimize the computer code of any model. An examination of the results show negligible differences between most models. Exceptions to this are the multiple-reservoir models proposed by Schwid and Geisler



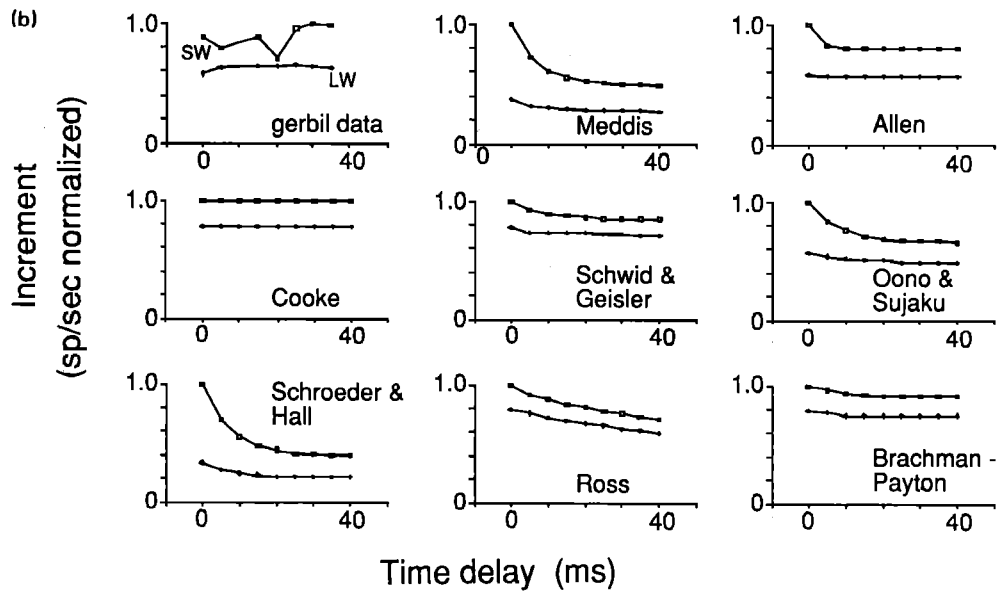


FIG. 9. Model responses and the empirical data of Smith *et al.* (1985) to the additivity stimuli. (a) Responses to decrements in level as a function of time delay. (b) Responses to stimulus increments as a function of delay. The small window (SW) is based on the response during the first ms after the stimulus change, and the large-window (LW) measure is based on the first 20 ms after the stimulus change.

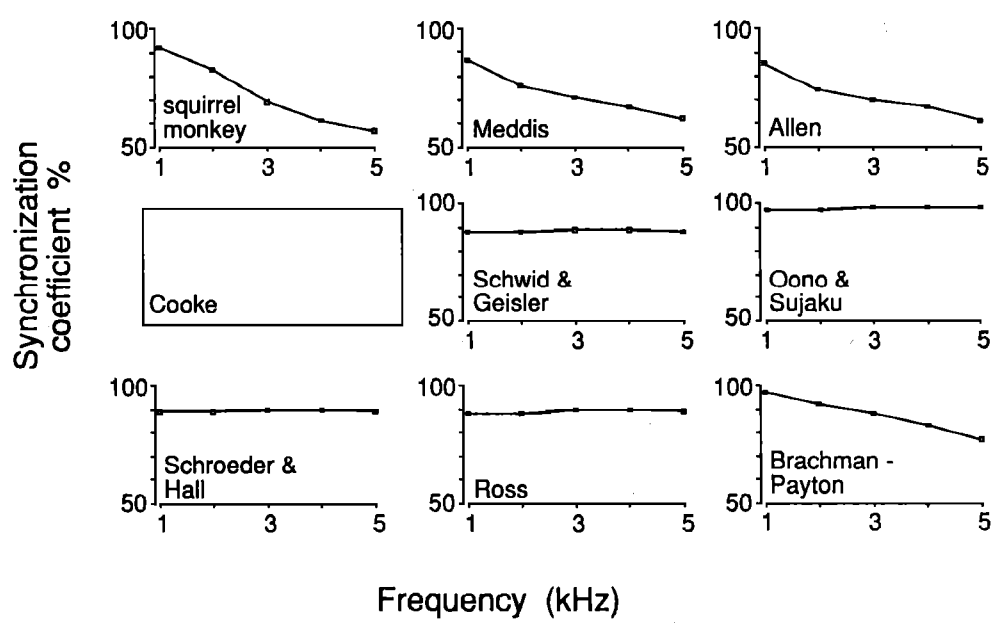


FIG. 10. Model synchronization coefficients as a function of frequency compared to squirrel monkey data from Rose *et al.* (1967). No synchronization data for Cooke's envelope model.

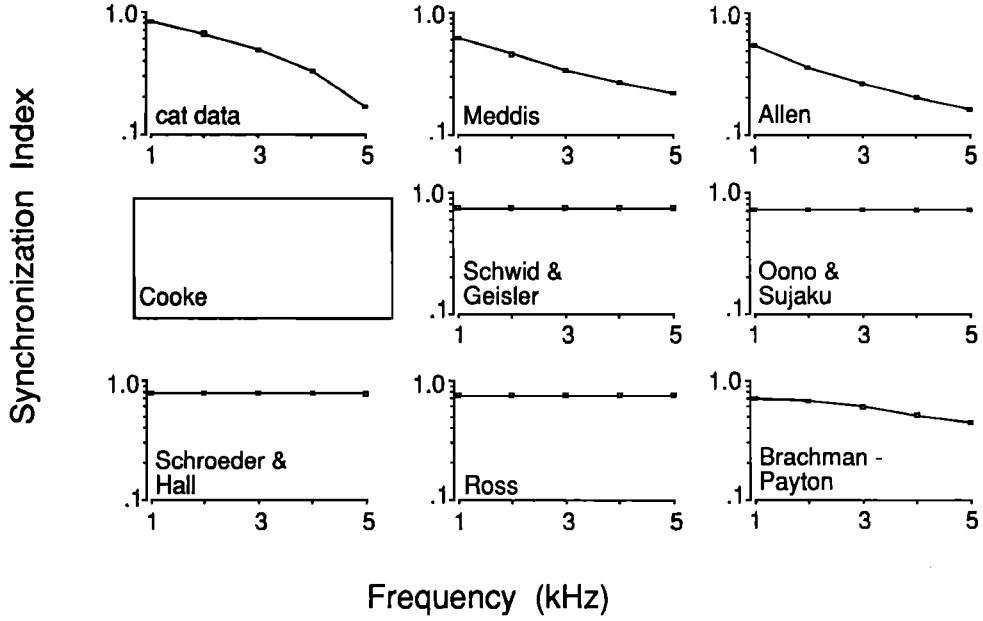


FIG. 11. Model synchronization indices as a function of frequency compared to Johnson's (1980) cat data.

(1982) and Brachman–Payton (Payton, 1988), which are relatively slow.

Although one particular implementation of a model may be optimized to run faster than another, it is clear that tracking the contents of discrete immediate sites (e.g., Brachman–Payton, Schwid and Geisler) reduces computational efficiency. The return for the increased computation time required to run multiple-site models is the time-independent response to incremental stimuli.

III. DISCUSSION

This paper has reported a detailed computational investigation of eight hair-cell models. The properties of each model are summarized in Table III.

A. Adaptation

AN-fiber responses adapt in response to a constant-level stimulus. This property is not exclusive to the hair-cell/AN

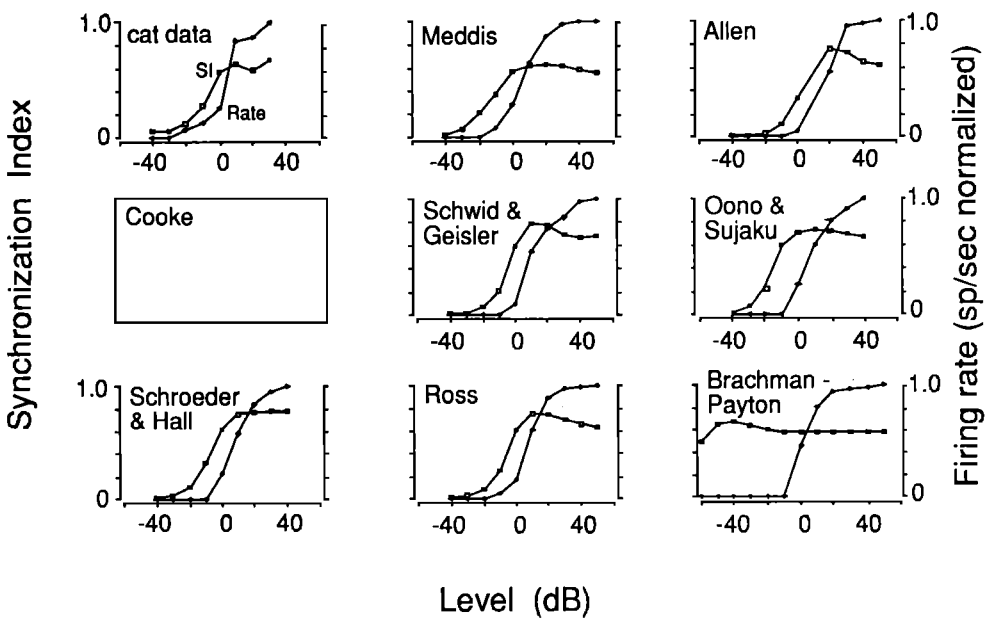


FIG. 12. Synchronization index as a function of stimulus level for model responses and empirical data from cat (Johnson, 1980), together with normalized steady-state rate-level functions.

TABLE II. Time taken by each model to process a 1-s tone burst presented at 40 dB.

Model	Time (s)
Schroeder and Hall	2.8
Oono and Sujaku	2.0
Allen	3.2
Schwid and Geisler	42.0
Ross	2.0
Brachman–Payton	320.0
Cooke	1.2
Meddis	4.0

complex, but is inherent in all sensory-receptor systems. Available experimental evidence suggests that spike activity in AN fibers is initiated by the release of a transmitter substance from the hair cell into the cleft that synapses with the fiber. The hair-cell receptor potential that is likely to cause the release of transmitter shows no adaptation (Russell and Sellick, 1978; Goodman *et al.*, 1982). It is, therefore, believed that adaptation is due to a progressive rundown of transmitter substance within the cell. Other available evidence (e.g., Furukawa and Matsuura, 1978) substantiates this viewpoint.

In general, the models evaluated here show an adaptation of response that can be explained by the dynamics of depletion and replenishment of transmitter substance. The mechanisms by which the specific adaptation components are modeled, however, deserve further comment.

Adaptation in the single-reservoir model proposed by Schroeder and Hall (1974) is produced by the depletion of transmitter quanta from the reservoir; the process continues until the loss of quanta is balanced by the replacement quanta. As the probability of release was directly related to stimulus intensity, the rate of adaptation varied with stimulus intensity in a way not consistent with neural data (e.g., Kiang *et al.*, 1965).

The proposal of Oono and Sujaku (1974) assumed a single-reservoir model with transmitter-release properties similar to those of the Schroeder and Hall model. However, more realistic adaptation properties are produced by the model as only a fraction of the available transmitter was released per time unit when the cell was activated. The principle of fractional release is inferred from the work of Furukawa and Matsuura (1978), and was adopted into the later multiple-reservoir models (Schwid and Geisler, 1982; Smith and Brachman, 1982; Cooke, 1986). Although both short-term and rapid time constants can be fitted to the post-stimulus time histograms of these models, a more careful consideration of two-component adaptation was shown by Meddis (1988), Ross (1982), and Smith and Brachman (1982).

The Meddis model quantitatively simulates Westerman and Smith's rapid adaptation data by fast reuptake of transmitter from the cleft by the presynaptic membrane. This limits the amount of free transmitter available for release. Transmitter taken up from the cleft is subject to a reprocessing delay before being available for release again. This delay process, as well as global replenishment of the free transmitter store (largely), determines the short-term adaptation time constant. There is no direct evidence to support this detail of the model; it was simply incorporated as a possible explanation. However, no other model has addressed the issue of what happens to the transmitter in the cleft.

The models proposed by Brachman–Payton and Ross simulate rapid adaptation by relying on small volume reservoirs lying close to the presynaptic membrane. During stimulation, these reservoirs (Brachman's immediate sites and reservoir v4 of Ross) become depleted of their transmitter with a time constant matched to that of rapid adaptation. These immediate stores are replenished by larger local stores. Short-term adaptation is simulated by the slower depletion of the local stores.

B. Recovery

While much time has been devoted to modeling adaptation, the simulation of recovery processes at the offset of a tone burst has largely been neglected. The presence of two different exponential recovery functions is not found in any modeled outputs. However, Westerman and Smith's (1988) analytical model provides some interesting insights to recovery processes. The model is based on the configuration of reservoirs proposed by Brachman, but differs in that the

TABLE III. Summary of the eight models response properties.

Model	Number of reservoirs	Number of parameters	Accept arbitrary stimulus	Two-component adaptation	Additive (to stimulus increments)	Physiological forward masking response	Low-frequency phase locking	Computationally efficient
Schroeder and Hall	1	3	yes	no	no	no	no	yes
Oono and Sujaku	1	3	yes	yes	no	no	no	yes
Allen	1	10	yes	no	no	no	yes	yes
Schwid and Geisler	6	15	yes	yes	yes	no	no	no
Ross	4	9	yes	yes	no	no	no	yes
Brachman–Payton	1024	19	yes	yes	yes	no	yes	no
Cooke	3	6	no	yes	yes	no	no	yes
Meddis	3	8	yes	yes	no	no	yes	yes

model parameters were determined directly from experimental data (see Westerman and Smith, 1988, for details). Their analysis suggests that two-component recovery can be simulated by the Brachman configuration if the rates of replenishment to the local and immediate stores are approximately equal. This implies that replenishment of transmitter is not preferential to one store.

C. Additivity

An important constraint on most models is the failure to replicate additivity phenomena. The most crucial assumption made by models with additive properties is the existence of many release sites spatially ordered by increasing threshold (cf. Furukawa and Matsuura, 1978). All other model configurations evaluated in this study could not mimic the main additivity data.

D. Synchronization

Another failure common to many models is the absence of reduced response synchronization to high-frequency stimuli. Only the models proposed by Meddis, Allen, and Brachman–Payton show the effect.

In the Meddis model, phase locking is affected by the rate at which transmitter can be cleared from the synaptic cleft. The rates of transmitter loss and re-uptake from the cleft are independent of stimulus frequency. When the stimulus frequency is high relative to these rates the ability of the system to reflect the waveform's fine structure is thus reduced.

Allen used a low-pass filter in his model to reduce synchrony at high frequencies. The source of the filter was ascribed to the diffusion of calcium ions from the cell wall to the synaptic region of the cell where transmitter vesicles reside.

The Brachman model used a low-pass filter to simulate the intrinsic resistance and capacitance of cell membranes. Payton (1988) examined the synchrony responses of the Brachman model before and after adaptation comparing the results to Johnson's synchronization data. The low-pass filter implemented produced results that showed excellent agreement with the empirically derived $SI = 1 - f/6$. However, the adaptation response imposed changes that were characteristic of high-pass filtering, which resulted in the loss of well-matched indices. To regain appropriate synchronization data, Payton introduced a second-order low-pass filter after the adaptation stage of the model. Excellent results were obtained with the time constant of the second-order filter set to 0.05 ms. Payton removed the second filter from her final implementation due to lack of physiological evidence for such a filter. It does, however, raise further questions about the nature of the filtering required to simulate synchrony data.

A number of factors have been proposed to account for the loss of synchronization in AN discharges with increasing frequency. Anderson (1973) presented a quantitative model based on the premise of temporal jitter within the periphery. Russell and Sellick (1978) noted from their inner hair-cell recordings that phase locking disappeared when the cell's ac response became small in comparison with the dc response. The attenuation of the ac response was ascribed to the capacitance of the cell membrane. Later, Palmer and Russell (1986) established from direct measurements of the membrane time constant and the ac/dc ratios that phase locking did indeed decline directly in proportion to the decline in the inner hair cells' ac/dc ratio, suggesting that the limiting factor for phase locking is the hair-cell time constant.

E. Analytical models of hair cell-functioning

Our study has confined itself to models that were implemented by their authors on digital computers. However, some analytic models deserve some comment. For example, Eggermont's (1985) model, which features processes known to operate at the neuromuscular junction, shows realistic adaptation and recovery properties. Transmitter is released from an hypothesized hair cell at a rate proportional to the amount of transmitter available for release and a stimulus (envelope) amplitude related permeability function. Transmitter quanta in the cleft bind with free receptor sites on the post-synaptic afferent-nerve fiber. After activation of sites by quanta, the receptor sites are converted to an inactive or occupied state. Inactivation occurs at a rate that depends on the number of receptors that can be occupied as well as on the number that are occupied. Enzymatic action on the receptor-transmitter complex frees receptor sites. The receptor site is only free to combine with transmitter quanta after a period of recovery. The number of receptor-transmitter complexes represents the post-synaptic excitatory potential.

Eggermont's model differs from other proposals in that it considers two adaptation mechanisms, the first located in the hair-cell/synapse and the second in the auditory nerve. In brief, it has the following properties: (i) additive responses to stimulus increments; (ii) short-term adaptation time constants independent of stimulus level; and (iii) forward masking responses similar to those found in the empirical studies of Harris and Dallos (1979) and Smith (1977).

A different approach to modeling AN data was taken by Westerman and Smith (1988). Using experimental data on gerbil AN responses, they were able to derive model parameters suited to a three-stage synaptic model such as the one proposed by Brachman. To accurately model two-stage adaptation in accordance with empirical data, they propose that both the volume and membrane permeability of the reservoirs should be stimulus level dependent. This differs from all previous synaptic models where it is considered that only one function is required to be level dependent.

F. Biological correlates

Many of the models evaluated in this study could be described as phenomenological. The models have been designed to replicate certain response properties of AN fibers and then the processing steps are correlated with known physiological processes. This approach has resulted in models that are computationally convenient and already well suited to applications such as speech recognition devices that use models of the peripheral auditory system as input devices. However, the models do not represent the biological processes accurately enough to suggest crucial experimental

designs to evaluate theories of AN spike generation. Future hair-cell/AN models will need to consider, for example, the role of intracellular calcium; potassium, and calcium-activated potassium channels; transmitter re-uptake and post-synaptic mechanisms. Even though such models would be complex and computationally expensive, they will provide a means of assessing the relative merits of any theory of AN spike generation.

G. Nonlinearity

It is well known that the release of transmitter at the hair-cell/synapse is preceded by a number of nonlinearities. For example, input–output functions for basilar membrane vibration show saturation of response at the best frequency for levels as low as 30 dB SPL (Sellick *et al.*, 1982; Yates *et al.*, 1990). The nonlinear preprocessing would have had an important role in shaping the neural responses reported in the literature, and, will have to be accounted for in future hair-cell modeling efforts. All existing models appear to assume that the basilar membrane response is linear. Nonlinear permeability functions, however, may conceal effects that might be better attributed to mechanical processes prior to the hair-cell transduction.

IV. CONCLUSIONS

We have examined the relative merits of eight computational hair-cell models by comparing the outputs with mammalian auditory-nerve data. The evaluation shows that the Meddis model shows only minor discrepancies with the empirical data. The Meddis model is also one of the most computationally efficient and is well suited to provide the primary input to larger systems such as speech-recognition devices and models of central-auditory processing.

One important constraint on the Meddis model is the failure to replicate the additivity effect for onset responses. This problem could be overcome by introducing multiple-release sites in accordance with Furukawa and Matsuura's conceptual model. However, this would have a detrimental effect on the model's relatively fast run time.

In this paper we have not considered the response characteristics of low spontaneous-rate fibers, or, the models that have been proposed to replicate their properties. Both Geisler (1981, 1990) and Meddis *et al.* (1990) have published preliminary findings on the response properties of their low spontaneous-rate models. Work is currently in progress in our laboratory to make a more detailed study of the response properties of the models and will be the subject of a future paper.

ACKNOWLEDGMENTS

We are indebted to Jont Allen, Martin Cooke, Dan Geisler, and Karen Payton for generously supplying program listings and/or model parameters. We are grateful to Bob Smith for helpful discussions. We thank Trevor Shackleton, Strange Ross, and an anonymous reviewer for valuable comments on an earlier version of this paper.

[1] One exception to this is Cooke's model which requires the input of half-wave-rectified stimulus envelopes.

[2] Even though it may not have been the intention of the authors who proposed the models, most, if not all, of the simulations evaluated here have been used to provide the primary input to larger systems (e.g., Payton, 1988).

Abbas, P. J. (**1979**). "Effects of stimulus frequency on adaptation in auditory-nerve fibers," J. Acoust. Soc. Am. **65**, 162–165.

Allen, J. (**1983**). "A hair cell model of neural response," in *Proceedings of the IUTAM/ICA Symposium*, edited by E. de Boer and M. A. Viergever (Martinus, The Hague, and Delft U. P., Delft, The Netherlands), pp. 193–202.

Anderson, D. J. (**1973**). "Quantitative model for the effects of stimulus frequency upon synchronization of auditory discharges," J. Acoust. Soc. Am. **54**, 361–364.

Brachman, M. L. (**1980**). "Dynamic response characteristics of single auditory-nerve fibers," Special Rep. ISR-S-19, Institute for Sensory Research, Syracuse University, Syracuse, NY 13210.

Cooke, M. P. (**1986**). "A computer model of peripheral auditory processing," Speech Commun. **5**, 261–281.

Eggermont, J. J. (**1973**). "Analogue modelling of cochlear adaptation," Kybernetic **14**, 117–126.

Eggermont, J. J. (**1985**). "Peripheral auditory adaptation and fatigue: A model orientated review," Hear. Res. **18**, 57–71.

Furukawa, T., Hayashida, U., and Matsuura, S. S. (**1978**). "Quantal analysis of the size of post-synaptic potentials at synapses between hair cells and afferent nerve fibers in goldfish," J. Physiol. (Lond.) **276**, 211–226.

Furukawa, T., and Matsuura, S. S. (**1978**). "Adaptive rundown of excitatory post-synaptic potentials at synapses between hair cells and eighth nerve in the goldfish," J. Physiol. (Lond.) **276**, 193–209.

Furukawa, T., Mikuki, K., and Matsuura, S. (**1982**). "Quantal analysis of a decremental response at hair cell-afferent fiber synapse in the goldfish sacculus," J. Physiol. (Lond.) **322**, 181–195.

Gaumond, R. P., Kim, D. O., and Molnar, C. E. (**1983**). "Response of cochlear nerve fibers to brief acoustic stimuli: Role of discharge-history effects," J. Acoust. Soc. Am. **74**, 1392–1398.

Geisler, C. D. (**1981**). "A model for discharge patterns of auditory-nerve fibers," Brain Res. **212**, 198–201.

Geisler, C. D. (**1990**). "Evidence for expansive power functions in the generation of the discharges of 'low- and medium-spontaneous' auditory-nerve fiber," Hear. Res. **44**, 1–12.

Geisler, C. D., Le, S., and Schwid, H. (**1979**). "Further studies on the Schroeder–Hall hair-cell model," J. Acoust. Soc. Am. **65**, 985–990.

Goodman, D. A., Smith, R. L., and Chamberlain, S. C. (**1982**). "Intracellular and extracellular responses in the organ of Corti of the gerbil," Hear. Res. **7**, 161–179.

Gray, P. R. (**1967**). "Conditional probability analyses of the spike activity of single neurons," Biophys. J. **7**, 759–777.

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

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115–1122.

Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). "Discharge patterns of single fibers in the cat's auditory nerve," Res. Monogr. No. 35, MITP, Cambridge, MA.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442–455.

Meddis, R. (**1986a**). "Exploiting physiological codes in ASR," Proc. Inst. Acoust. **8**, 581–588.

Meddis, R. (**1986b**). "Simulation of mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **79**, 702–711.

Meddis, R. (**1988**). "Simulation of auditory-neural transduction: Further studies," J. Acoust. Soc. Am. **83**, 1056–1063.

Meddis, R., Hewitt, M. J., and Shackleton, T. M. (**1990**). "Implementation details of a computational model of the inner hair-cell/auditory-nerve synapse," J. Acoust. Soc. Am. **87**, 1813–1816.

Oono, Y., and Sujaku, Y. (**1974**). "A probabilistic model for discharge patterns of auditory nerve fibers," Trans. Inst. Elect. Comm. Eng. (Japan) **57**, 35–36.

Oono, Y., and Sujaku, Y. (**1975**). "A model for automatic gain control observed in the firing of primary auditory neurons," Trans. Inst. Elect. Comm. Eng. **58**, 61–62.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear-nerve of the guinea pig and its relation to the receptor potential of inner

hair cells," Hear. Res. **24**, 1–15.
Payton, K. L. (**1988**). "Vowel processing by a model of the auditory periphery: A comparison to eight-nerve responses," J. Acoust. Soc. Am. **83**, 145–162.
Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locking response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769–793.
Rose, J. E., Hind, J. E., Anderson, D. J., and Brugge, J. F. (**1971**). "Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey," J. Neurophysiol. **34**, 685–699.
Ross, S. (**1982**). "A model of the hair cell-primary fiber complex," J. Acoust. Soc. Am. **71**, 926–941.
Russell, I. J., and Sellick, P. M. (**1978**). "Intracellular studies of hair cells in the mammalian cochlea," J. Physiol. (Lond.) **264**, 261–290.
Schroeder, M. R., and Hall, J. L. (**1974**). "Model for mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **55**, 1055–1060.
Schwid, H. A., and Geisler, C. D. (**1982**). "Multiple reservoir model of neurotransmitter release by cochlear inner hair cell," J. Acoust. Soc. Am. **72**, 1435–1440.
Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (**1982**). "Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique," J. Acoust. Soc. Am. **72**, 131–141.
Siebert, W. M. (**1965**). "Some implications of the stochastic behavior of primary auditory neurons," Kybernetic **2**, 206–215.
Smith, R. L. (**1977**). "Short-term adaptation in single auditory-nerve fibers: some post-stimulatory effects," J. Neurophysiol. **40**, 1098–1112.
Smith, R. L. (**1988**). "Encoding of sound intensity by auditory neurons," in *Neurobiological Basis of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 243–274.
Smith, R. L., and Brachman, M. L. (**1982**). "Adaptation in auditory nerve fibers: A revised model," Biol. Cybern. **44**, 107–129.
Smith, R. L., Brachman, M. L., and Goodman, D. A. (**1983**). "Intensity functions and dynamic responses from the cochlea to the cochlear nucleus," in *Hearing—Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Springer-Verlag, New York), pp. 112–118.
Smith, R. L., Brachman, M. L., and Frisina, R. D. (**1985**). "Sensitivity of auditory nerve fibers to changes in intensity: A dichotomy between decrements and increments," J. Acoust. Soc. Am. **78**, 1310–1316.
Smith, R. L., and Zwislocki, J. J. (**1975**). "Short-term adaptation and incremental responses of single auditory-nerve fibers," Biol. Cybernet. **17**, 169–182.
Weiss, T. F. (**1966**). "A model of the peripheral auditory system," Kybernetic **3**, 153–157.
Westerman, L. A. (**1985**). "Adaptation and recovery of auditory nerve responses," Special Rep. ISR-S-24. Institute for Sensory Research, Syracuse University, Syracuse, NY.
Westerman, L. A., and Smith, R. L. (**1984**). "Rapid and short-term adaptation in auditory nerve responses," Hear. Res. **15**, 249–260.
Westerman, L. A., and Smith, R. L. (**1988**). "A diffusion model of the transient response of the cochlear inner hair cell synapse," J. Acoust. Soc. Am. **83**, 2266–2276.
Yates, G. K., and Robertson, D. (**1980**). "Very rapid adaptation in auditory ganglion cell," in *Psychophysical, Physiological and Behavioral Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft U. P., Delft, The Netherlands).
Yates, G. K., Robertson, D., and Johnstone, B. M. (**1985**). "Very rapid adaptation in the guinea pig auditory nerve," Hear. Res. **17**, 1–12.
Yates, G. K., Winter, I. M., and Robertson, D. (**1990**). "Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range," Hear. Res. **45**, 203–222.
Zwislocki, J. J. (**1973**). "On intensity characteristics of sensory receptors: A generalized function," Kybernetic **12**, 169–183.