Enhanced coding in a cochlear-implant model using additive noise: Aperiodic stochastic resonance with tuning

Robert P. Morse*
Mackay Institute of Communication and Neuroscience, School of Life Sciences, Keele University, Keele ST5 5BG, United Kingdom

Peter Roper†
Laboratory of Computational Neuroscience, Babraham Institute, Babraham CB2 4AT, United Kingdom
(Received 18 October 1999; revised manuscript received 25 January 2000)

Analog electrical stimulation of the cochlear nerve (the nerve of hearing) by a cochlear implant is an effective method of providing functional hearing to profoundly deaf people. Recent physiological and computational experiments have shown that analog cochlear implants are unlikely to convey certain speech cues by the temporal pattern of evoked nerve discharges. However, these experiments have also shown that the optimal addition of noise to cochlear implant signals can enhance the temporal representation of speech cues [R. P. Morse and E. F. Evans, Nature Medicine **2**, 928 (1996)]. We present a simple model to explain this enhancement of temporal representation. Our model derives from a rate equation for the mean threshold-crossing rate of an infinite set of parallel discriminators (level-crossing detectors); a system that well describes the time coding of information by a set of nerve fibers. Our results show that the optimal transfer of information occurs when the threshold level of each discriminator is equal to the root-mean-square noise level. The optimal transfer of information by a cochlear implant is therefore expected to occur when the internal root-mean-square noise level of each stimulated fiber is approximately equal to the nerve threshold. When interpreted within the framework of aperiodic stochastic resonance, our results indicate therefore that for an infinite array of discriminators, a *tuning* of the noise is still necessary for optimal performance. This is in contrast to previous results [Collins, Chow, and Imhoff, Nature **376**, 236 (1995); Chialvo, Longtin, and Müller-Gerking, Phys. Rev. E **55**, 1798 (1997)] on arrays of FitzHugh-Nagumo neurons.

PACS number(s): 43.72.+q, 87.10.+e, 87.19.Dd, 02.50.−r

I. INTRODUCTION

The auditory mechanisms of the inner ear, such as the hair cells, are physiologically vulnerable and can be severely damaged by disease, by noise exposure, or by the side effects of some pharmaceuticals [1–5]. Severe loss of hair cells leads to profound deafness and is associated with a partial loss of the innervating cochlear nerve fibers [5–7]. Functional hearing can, however, be partially restored by direct electrical stimulation of the surviving cochlear nerve fibers [8–11], and this can be achieved with a cochlear implant. A cochlear implant consists of three parts: an externally worn sound processor (or ''speech processor''), a set of surgically implanted electrodes in the inner ear (or cochlea), and a means of transmitting signals across the skin to the electrodes (for reviews, see Refs. [12–15]). The sound processor converts audio signals into ''appropriate'' electrical currents that are then transmitted to one or more electrodes situated near the cochlear nerve fibers. Even though nerve fibers in a deafened ear will have partially degenerated, a sound processor must somehow code speech information into a form that is usable by the brain.

Speech consists of sequences of acoustic features that must be identified to allow comprehension of an utterance. During an utterance, the amplitude of the sound rises and falls and, for some components of speech, the vocal chords vibrate [16]. The fundamental frequency of vocal chord vibration is perceived as the voice pitch and its variation provides syntactic information. Although the amplitude and pitch cues alone provide substantial information, a listener must determine the finer spectral characteristics of an utterance to understand it. These finer spectral characteristics are evident in the amplitude spectrum of a voiced speech component as harmonics at multiples of the voice fundamental. The peaks in the envelope of an amplitude spectrum, which result from vocal-tract resonances, are known as formants. The frequencies of the formants are used to categorize vowels [17] and their frequency transitions are used to categorize some consonants [18]. In the normal ear, spectral speech cues are probably coded both in terms of which nerve fibres are most active, i.e., place coding, and also by the temporal discharge patterns of each fiber, i.e., time coding (for reviews, see Refs. [19–21]).

The appropriate coding strategy for a cochlear implant depends on whether it is primarily intended to aid lip reading, to convey only speech cues by audition alone, or to enable discrimination of all sounds including environmental sounds (e.g., door bells). Furthermore, since the performance of cochlear implants depends on many factors (including the amount of cochlear-nerve degeneration), it is now evident that a single processing strategy will not be suitable for all cochlear implantees [22–25]

In cochlear implants intended to restore functional hearing of all sounds, rather than just speech signals, the input signal is filtered by a bank of up to about 20 bandpass filters

*Electronic address: r.p.morse@cns.keele.ac.uk
†Electronic address: pr230@cam.ac.uk

 ©2000 The American Physical Society

that crudely mimic the frequency mapping in the normal ear [26–28,22,29]. The output of each filter determines the stimulus current of an electrode in accordance with the normal tonotopicity of the cochlea: high frequencies are used to stimulate nerve fibers at the base of the cochlea, and low frequencies are used to stimulate apical fibers.

In some non-speech-specific strategies [28,30,29], it is assumed that high frequency temporal information is not usable by implantees, and formants above 400 Hz are coded only by the place of stimulation. However, in other non-speech-specific strategies [26,31,27,23], formants are intended to be coded by both place *and* time cues. The compressed analog filter outputs are used to stimulate segments of the cochlear nerve, and the formant information in each channel is presumed to be retained in the pattern of evoked nerve discharges. Our recent experiments using a toad sciatic nerve as a physiological model of the cochlear nerve [32] and simulations using the Frankenhauser-Huxley nerve model [33] have shown that conventional analog cochlear implants are unlikely to convey formant information by time coding (as seen in amplitude spectra of the nerve discharges). Under the assumption that high frequency temporal information is usable by implantees, we advocated that noise should be deliberately added to cochlear implant signals, and we showed that formant representation by time coding was enhanced by the optimum addition of noise to the outputs of the cochlear implant filters. It is the aim of the present work to indicate how this might occur.

Noise-enhanced transmission of information is not a new phenomenon, and has frequently been investigated under the generic umbrella of stochastic resonance (SR) (see e.g. [34–37]). Typical SR studies focus upon the transduction of subthreshold signals that have a fixed amplitude and are periodic. However, speech has neither of these characteristics: the signal amplitude varies continually throughout an utterance, and furthermore even though each formant corresponds to a narrow frequency band [56], an utterance such as a vowel will typically comprise more than one formant as well as finer spectral information.

Collins, Chow, and Imhoff [38] demonstrated that a single noisy, excitable unit can optimally detect a slowly varying, aperiodic signal, and they introduced the normalized power norm C, a measure that characterizes signal detection. The normalized power norm exhibits a nonmonotonic and unimodal dependence upon the noise strength, and the optimization (by tuning of the noise strength) of this measure was called aperiodic stochastic resonance (ASR). C is the normalized correlation between the input and output signals and there is therefore an optimal noise strength for which this correlation is maximal. However, this optimal noise depends upon the mean and variance of the signal, and so must be actively modulated when the unit is presented with a signal for which these parameters vary. Subsequent work by the same authors [39] shows that the summed output of a parallel array of noisy FitzHugh-Nagumo (FHN) neurons also exhibits ASR when each unit is presented with the same aperiodic input signal but each is subject to a different noise source. The authors of Ref. [38] further suggest that the dependence upon noise strength of the normalized power norm asymptotically approaches a plateau near 1.0 as a function of increasing array size (see, e.g., Ref. [39] Fig. 2). They thus concluded that ASR for a large array does not require a *tuning* of the noise, but instead that a fixed noise level can optimally transduce an aperiodic signal. Further analysis in Ref. [40] showed that ASR with slowly varying signals can be interpreted as a noise-induced linearization of the unit's transfer function, and thus may be considered as a special case of dither [41,42]. Such a transfer function for the FHN array cannot be analytically determined, but must instead be found numerically. By replacing the transfer function with a linear ansatz, the authors showed that they could reliably predict the optimal normalized power-norm for the FHN array.

In this paper we model the set of nerve fibres that comprise the cochlear nerve as an infinitely large parallel array of thresholding devices (discriminators). We first confirm that ASR for this system derives from a noise induced linearization of the transfer function [40]. However, in contrast to Ref. [40], an analytic expression for the transfer function of our network may be found. Consequently we do not assume a linear ansatz but can instead show that the transfer function has minimal curvature (i.e. it is approximately linear over some region) for some optimal noise, say $\sigma=\sigma_c$, but is convex for $\sigma<\sigma_c$ and concave for $\sigma<\sigma_c$. We subsequently show that the normalized power norm for this array does not possess a plateau (in contrast to the findings of Ref. [39]), but instead has a well-defined maximum at some critical noise intensity, σ_c. However, it should be noted that the elements of our discriminator array each possess an explicit threshold, which is in contrast to the parallel FHN array previously studied [40,39]. We determine σ_c for the infinite discriminator array, and we compare our prediction with numerical simulations of a large array. Finally, we investigate the effect of stimulus level on the correlation between the summed outputs and the input waveform.

Nerve models more complex than simple discriminators are more common in studies of neural behavior. Such models are used to investigate such phenomena as, e.g., strength-duration compensation, refractoriness, and accommodation. However, here we are primarily concerned with the temporal pattern of nerve activity, and we have previously [43] demonstrated that, in response to simulated cochlear implant stimulation, the temporal response of a discriminator is strikingly similar to both that of real neurons (the toad sciatic nerve) and also complex model neurons (the Frankenhauser-Huxley model). Our findings should therefore have direct relevance to the choice of noise level for cochlear-implant coding. We further suggest that our model provides insight into the cause of improved time coding by the addition of noise to cochlear implant stimuli.

Nevertheless, while the temporal response of a nerve can be well predicted by a system with an explicit threshold alone, more detailed threshold characteristics and secondary nerve properties (such as membrane filtering, refractoriness, and accommodation) must be considered if more complex nerve behavior is to be investigated.

II. APERIODIC STOCHASTIC RESONANCE IN AN ARRAY OF THRESHOLDING DEVICES

The temporal response to a filtered vowel (with or without noise) of both the sciatic nerve and the Frankenhauser-Huxley nerve model can be well modeled as a discriminator

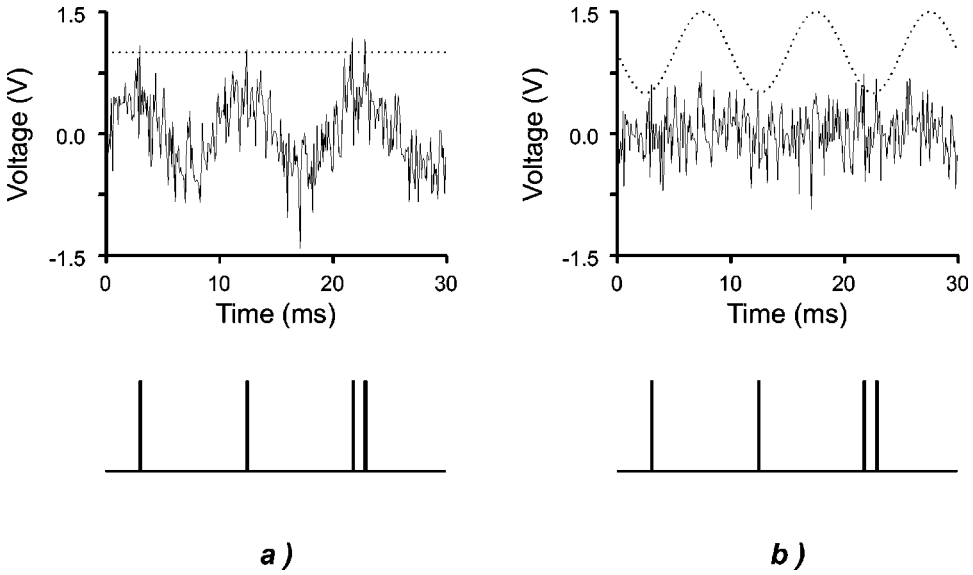

FIG. 1. Thresholding operation of a discriminator: (a) a subthreshold sinusoidal signal with additive noise produces a pulse (lower figure) whenever the stimulus crosses the threshold (dashed line upper figure) in the increasing direction; (b) a variable stimulus [such as that in (a)] may be reinterpreted as a pure noise input to a discriminator with a variable threshold.

(or level-crossing detector) [43]. Such a discriminator is a thresholding device that produces a narrow voltage pulse, whenever its input signal $S(t)$ exceeds some threshold value, say Δ. The output $R(t)$ of a single discriminator therefore consists of a time series of voltage pulses. For a subthreshold signal with additive noise, crossings of the threshold are more probable when the signal is high than when it is low [Fig. 1(a)], and this is reflected in the corresponding instantaneous pulse frequency of $R(t)$.

A convenient model to study time-coding by a set of N parallel nerve fibers is a parallel array of N such discriminators. We compute the mean activity of the array at a time t by

$$R_\Sigma(t)=\frac{1}{N}\sum_{i=1}^{N} R_i(t), \tag{1}$$

where $R_i(t)$ is the response of the ith discriminator.

Following Ref. [39], we introduce the normalized power norm C, to characterize the response of the array

$$C=\frac{\overline{S(t)R_\Sigma(t)}}{[\overline{S(t)^2}]^{1/2}[\overline{(R_\Sigma(t)^2-\overline{R_\Sigma(t)})^2}]^{1/2}}, \tag{2}$$

where the overbar denotes a temporal average.

It was previously found [43] that if each of a set of parallel discriminators are given the same input signal (a filtered-vowel stimulus, representing the output of a single channel of an implant), then the addition of an independent additive-noise source to each discriminator can increase the correlation, C. Furthermore, as shown in Fig. 2, with many discriminators and a subthreshold input it is possible to obtain an almost perfect correlation between the summed outputs and the input when the noise added to each discriminator input is above a *minimum* level (see also Refs. [40], [39]).

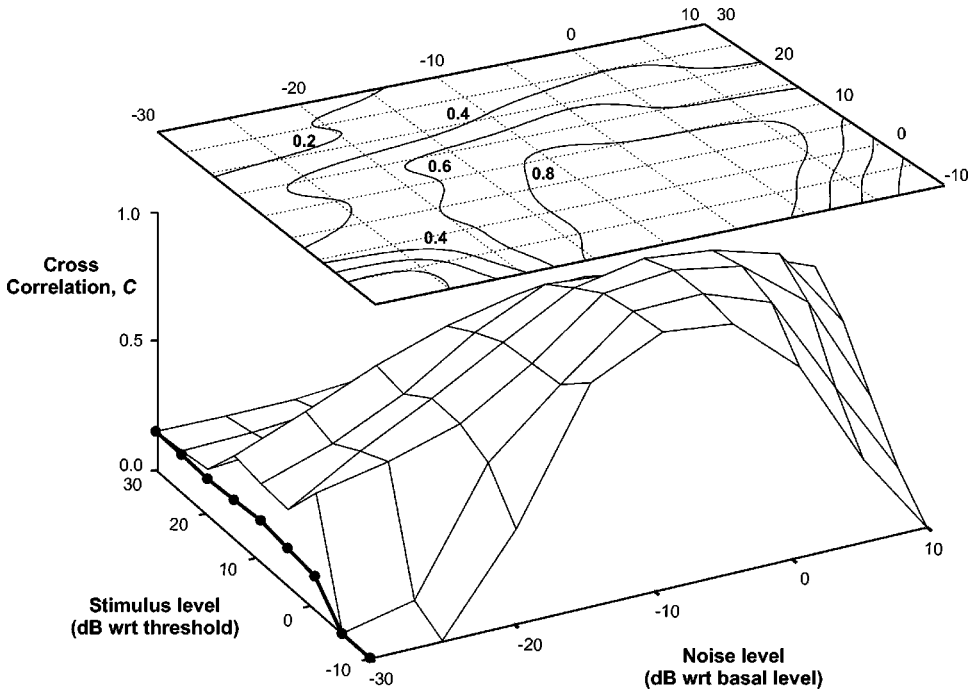

FIG. 2. Measured cross-correlation between the input and output of a 1000-element discriminator array, in response to the filtered vowel /æ/ (after filtering by the first channel of the simulated cochlear implant; see Fig. 3). The absolute filtered-vowel threshold was 0.39 V_{rms} and the base noise level was $\sqrt{5}$ V_{rms}. Note the contour plot above the main figure, which clearly shows that there is a finite noise for optimal correlation for this large but finite array.

In other words, with optimal conditions, information in a filtered-vowel stimulus can be faithfully represented by the temporal pattern of output pulses from the discriminators. This therefore suggests that it may be possible to enhance the temporal coding of speech cues by stimulating many cochlear nerve fibers with the same subthreshold information-bearing signal, but also adding to each a different noise waveform. However, for the faithful transmission of information, the noise level added to each nerve fibre must be above some minimum value [39]. The authors of Ref. [29] also conjectured that the cross correlation measure C for an array of FitzHugh-Nagumo elements would asymptotically approach a plateau near 1.0 in the limit that the array size tended to infinity. Despite this, Fig. 2, which comprises the output of 1000 discriminators, clearly shows that for this array, C possesses a maxima.

It is interesting to investigate whether the conjecture of Ref. [39] will hold true for an infinite array of this type. We may estimate the power norm C for an infinitely large discriminator array by using an extension to the Rice rate equation [44,45] that is due to Ref. [46] (see also Ref. [35]). Rice's equation predicts the ensemble-average of the threshold-crossing rate due to noise, ν, of an infinite number of identical discriminators, each subject to independent noise inputs (but such that each noise source is drawn from the same band-limited Gaussian distribution)

$$\langle \nu \rangle = \frac{B}{\sqrt{3}} \exp\left(-\frac{\Delta^2}{2\sigma^2} \right), \tag{3}$$

where $\langle\ \rangle$ denotes the ensemble mean, Δ the threshold level, σ the root-mean-square (rms) noise level, and B the noise bandwidth of each discriminator. Following Ref. [46], an instantaneous change in stimulus-amplitude is equivalent to a change in the threshold of each discriminator [Fig. 1(b)], and so Eq. (3) can be extended to predict the mean threshold-crossing rate of the discriminator system at each instant of some stimulus $S(t)$, such that

$$\langle \nu(t) \rangle = \frac{B}{\sqrt{3}} \exp\left(-\frac{(\Delta - S(t))^2}{2\sigma^2} \right), \tag{4}$$

and since this is an ensemble mean, the mean crossing rate $\langle \nu(t) \rangle$ may be identified with the mean activity of the array, $R_\Sigma(t)$.

III. OPTIMAL DETECTION OF SPEECH

The five English vowels /*a*/, /i/, /u/, /æ/ and /ə/ (as in ''cart,'' ''seat,'' ''rude,'' ''hat,'' and ''the'') were synthesized with a software implementation of a cascade formant synthesizer [47], for use as an input signal to the discriminator array. These synthetic vowels, which had a sample period of 25 μs and a fundamental frequency of 100 Hz, were then digitally filtered by eighth-order bandpass filters (such filters are identical to those used in an experimental cochlear implant [27]). Thus five steady-state filtered vowels, each corresponding to one channel output, were generated for each vowel (see Fig. 3). These filtered vowels were identical to those used in discriminator simulations [43] and, except for the lower sampling frequency, identical to those previously

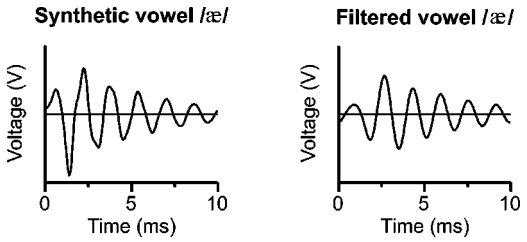

FIG. 3. One period of a synthetic vowel (left) and a filtered vowel (right): filtering was between 200 Hz and 671 Hz, and corresponds to the first channel of a cochlear implant.

used in sciatic nerve experiments and Frankenhauser-Huxley simulations [32,33]. root-mean-square stimulus amplitudes ranged from 10 dB below the stimulus threshold to 20 dB above it, in steps of 5 dB. Noise levels ranged from 30 dB below a base noise level to 10 dB above in steps of 5 dB. To maintain consistency with the discriminator simulations [43], the threshold level for our model was chosen to be 1 V and the noise bandwidth B, was 20 kHz; thus the stimulus thresholds and base noise level ($\sqrt{5}\ V_{rms}$) were also identical to those used in our previous work.

For every permutation of filtered-vowel waveform, filtered-vowel amplitude, and noise level, Eq. (4) was used to predict a time series of mean threshold-crossing rates for an infinite discriminator array in response to 4000 consecutive samples (i.e., 100 ms) of the filtered-vowel stimulus. The resulting normalized power-norm, C, for the array was then calculated. Figure 4 shows how C for the filtered vowel /æ/ varies with stimulus amplitude and noise level, and it is clear that for stimulus amplitudes between $\sim$ 10 dB below the stimulus threshold and $\sim$15 dB above it, our model predicts that C should be enhanced by the addition of noise. C *appears* to plateau when the noise added to each discriminator input exceeds a certain level (which is about 10 dB below the base noise level for the results shown in Fig. 4), and this accords with the findings of Refs. [39] and [40] for a large, parallel FHN array. For vowel stimuli below or near threshold, an almost perfect correlation between the input and output spectra is obtained when the noise level of each discriminator is above this critical level. With lower amplitude vowel stimuli, the minimum noise level required to obtain a given cross-correlation is lower.

To understand this noise-assisted enhancement of C, consider first the noise to be internal to each discriminator and the input to be noise free. When there is no stimulus present, the array exhibits a mean threshold-crossing rate $\langle \nu \rangle$ given by Eq. (3), which is determined by both the threshold and by the noise level of each discriminator (Fig. 5). Thus when the array is subject to some time-varying signal, its output may be considered as the difference between the mean crossing rate at any instant, $\langle \nu(t) \rangle$, and the mean crossing rate in the absence of a signal, $\langle \nu \rangle$. Consequently, using ideas from communication theory [48], we may consider the information-bearing input to the array [Fig. 5(c)] as modulating the threshold of each discriminator about an operating point [Fig. 5(a)]. The relation between the instantaneous stimulus-level and the mean threshold-crossing rate [Eq. (4)] therefore defines a *modulation transfer function* (MTF). A related transfer function for the FHN array was numerically computed in Ref. [40] (see, e.g., Sec. IV and Fig. 7 of that

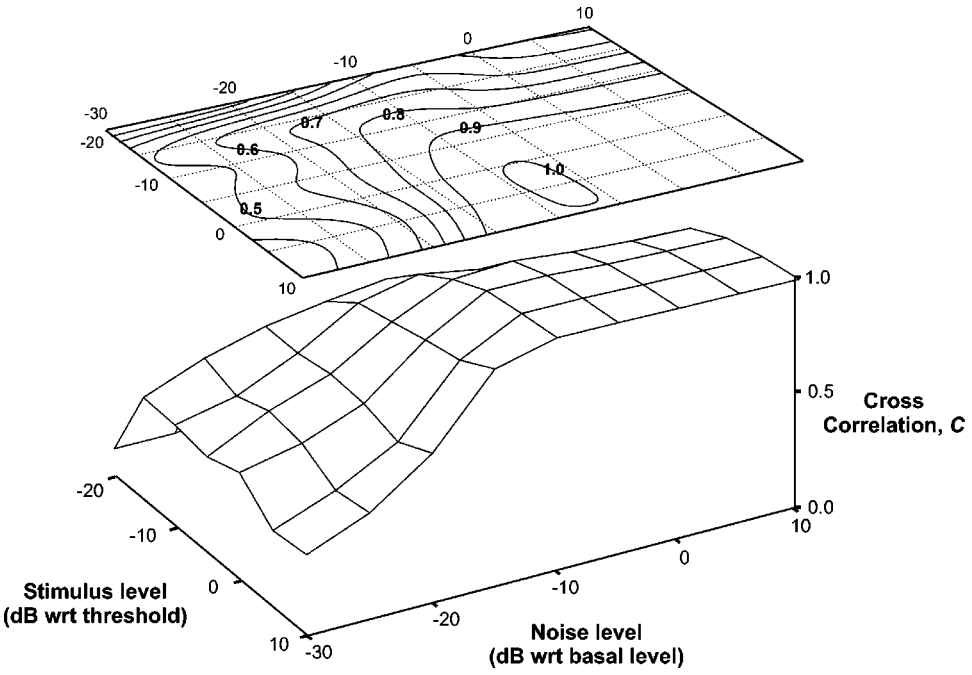

FIG. 4. Predicted cross-correlation between the input and output of an infinite discriminator array, in response to the filtered vowel /æ/ (after filtering by the first channel of the simulated cochlear implant; see Fig. 3). The absolute filtered-vowel threshold was 0.39 V_{rms} and the base noise level was $\sqrt{5}$ V_{rms}. Note the contour plot above the main figure which clearly shows that there is a finite noise for optimal correlation.

paper). The peak-to-peak amplitude of the input signal defines a interval about the operating point and so fixes the region of interest of the MTF. The form of the MTF over this interval [shown, for example, by the highlighted part of Fig. 5(a)] reflects the similarity between the input and the output of the array, and so determines the power norm C. For example, if it were linear over this region, then some change in the stimulus amplitude would induce a proportional change in $\langle \nu(t) \rangle$, and hence the correlation between the input and the output would be perfect (i.e., $C=1$).

It is instructive to examine how increasing the noise varies the shape of the MTF over the interval about the operating point. At low noise levels, the operating point lies toward the tail of the MTF (curve **L**, Fig. 5) in a region where the transfer function is concave. However, when the noise is increased, the operating point moves away from the tail and toward a convex region of the MTF (curve **H**, Fig. 5). Normalized [57] plots (Fig. 6) of the transfer function for various noise strengths show this situation more clearly: for low noise levels the MTF is concave, while for high noise it is convex. Furthermore, in the limit $\sigma \to \infty$, the MTF asymptotically approaches a fixed convex function. In Ref. [40] the transfer function was replaced with a linear ansatz; however, for our array the MTF can never be truly linear except in some infinitesimal region about the operating point. Optimal C is instead achieved when the curvature of the MTF over the interval is minimized, which occurs when the length, l, is minimal, where

$$l = \int_b^a d\Delta \left[1 + \left(\frac{d}{d\Delta} \langle \nu \rangle \right)^2 \right]^{1/2}, \tag{5}$$

and $[a,b]$ delimit the interval. Application of the Euler equation shows that minimal l, and hence optimal C, occur when $\sigma = \Delta$.

For optimal noise (i.e., $\sigma = \Delta$), the curvature of the MTF vanishes at the operating point. Therefore, the degree to which the linear ansatz fails at this time depends upon the size of the interval, and hence upon the signal amplitude $A = 1/2[S(t)_{max} - S(t)_{min}]$. For very small A the ansatz is valid and $C \to 1$; however, as A increases the approximation becomes worse and C falls below unity. This implies that the transduction quality of the network diminishes with signal amplitude. This effect is best illustrated by calculating the cross-correlation between one period of an input sine wave, $A \sin(\omega t)$, and the corresponding time series of mean threshold-crossing rates. The amplitude A of the input determines the size of the interval (Fig. 5), and the effect of discriminator noise level on the power norm, C, is shown in Fig. 7 for three values of A. It is clear that the network transducers larger amplitude signals less well.

A second, more important, point to note is that for this network the noise dependence of the power norm, C, does not in fact plateau, but instead has a well-defined maximum at $\sigma = \Delta$ [see Fig. 7(b) and the contour plot in Fig. 4]. This is in contrast to the findings of Refs. [39] and [40] for the FHN model, and implies that for optimal transduction of a variable amplitude signal, a *tuning* of the noise is still necessary for this array (cf. Ref. [39]). However, for small amplitude signals, the decrease of C with increasing noise is small [typically less than $\sim$3%; see, e.g., Fig. 7(b)] and so might be ignored for practical purposes.

IV. SUPRATHRESHOLD SIGNALS

Our prediction of C for subthreshold vowel stimulation agrees well with numerical results from a discriminator system; we also obtained good agreement for suprathreshold signals with amplitudes less than $\sim$10 dB above threshold [43]. However, the accuracy of our predictions for suprathreshold stimuli noticeably diminishes for higher amplitude (i.e., $>$10 dB) signals [43]; for example recall the large stimuli regions of Figs. 2 and 4.

This degradation with stronger suprathreshold stimuli should be expected since Rice's rate equation (4) depends only upon the absolute magnitude of the distance between

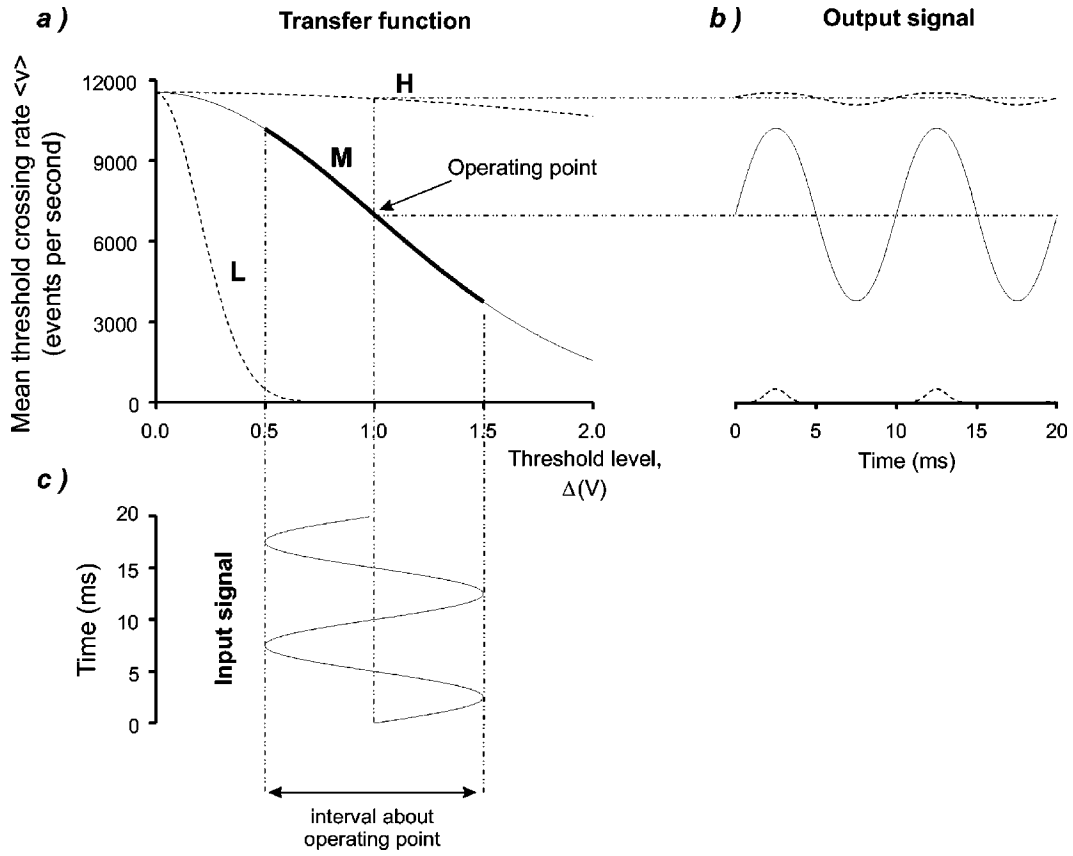

FIG. 5. A discriminator array subject to a noise of rms level σ, and bandwidth B, has a modulation transfer function (MTF) given by Eq. (3). The noise-free input signal to the system may be interpreted as modulating the threshold level of each discriminator about some mean level: the operating point. The operating point is equal to the mean threshold crossing rate and so is determined by σ and by the threshold, Δ, of each discriminator. The input to the array [figure (c)] is a 100 Hz sine wave of amplitude $A=0.5$ V. (a) shows MTF's for noise levels: $\sigma=0.2\ V_{rms}$ (**L**), $\sigma=1.0\ V_{rms}$ (**M**) and $\sigma=5.0\ V_{rms}$ (**H**). The highlighted part of the $\sigma=1.0$ MTF shows the interval about the operating point, and optimal C is achieved when the length of this interval is minimal. (b) shows the output of the array. Each discriminator has threshold $\Delta=1$ V and noise bandwidth $B=20$ kHz.

signal and threshold. Therefore, if the signal is allowed to be suprathreshold, two values of instantaneous stimulus amplitude (one above threshold and the other equally far below) give the same mean threshold-crossing rate. Rice's equation (4) is therefore symmetric with respect to the threshold. However real discriminators (and real neurons) are *asymmetric* with respect to threshold: although the rising phase of a suprathreshold stimulation will always trigger at least one pulse as the signal (plus noise) exceeds threshold, the threshold may not be crossed in the positive direction during the

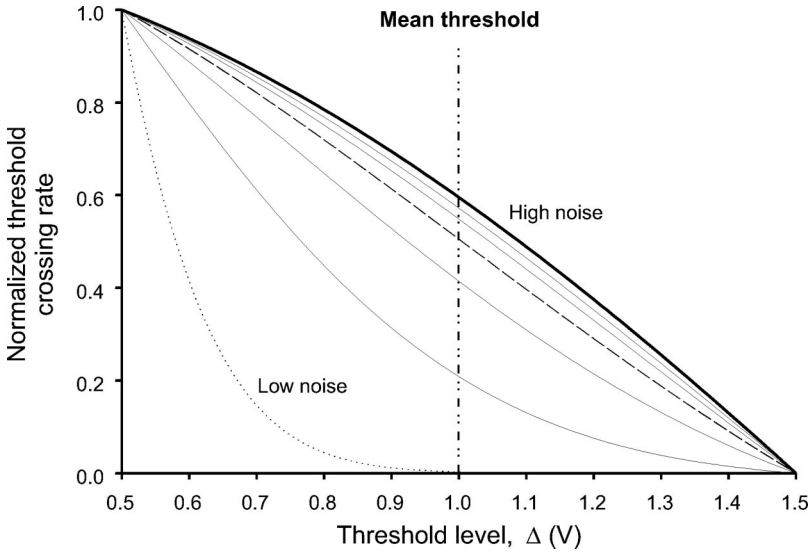

FIG. 6. Normalized (i.e., such that the rates lie between 0 and 1 over the range of threshold levels shown) plots of the transfer function for various noise strengths. For low noise levels (dotted line) the MTF is concave, while at high noise levels (bold line) it is convex. Intermediate noise levels (dashed line) best approximate a linear function, and in this regime the correlation C is optimal.

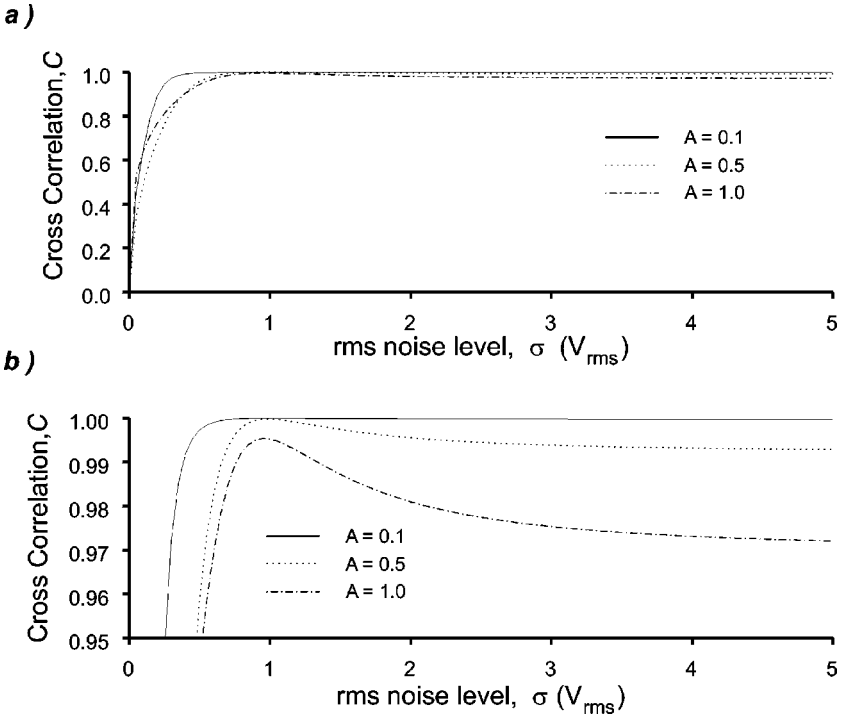

FIG. 7. Predicted cross-correlation between the input and output of an infinite discriminator array, in response to three different amplitude sine waves, $A\sin(\omega t)$. The mean threshold is $\Delta = 1.0\ V_{rms}$. Figure (b) is a magnification of figure (a) over the range $C \in [0.95,1.0]$, and clearly shows that there is an optimal noise at $\sigma = \Delta$, and furthermore that C is a function of stimulus amplitude.

falling phase of the signal. Thus Rice's equation will overestimate the threshold crossing rate. Furthermore, this asymmetry of real discriminators means that the pulse rate for a suprathreshold stimulus will also be dependent upon the stimulus gradient. Signals that change slowly in the vicinity of threshold will be more likely to undergo crossings and recrossings of the threshold, and subsequently will produce more spikes than will signals that vary more sharply at threshold.

A further consequence of the symmetry of Eq. (4) is that the response to a suprathreshold stimulus is a nonmonotonic function of instantaneous amplitude: e.g., the response to the peak of a suprathreshold signal will be a minimum in the mean threshold-crossing rate, even though the instantaneous stimulus amplitude is maximal (Fig. 8). This is therefore another source of poor correlation between the time series of mean threshold-crossing rates and the suprathreshold stimulus. In contrast, if the input is constrained to be subthreshold, there is a one-to-one relationship between the instantaneous stimulus amplitude and the mean threshold-crossing rate. Thus, for subthreshold signals, if the threshold and bandwidth of each discriminator are known, then reconstruction of the input signal is completely determined by the time series of mean crossing rates. However, in practice, the mean threshold-crossing rates must be estimated, and so signal information may be lost.

V. FINITE SIZE EFFECTS

Our model predicts the effects of stimulus amplitude and of noise strength on the output of a infinitely large array. However, real cochlear implants are finite in size, and so it is important to consider how reducing the array affects our predictions. For small-amplitude, subthreshold signals, we have shown that large discriminator array exhibits almost perfect correlation when operating in a high noise regime. However, the observed correlation for a small (e.g., < 1000 elements) array is markedly lower (Fig. 2 and Ref. [43]). Such a failure of our model is a straightforward consequence of the failure of the central limit theorem for such a small array size, and shows that for optimal performance we need to combine the outputs of many parallel discriminators.

VI. DISCUSSION

The effect of noise level and interval size on stimulus amplitude on the curvature of the MTF close to the operating point accounts for the cross-correlation profile calculated for filtered vowels in background noise (see, e.g., Fig. 4). The MTF was convex when the rms noise level of each discriminator was much lower than the optimal noise σ_c. The nonlinear modulation therefore emphasized the maxima of a stimulus compared with the minima. When a high amplitude (but subthreshold) filtered vowel is presented, only the largest peaks cause a noticeable change in the mean threshold-crossing rate of the array. Thus the normalized amplitude spectrum of the time series of mean threshold-crossing rates is dominated by the vowel fundamental and by its harmonics, and the cross-correlation between the output and input of the discriminator array is poor. Increased noise reduces the curvature of the MTF close to the operating point, which therefore increases the cross-correlation C. With a larger stimulus amplitude, more noise is required to minimize the MTF curvature.

We have previously shown [43] that the response to filtered vowels of the discriminator array and the sciatic nerve are similar. It is therefore likely that the modulation mechanism by which noise leads to improved time coding in a discriminator system may also account for the improved coding in the nerve data. The addition of noise to some suprathreshold nerve stimuli, however, caused the harmonic closest to a formant to be transmitted preferentially compared with proximal harmonics, and this result is not fully explained by the modulation process. If the operating point of each dis-

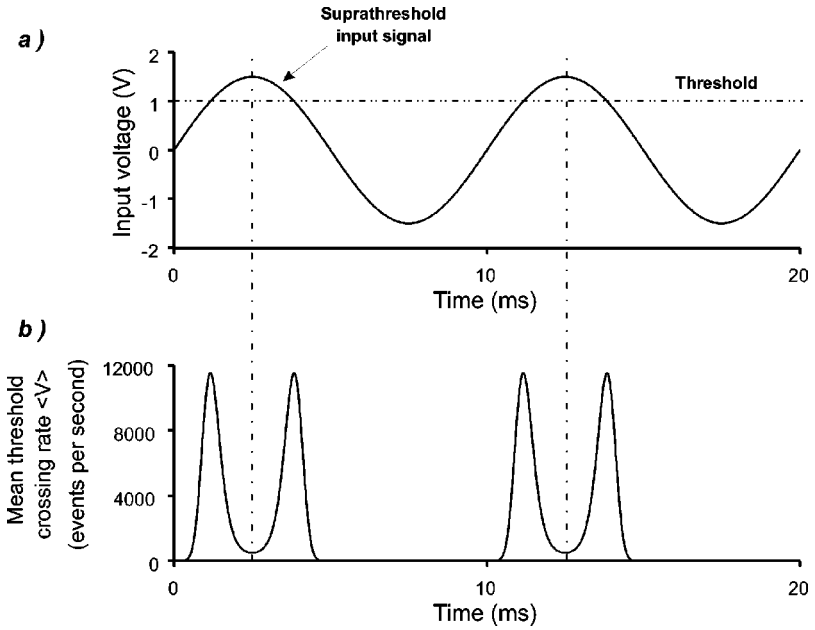

FIG. 8. The mean threshold crossing rate of a discriminator array subject to a suprathreshold signal. The upper figure shows the noise-free input to the array (a 100-Hz sine wave of amplitude 1.5 V), the lower one shows the discriminator array's response. The response to the peak of the signal is a minimum in the mean rate, even though the instantaneous stimulus amplitude is maximal. The mean noise level of each discriminator is $\sigma = 1\ V_{rms}$ and the threshold is $\Delta = 1$ V.

criminator is near the point of zero curvature of the MTF (Fig. 5, curve **M**) then, because of the sigmoidal shape of the MTF, the time series of mean threshold-crossing rates for the discriminator system will be a compressed form of the information-bearing input. This compression probably accounts for the limited degree of preferential transmission observed in the computational discriminator simulations [43], and may account for some of the preferential transmission observed in the sciatic nerve experiments. As noted previously [43], the additional preferential transmission may be the result of the compressive nonlinearity associated with an inherent neural threshold.

Previous studies of cochlear implant coding using additive noise concerned only the benefits of adding noise to vowel stimuli [32,33]. However, the modulation process described is not stimulus specific: additive noise should be expected to enhance the implant coding of all components of speech and even of any other sound. Our results suggest that, for optimal signal transduction, the internal rms noise level of each stimulated nerve fiber should be approximately equal to the nerve threshold. The electrode current required to obtain a given internal noise level will depend on the position of the electrodes and the distribution of surviving nerve fibers (e.g., Ref. [49]). However, it should be possible to gain some estimate of the required electrode currents for optimal operation, by means of psychophysical measurements of the nerve firing thresholds to sinusoidal stimulation.

People with a profound deafness have typically $\sim$15 000 surviving nerve fibers [5–7], and many cochlear implantees may therefore have a sufficient number of fibers per cochlear implant channel to represent speech cues by time coding reliably. To benefit from the addition of noise to cochlear implant signals, the noise waveforms used to stimulate each fiber should ideally be independent of each other. Recent computer simulations have shown that stimulation of each fiber by an independent noise current is largely obtainable if the spread of current through the conductive fluids of the cochlea is exploited ([50]). Our next step is to determine whether cochlear implantees can use the extra information transmitted by analog cochlear implants using additive noise.

Equation (4) predicts the mean threshold-crossing rate of a theoretical discriminator system without refractoriness. However, a real neuron produces a finite-width output pulse (an action potential) in response to a detected threshold crossing, and furthermore this pulse is generally followed by both an absolute and a relative refractory period. Threshold crossings that occur during either the output pulse or the absolute refractory period are not detected, and threshold events during the relative refractory period are less likely to produce a pulse. The firing rate of a real array of neurons may therefore be lower than that of the corresponding discriminator array. Therefore, to model a real system, its firing rate must be calculated from knowledge of the threshold-crossing rate, the pulse width, and the refractory parameters. The inverse problem has been studied, that is, estimation of the instantaneous threshold-crossing rate given the pulse rate (spike rate) and refractory parameters [51–53]. The study by Johnson and Swami [52] was based on extensions to renewal theory [54] that enable nonstationary point processes to be studied. It may therefore be possible to extend this model so that the pulse rate can be calculated from the threshold-crossing rate.

VII. CONCLUSIONS

Our preliminary study [55] of discriminator systems and a more detailed examination of neural systems (with implicit rather than explicit thresholds) in Ref. [40] have also focused on the linearizing effect of noise on the effective transfer function. Both studies demonstrated that the similarity between the summed outputs and the input of some parallel nonlinear systems can be increased by the addition of noise. However, in contrast to the analysis of Ref. [40], the MTF for our system was not assumed to be linear in the region about the operating point. Our study shows that, even for an infinite number of parallel discriminators, there is an optimal noise level for maximal correlation between a finite input signal and the summed output. The maximum correlation occurs when the curvature of the MTF in the region of the operating point is minimal. For a set of parallel discrimina-

tors, the least MTF curvature, and therefore maximum correlation, occurs when the rms noise level is equal to the threshold of each discriminator. With this in mind, it would be interesting to re-examine the FHN network (and other more tractable alternatives) to examine whether our results carry over to that domain, or whether the discriminator array represents a special case.

ACKNOWLEDGMENTS

The authors would like to thank Laura Doherty, Georg Meyer, and Ted Evans for their comments on earlier drafts of this manuscript; Ted Evans for the use of a computer during the initial stages of this investigation; and Nigel Stocks for many helpful discussions and comments during the preparation of this work.

[1] L. Bergstrom, Can. J. Otolaryngology **1**, 1 (1975).
[2] R. Hinojosa and J. R. Lindsay, Archives Otolaryngology - Head and Neck Surgery **106**, 193 (1980).
[3] A. Kerr and H. Schuknecht, Acta Oto-Laryngologica **65**, 586 (1968).
[4] N. Y.-S. Kiang, E. C. Moxon, and R. A. Levine, in *Sensorineural Hearing Loss*, edited by G. E. W. Wolstenholme and J. Knight (Churchill, London, 1970), pp. 241–268.
[5] J. Otte, H. F. Schucknecht, and A. G. Kerr, Laryngoscope **8**, 1231 (1978).
[6] R. Hinojosa and M. Marion, in *Cochlear Prostheses: an International Symposium*, edited by C. W. Parkins and S. W. Anderson (Annals of the New York Academy of Sciences, New York, 1983), Vol. 405, pp. 459–484.
[7] J. B. Nadol, Jr. and W.-Z. Xu, Ann. Otology, Rhinology and Laryngology **101**, 988 (1992).
[8] J. C. Ballantyne, E. F. Evans, and A. W. Morrison, J. Laryngology Otology Suppl. **1**, 1 (1978).
[9] R. C. Bilger, F. O. Black, N. T. Hopkinson, E. N. Myers, J. L. Payne, N. R. Stenson, A. Vega, and R. V. Wolf, Ann. Otology, Rhinology, Laryngology **86**, 1 (1977).
[10] B. J. Gantz, R. S. Tyler, J. F. Knutson, G. Woodworth, P. Abbas, B. F. McCabe, J. Hinrichs, N. Tye-Murray, C. Lansing, F. Kuk, and C. Brown, Laryngoscope **98**, 1100 (1988).
[11] R. S. Tyler, B. C. J. Moore, and F. Kuk, J. Speech Hear. Res. **32**, 887 (1989).
[12] B. Gantz, Archives Otolaryngology–Head and Neck Surgery **1**, 171 (1987).
[13] J. B. Millar, Y. C. Tong, and G. M. Clark, J. Speech Hear. Res. **27**, 280 (1984).
[14] B. C. J. Moore, *Cochlear Implants*, edited by R. F. Gray (Croom Helm, London, 1985), pp. 163–179.
[15] S. Rosen, in *Scott-Brown's Otolaryngology*, 6th ed., edited by D. Stephens (Butterworth Heinemann, Oxford, 1997), Vol. 2.
[16] F. B. Fry, *The Physics of Speech* (Cambridge University Press, Cambridge, 1979).
[17] P. Delattre, A. M. Liberman, F. S. Cooper, and L. J. Gerstman, Word **8**, 195 (1952).
[18] A. M. Liberman, P. C. Delattre, F. S. Cooper, and L. J. Gerstman, Psychological Monographs, **68**, 1 (1954).
[19] E. F. Evans, Audiology **17**, 369 (1978).
[20] W. M. Hartmann, J. Acoust. Soc. Am. **100**, 3491 (1996).
[21] E. Javel and J. B. Mott, Hear. Res. **34**, 275 (1988).
[22] M. M. Merzenich, in *Cochlear Implants*, edited by R. A. Schindler and M. M. Merzenich (Raven, New York, 1985), pp. 121–129.
[23] R. A. Schindler and D. K. Kessler, Am. J. Otology **14**, 263 (1993).
[24] Q. Summerfield, in *Cochlear Implants* (Ref. [25]), pp. 195–222.
[25] B. S. Wilson, R. A. Schindler, C. C. Finley, D. K. Kessler, D. T. Lawson, and R. D. Wolford, in *Cochlear Implants*: *Current Situation*, edited by P. Banfai (Springer, Berlin, 1987), pp. 395–427.
[26] D. K. Eddington, W. H. Dobelle, D. E. Brackmann, M. G. Mladejovsky, and J. L. Parkin, Annals Otology, Rhinology, Laryngology **87**, 5 (1978).
[27] E. F. Evans, IEE Dig. **179**, 2/1 (1991).
[28] H. J. McDermott, C. M. McKay, and A. E. Vandali, J. Acoust. Soc. Am. **91**, 3367 (1992).
[29] B. S. Wilson, C. C. Finley, D. T. Lawson, R. D. Wolford, D. K. Eddington, and W. M. Rabinowitz, Nature (London) **352**, 236 (1991).
[30] L. A. Whitford, P. M. Seligman, C. E. Everingham, T. Antognelli, M. C. Skok, R. D. Hollow, K. L. Plant, E. S. Gerin, S. J. Staller, H. J. McDermott, W. R. Gibson, and G. M. Clark, Acta Oto-Laryngologica **115**, 629 (1995).
[31] J. L. Parkin, G. A. McCandless, and J. Youngblood, in *Cochlear Implants: Current Situation* (Ref. [28]), pp. 429–461.
[32] R. P. Morse and E. F. Evans, Nature Medicine **2**, 928 (1996).
[33] R. P. Morse and E. F. Evans, Hear. Res. **133**, 107 (1999).
[34] F. Moss, in *Contemporary problems in Statistical Physics*, edited by G. H. Weiss (SIAM, Philadelphia, 1994), Chap. 5, pp. 205–253.
[35] F. Moss, D. Pierson, and D. O'Gorman, Int. J. Bifurcation Chaos Appl. Sci. Eng. **4**, 1383 (1994).
[36] K. Wiesenfeld and F. Moss, Nature (London) **373**, 33 (1995).
[37] K. Wiesenfeld and F. Jaramillo, Chaos **8**, 539 (1998).
[38] J. J. Collins, C. C. Chow, and T. T. Imhoff, Phys. Rev. E **52**, R3321 (1995).
[39] J. J. Collins, C. C. Chow, and T. T. Imhoff, Nature (London) **376**, 236 (1995).
[40] D. R. Chialvo, A. Longtin, and J. Müller-Gerking, Phys. Rev. E **55**, 1798 (1997).
[41] L. Gammaitoni, Phys. Rev. E **52**, 4691 (1995).
[42] R. A. Wannamaker, S. P. Lipshitz, and J. Vanderkooy, Phys. Rev. E **61**, 233 (2000).
[43] R. P. Morse and E. F. Evans, Hear. Res. **133**, 120 (1999).
[44] S. O. Rice, Bell Syst. Tech. J. **23**, 282 (1944).
[45] S. O. Rice, Bell Syst. Tech. J. **24**, 46 (1945).
[46] Z. Gingl, L. B. Kiss, and F. Moss, Nuovo Cimento D **17**, 795 (1995).
[47] D. H. Klatt, J. Acoust. Soc. Am. **67**, 971 (1980).
[48] M. Schwartz, *Information Transmission, Modulation and Noise* (McGraw Hill, New York, 1959).
[49] R. K. Shephard, S. Hatsushika, and G. M. Clark, Hear. Res. **66**, 108 (1993).

[50] R. P. Morse and G. F. Meyer, Chaos, Solitons, Fractals (to be published).
[51] P. R. Gray, Biophys. J. **7**, 759 (1967).
[52] D. H. Johnson and A. Swami, J. Acoust. Soc. Am. **74**, 493 (1983).
[53] R. P. Gaumond, D. O. Kim, and C. E. Molnar, J. Acoust. Soc. Am. **74**, 1392 (1983).
[54] D. R. Cox, *Renewal Theory* (Metheun, London 1962).
[55] R. P. Morse and E. F. Evans, B. J. Audiology **30**, 97 (1996).
[56] The peak of this band is referred to as the formant frequency.
[57] That is, such that the mean threshold crossing rate lies between 0 and 1.