Towards a better speech processor for cochlear implants: Auditory- nerve responses to high-rate electric pulse trains

By

Leonid Litvak

B.S. Electrical Engineering, Rice University, 1995
B.A. Applied Mathematics, Rice University, 1995

Submitted to the Department of Health, Science, and Technology in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
3/5/2002

Author……
Harvard-M.I.T. Division of Health Sciences and Technology
March 5, 2002

Certified by ……...
Bertrand Delgutte
Associate Professor of Otology and Laryngology, Harvard Medical School
Thesis Supervisor

Certified by ……...
Donald K. Eddington
Associate Professor of Otology and Laryngology, Harvard Medical School
Thesis Supervisor

Accepted by …...……...
Martha L. Gray
Edward Hood Taplin Professor of Medical and Electrical Engineering
Co-director, Harvard-M.I.T. Division of Health Sciences and Technology

Towards a better speech processor for cochlear implants: Auditory nerve responses to high-rate electric pulse trains

By

Leonid Litvak

Submitted to the Department of Health, Science, and Technology
on March 5, 2002
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Abstract

Cochlear implants are prosthetic devices that seek to restore hearing in profoundly deaf patients by electrically stimulating the auditory-nerve (AN). With current implants, the representation of the sound waveform in temporal discharge patterns of the auditory nerve is severely distorted. The distortion is particularly significant at higher (>600 Hz) frequencies, for which the period of the electric stimulus is near the AN refractory period. For example, in response to a 1000-Hz pulse train, most AN fibers may fire on every other stimulus cycle, so that the AN population would represent half of the stimulus frequency rather than the actual frequency.

Rubinstein et al. [Hearing. Res. 127, 108] proposed that the coding of electric waveforms in cochlear implants can be improved if a sustained, electric high-rate (5 kpps) desynchronizing pulse train (DPT) is presented in addition to the information-carrying electric stimulus. The DPT may amplify the inherent noise in AN fibers so as to produce ongoing, stochastic discharges similar to the spontaneous activity in a healthy hear. We tested this hypothesis by recording responses of AN fibers of deafened cats to sustained electric pulse trains.For most fibers, responses to the DPT showed adaptation during the first 2minutes, followed by a sustained response for the remainder of the 10-minute stimulus. These sustained responses partially resembled spontaneous activity in terms of discharge rate and interspike interval distributions. AN fibers were extremely sensitive to modulations of the DPT, responding tomodulations as small as 0.5%. Responses to sinusoidal modulations resembled AN responses to pure tones over a 15-25 dB range of modulation depths.

Responses to complex modulations simultaneously represented several spectral components of the modulator in their temporal discharge patterns. However, for modulation depths above 10%, the representation of both sinusoidal and complex modulators was more distorted. These results demonstrate that strategies that incorporate a DPT, and that use low modulation depths to encode sounds, may evoke AN responses that more accurately represent the modulator in their temporal discharge patterns. If the central nervous system can utilize this information, then these strategies may substantially improve performance enjoyed by cochlear implant users.

Thesis Supervisor: Bertrand Delgutte
Title: Associate Professor of Otology and Laryngology, Harvard Medical School

Thesis Supervisor: Donald Eddington
Title: Associate Professor of Otology and Laryngology, Harvard Medical School

Acknowledgements

This work benefited from the input of many kind, knowledgeable individuals. Foremost, I must thank my advisors Bertrand Delgutte and Donald Eddington for their tireless efforts to corral my work into a focused, coherent, piece of research. My path to this goal was far from straight. Throughout the many detours, they have been simultaneously straightforward about what they thought, supportive of my wanderings, and confident that eventually I would strike "gold." For me, that "gold" was reading, in 1999, Jay Rubinstein's paper entitled "Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation." The ideas expressed in that paper constitute a starting point of my work. I am also thankful to Dr. Rubinstein that his support of my exploration of this topic extended to serving on my thesis committee.

Throughout my 6½ year stay, I have benefited from my discussions with Peter Cariani. The focus of this work on the temporal discharge patterns is also in part due to Peter. Peter also helped me run the early experiments with electric stimulation. Although those experiments didn't go well, I remember them fondly.

I would like to thank all the fellow students in the Speech and Hearing Sciences program (which is the name by which I will remember it, although I like the new name better). I am especially grateful to Barry Jacobson, who has become my close friend through the years. I would also like to acknowledge the help and the friendship of Zachary Smith. Part of his work is actually incorporated into this thesis.

This work has also benefited greatly from the assistance of Leslie Liberman. Leslie has, through personal sacrifices, directly contributed to these experiments in more ways than what meets the eye. Your efforts are greatly appreciated, Leslie!

I would like to thank my parents for believing in me and encouraging me throughout my graduate school experience. Although it seemed at times that I would never graduate, I have never heard from them a word of impatience. I am also grateful to the Boston Jewish community for its warmth, acceptance, and support. I would also like to thank Uri Blank for his close friendship.

It would not be an exaggeration to say that this work would have been impossible without constant support of my wife Yuliya. On May 26th of 2001, our life has been greatly enriched by the birth of Abigail Zelda. Abigail provided the motivation to finish this thesis as quickly as possible. She has also made our lives very entertaining.

I am deeply thankful to G-d for surrounding me with wonderful people, and for giving me many opportunities to learn and to grow. May G-d grant that my thesis be of help to the 40,000 cochlear implant users throughout the world, who are the ultimate inspiration for this work.

This work was supported by the NIH training grant T32 DC00038, as well as by the W.M. Keck Foundation, and NIH research grants DC00361, DC02258 and N01-DC-6-2100.

Contents

Chapter 1

Introduction

In a normally-functioning ear, sound causes motion of the basilar membrane in the cochlea. Specialized hair cells translate basilar membrane vibrations into auditory-nerve electrical discharges. The properties of the basilar membrane are such that sounds of different frequencies excite different cochlear locations. Thus, when the stimulus is a pure tone, the location of the peak in neural excitation progresses from the apex to the base of the cochlea as frequency is increased (Bekesy, 1960). In addition to this spatial code, for frequencies below 5 kHz, the frequency of a pure tone is also coded in the temporal discharge patterns of auditory-nerve fibers (ANFs) because action potentials (spikes) tend to occur at a particular phase of the sound (Rose et al., 1967).

The most frequent cause of deafness is a loss of cells that translate mechanical vibrations into electrical potential. The idea of a *cochlear implant* is to bypass these nonfunctioning transducers and excite the auditory nerve directly using electric stimulation. An important goal of cochlear-implant design is to produce patterns of spike activity on the auditory nerve that resemble normal acoustic responses because such responses will presumably be easily interpreted by the brain. However, recordings from electrically excited ANFs indicate that temporal discharge patterns evoked by electric stimulation differ greatly from acoustic. For example, in a healthy ear the frequency of a 1 kHz tone is coded unambiguously in the temporal patterns of ANF responses, but responses to an electric 1 kpps pulse train may occur on every second pulse and resemble responses to a 500 pps pulse train (Wilson et al., 1997).

Rubinstein and his colleagues (1999b) proposed that the naturalness of responses evoked by cochlear implants might be improved by adding a sustained "desynchronizing"

pulse train (DPT) to a conventional stimulating strategy (Rubinstein et al., 1999a; Rubinstein et al., 1999b). The rationale for the DPT is that across-fiber differences in refractory, sensitivity and other properties, as well as noise present in the neural membrane will result in the responses across fibers being desynchronized with DPT stimulation. Such desynchronization would lead to a more natural representation of the stimulus waveform in temporal discharge patterns in a way that is akin to stochastic resonance (Collins et al., 1995; Wiesenfeld and Moss, 1995). In this thesis, we experimentally test Rubinstein's idea by recording from single units in the electrically stimulated auditory nerve of deaf, anaesthetized cats in response to stimuli incorporating a DPT.

The remainder of this introduction provides a framework for thinking about issues associated with temporal discharge patterns evoked by electric and acoustic stimulation. We start with an overview of cochlear implants. Next, we review coding of electric stimuli in auditory nerve responses. We then compare the responses evoked by electric stimulation to sound responses in a normally-functioning ear. Finally, we outline Rubinstein's DPT idea in more detail, and review previous experiments that are most relevant to that idea.

I. Cochlear implants

The concept of electrically stimulating the ear has existed since the 1800s (Volta, 1800); however, the first clinical use was reported in the late 1950s (Djourno and Eyriès, 1957), with others following soon after (Simons, 1964; House and Urban, 1973).

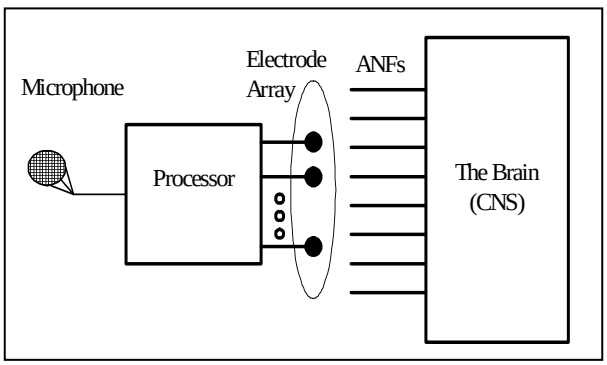


Figure 1. General scheme of all modern cochlear implant devices.

The general scheme of a multi-channel cochlear implant is shown in Fig. 1. A sound processor translates the microphone signal into electric stimuli that are delivered to an implanted electrode array. Currently, electrode arrays consist of up to 22 electrode contacts implanted in the cochlea such that the

electrodes are located along the “tonotopic axis.” It is found that an identical stimulus applied to different electrodes causes a different pitch sensation for most subjects for at least some electrodes (Eddington et al., 1978). The processor seeks to take advantage of this "electrode to pitch" mapping by presenting high-frequency components of the incoming sound on more basal electrodes, and low frequency components on more apical electrodes.

Sound percepts evoked by a cochlear implant also depend somewhat on the waveform of the electric stimulus. For electric pulse trains with pulse rates between 100 and 400 pulses per second (pps), perceived pitch increases monotonically with stimulation frequency for most patients (Eddington et al., 1978; Shannon, 1983; Dorman et al., 1994). For higher frequencies, pitch sensations tend to saturate.

A. CA and CIS processing strategies

The compressed analog (CA) strategy (Fig. 2) is used in several cochlear implant systems (e.g. (Wilson, 1993)). In CA, (a possibly compressed) waveform at the output of each filter is directly presented to the electrodes. Consequently, band-limited signals can be nearly reconstructed from the electrode waveforms. Thus, no information is lost in the output of the stimulating electrodes.

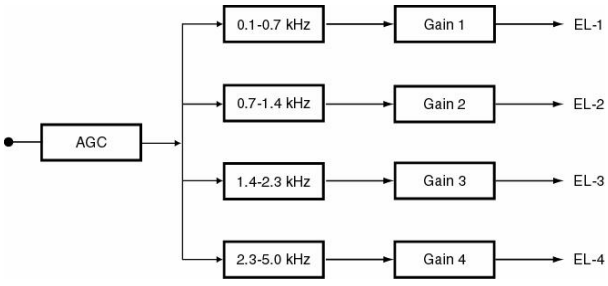


Figure 2. Block diagram of a compressed analog (CA) processor.

In the CA strategy, current is presented simultaneously on all electrodes. Although each neuron is primarily stimulated by the closest electrode, other electrodes can significantly influence its response. Psychophysical experiments reveal strong electrode interaction for patients fitted with the Ineraid device that utilizes a CA strategy (Favre and Pelizzone, 1993). The use of the continuous interleaved strategy (CIS) was advocated for these patients over the CA strategy to reduce electrode interactions (Fig. 3). In CIS, each electrode is stimulated by current pulses. Pulses are presented in an “interleaved” fashion across electrodes so that no two electrodes are stimulated simultaneously (Wilson et al., 1991). Because non-

simultaneous electrode interactions are much smaller than the simultaneous interactions and become negligible when the temporal separation between the pulses exceeds 0.5 ms, (Eddington et al., 1978; Favre and Pelizzone, 1993) one would expect less channel interaction with CIS.

In CIS, temporal information about the incoming sound is represented in the modulations of a high-frequency pulse train. Because the cutoff of the lowpass filter used in the envelope detection stage of the current systems is low (<400 Hz), only limited temporal information is available on electrode outputs. During a vowel, for example, a CIS modulator waveform might not contain periodicities corresponding to formant frequencies. In contrast, CA preserves formant information on electrode outputs.

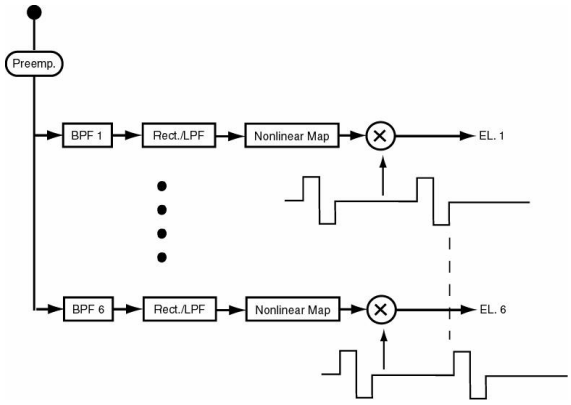


Figure 3: Block diagram of a continuous interleaved sampling (CIS) processor. Adapted from Wilson et al (1991).

Despite apparent loss of information in CIS, it was found to be a significant improvement over CA for subjects that have strong interaction of current fields between adjacent electrodes (e.g. users of the Ineraid device that uses monopolar stimulation), (Wilson et al., 1991). When electrode interaction can be reduced through other means (e.g. by using closely spaced bipolar electrodes), some patients perform similarly using CA and CIS strategies (Battmer et al., 1999). These results suggest that subjects do not benefit from the additional temporal information provided by the CA processor. We will argue in this thesis that the information provided by CA is not properly transferred to ANF responses. Better coding of temporal information on the auditory nerve, while preserving the reduced channel interactions of CIS, could lead to better performance by the implantees.

B. Can one improve on CIS?

The CIS strategy is far from perfect. Although the best CIS users of today achieve high speech-recognition scores, only about 15 to 20% of the users are able to carry on

fluent conversation without lip-reading[1]. In addition, performance of all CIS patients degrades rapidly in noise (Friesen et al.,). With 7-8 channels, performance of the best CIS users is comparable to that of normal-hearing subjects listening to sounds that contain the same reduced spectral information (Eddington et al., 1997; Friesen et al., 2001). While performance of normal-hearing listeners continues to improve for up to 20 channels, the performance of implantees saturates at 7-8 channels for the highest-performing patients, and 3-4 channels for the worst (Friesen et al., 2001). This result suggests that electrode interaction limits the quality of spectral representation in cochlear implant listeners. Fortunately, new modiolus-hugging electrodes may decrease channel interaction, and allow more independent channels.

Several studies have addressed implantees' perceptions of music. Implantees report enjoying music less after implantation than prior to their hearing loss (Gfeller et al., 2000). Melody recognition is likewise poor for normal-hearing listeners of spectrally distorted sounds (Smith et al., 2002). By better representing musical pitches, novel processors may improve implantees' recognition of tonal languages such as Mandarin Chinese. Although representing pitch in the spatial pattern of neural activity requires a large number of independent channels, future cochlear implants may improve coding of musical pitch by improving the representation of the temporal structure in the auditory nerve responses.

II. Neural responses to electric stimulation

In humans, ANF responses to electric stimulation have been investigated using electric compound action potentials (ECAPs) evoked by electric pulse trains. A low-rate (<400 pps) electric pulse train evokes responses to each stimulus pulse of approximately equal amplitude (Wilson et al., 1997). For pulse trains of intermediate rates (>400 pps and < 1000 pps), human ECAPs corresponding to each pulse alternate between strong and weak responses. Alternation suggests that while many neurons respond to the first pulse, these neurons are in a refractory state during the second pulse. If, as the ECAP data suggest,

[1] However, more than 80% are able to communicate fluently with lip-reading in quiet conditions.

most neurons fire together, then the most auditory-nerve fiber population would represent half of the stimulus frequency rather than the actual frequency [2].

Single-neuron responses to electric pulse trains and sinusoids have been investigated in animal models. Several trends are common to studies of pulsatile and sinusoidal stimuli. First, electric stimulation can evoke average discharge rates as high as 1000 spikes/s for at least several seconds (Moxon, 1967; Hartmann et al., 1982; Dynes and Delgutte, 1992). Second, electric responses are more synchronized to the stimulus period than their acoustic counterparts (Hartmann et al., 1982; Dynes and Delgutte, 1992). Finally, the range over which the discharge rate changes with stimulus level is much smaller for the electric responses (Kiang and Moxon, 1972; Javel et al., 1987; Dynes, 1995).

Knauth et al. studied coding of compressed, filtered electric vowels in ANF responses (Knauth et al., 1994). In that study, synthetic vowels were produced using a single-channel CA implant, and presented through a round-window electrode. Knauth et al. found that, for a range of stimulus levels, fibers fired once per vowel period. At these levels, the neural pattern contains information about the fundamental frequency, but not about vowel identity. At higher levels, the neuron could respond more than once per vowel period. The time of the second spike was related to the frequency of either the second or the third formant of the vowel, and to fiber refractory properties. At the highest levels, neurons fired several times during a single period, and afforded some representation of the vowel fine-time structure. However, this representation was strongly distorted by neural refractory properties. Because these highest levels were above threshold for facial contractions, the authors concluded that they are larger than the maximum comfortable levels that would be used by a CI user. However, because intracochlear electrodes may produce electric fields that are more localized to the cochlea, this consideration may not apply for most implantees. Overall, the Knauth et al.

[2] Alternation in the ECAP is a less consistent finding for animals (Abbas et al., 1999), indicating that their nerves may be able to represent higher frequencies.

(1994) data suggest that with CA, only low-frequency voice pitch is robustly coded in ANF responses.

III. Neural representation of acoustic stimuli in a healthy cochlea

Sound-evoked responses of ANFs in a healthy ear may be very different from their responses to electric stimulation. In the following section we briefly review ANF responses to sound in a normally-functioning ear, and compare them with ANF responses to electric stimulation.

A. Spontaneous activity

In a healthy cochlea, ANFs discharge even in absence of sound (Kiang et al., 1965). Discharge rates during spontaneous activity can range from 0 to over 140 spikes/s (Liberman, 1978). Temporal properties of spontaneous activity have been characterized by Kiang et al. (Kiang et al., 1965). Interval histograms of spontaneous activity are exponential in shape for intervals that are longer than the relative refractory period. Kiang et al. also characterized the dependence of each interval on the immediately preceding interval, and found no evidence for correlation between successive intervals. The latter finding suggests that spontaneous activity can be approximately described by a renewal process. Recently, both long-term and short-term correlations in spontaneous activity have been investigated (Teich, 1989; Lowen and Teich, 1992; Kelly et al., 1996). These studies report small deviations from the classical findings.

Classically, it has been reported that the vast majority (96%) of fibers in a deaf cochlea have no spontaneous activity (Kiang et al., 1970b). This view has been challenged recently by the results of Shepherd and Javel (1997), which suggested that nearly 50% of neurons from long-term deafened cats do exhibit some spontaneous responses. These responses tended to occur at low rates (<5 spikes/s), be regular, and occasionally exhibited bursting. Because they found few hair cells in serial sections of the cochlea, they suggested that spontaneous activity must be of non-synaptic origin. However, these ANFs were not specifically tested for sound responses. In addition, because the aminoglycosides used in these studies may directly affect auditory-nerve

fibers (Leake et al., 1997), it is unclear whether the non-synaptic spontaneous activity would be seen in patients with different etiologies of deafness. While it is difficult to ascertain the mechanisms responsible for the differences in the results of the two studies, it seems clear that spontaneous activity in a deaf cochlea can differ from that of a healthy ear, and that in at least in some deaf ears, there is much less spontaneous activity as compared to that in a healthy ear.

B. Pure-tone responses

Pure tones can be used to characterize frequency selectivity of ANFs. At low levels, each fiber responds only to a narrow range of frequencies; the response range is significantly broader at higher levels. The tuning characteristics of an ANF depend strongly on its cochlear place, and resemble the mechanical filtering of the cochlea (Ruggero et al., 1990). In comparison, tuning to electric sinusoids is broad, and differs little from one fiber to the next (Kiang and Moxon, 1972).

As the level of a pure tone is raised above threshold, discharge rate increases over a 20-40 dB range and then saturates at approximately 100-300 spikes per second (Sachs and Abbas, 1974; Liberman, 1978). The rate versus level curve can be accounted for by the properties of the hair-cell/ANF synapse and by the mechanics of the inner ear (Ruggero, 1992). In contrast, for sinusoidal electric stimulation, discharge rate grows rapidly (3-5 dB) and saturates at much higher rates (close to 800 spikes/s) (Kiang and Moxon, 1972).

In a healthy ear, for tones of frequency up to 4 kHz, the probability of instantaneous discharge tends to increase over spontaneous rate for one portion of the stimulus cycle and decrease for the opposite portion (Rose et al., 1967). Thus, the temporal discharge pattern reflects the periodicity in the stimulus waveform, and the response is said to be “phase locked” to the stimulus. Phase locking can be revealed by computing a “period histogram” locked to the stimulus period. At very low levels, responses of high spontaneous-rate fibers faithfully resemble the stimulus waveform (Johnson, 1980). At higher levels, although period histograms partially reflect the sinusoidal waveform of the tone, they differ from the sinusoid in several respects. First, for moderate and high

stimulus levels, positive excursions from spontaneous rate are generally greater than negative excursions; thus, the period histogram representation of the tone is half-wave rectified. Secondly, especially at low frequencies, the period histogram is somewhat skewed. Lastly, for very high levels and low frequencies, histogram can show "peak splitting," where several peaks appear in the histogram.

In contrast, ANF responses to electric sinusoids and pulse trains are more phase locked than responses to tones (Hartmann et al., 1984). The representation of the sinusoidal waveform is much more distorted in these responses, even at levels very close to threshold.

C. Steady-state responses to speech-like waveforms

Speech possesses nearly-steady segments over durations of 50-200 ms. Each steady segment can be distinguished by its spectral structure. When presented in isolation, these steady portions can be spontaneously identified as vowels or consonants. Thus, spectral structure is, in some cases, sufficient for phoneme recognition. The most important features for vowel perception are the frequencies of the spectral maxima associated with the first three formants (Peterson and Barney, 1952). How these peaks in the spectrum are represented in the auditory nerve, and which representation is actually used by the central processor has been the subject of extensive study and debate.

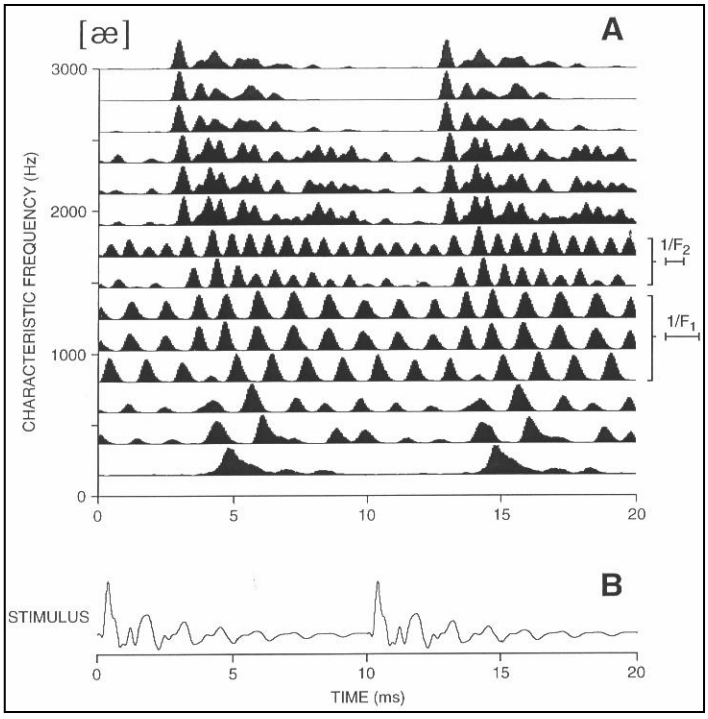


Figure 4. Temporal pattern of response of ANFs of differing CFs during the vowel [ae] stimulus (panel B). Adopted from Delgutte (1997).

One possibility is that the spectral peaks are represented spatially, i.e. as peaks of activity of the ANFs at the appropriate

CF regions. This spatial representation has been explored in detail by Sachs and colleagues (Sachs and Young, 1979; Sachs et al., 1983). They found that, while peaks at the formant frequencies are clearly represented at low levels, they are poorly represented either at high levels or in the presence of noise for the majority of fibers. Although these results represent a serious challenge to the "spatial" representation, they by no means rule out that such representation is in fact used by the auditory system. There exists a population of high threshold fibers that do not saturate at even the highest levels (Sachs et al., 1988). By selectively weighting these fibers at high levels, it may be possible to realize an accurate spatial representation of the formants over a wide range of levels (Delgutte, 1982).

Spectral representation is coded in a CI by first filtering the incoming sounds into several bands and then presenting the output of these bands to different electrodes along the cochlear partition. Because the number of channels used in CA or CIS strategies is typically small (4-7), representation of spectral shape is inherently less precise than in a healthy ear. However, such reduced spectral information has been shown to be sufficient for good recognition of simple sentences in normal listeners (Shannon et al., 1995).

Spectral information is also represented in the temporal patterns of ANF responses. Panel A of Fig. 4 shows a neurogram of the ANF responses to the vowel [ae]. Two periods of the stimulus waveform are shown in panel B. First, responses are clearly phase locked to stimulus components. Which components are most prominent in the response depends on the CF of the fiber. For fibers with CFs below 2000 Hz, the neurogram has peaks that are spaced at approximately the inverse of the closest formant frequency. For fibers with CFs above 2000 Hz, responses are composed of several harmonics. The pitch of the vowel can be readily discerned from responses of these fibers.

As discussed previously, responses to complex electric stimuli differ from acoustic responses. First, although pitch is coded in the interval pattern of responses to electric stimulation, the formants are poorly represented (Knauth et al., 1994). Second, refractory properties of the auditory nerve distort the representation of electric harmonics above 400 Hz (Knauth et al., 1994).

In summary, in a healthy ear, ANFs encode the sound's fine-time structure information in their temporal discharge patterns. The same may not be true in the ear that is stimulated by a cochlear implant with either a CA or a CIS strategy.

IV. Desynchronizing Pulse Train

Rubinstein et al. (1999b) proposed that the representation of the fine-time structure of an electric stimulus in auditory-nerve responses might be improved by introducing a sustained, high-frequency, "desynchronizing" pulse train (DPT) in addition to the electric stimulus. When enhanced by a DPT, the responses to modulated electric pulse trains may more accurately resemble the responses of a healthy ear to the corresponding sound. The rationale for this improvement comes from theories of stochastic resonance (Collins et al., 1995). Stochastic resonance applies to many threshold systems, and in particular to nerve fiber responses. In so far as a neuron behaves as a simple threshold device, only threshold crossings will be represented in the neural responses. Portions of the waveform that are either significantly above or below threshold will not be represented. However, if independent noise is added to the stimulus, then the instantaneous stimulus amplitude can be represented in the probability that the threshold is crossed. Thus,

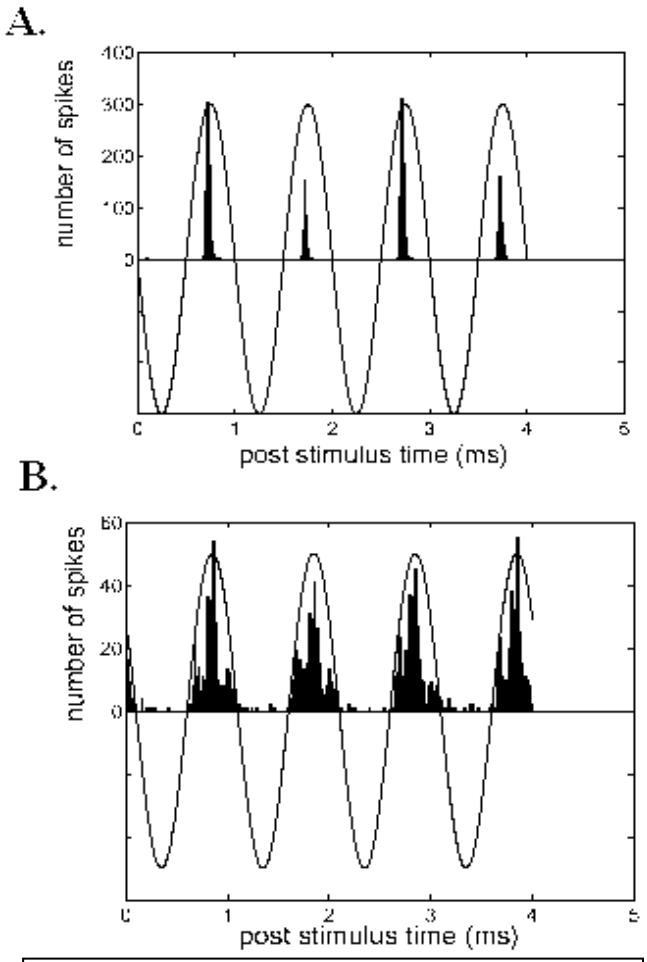


Figure 5. Response of Rubinstein's model to a sinusoid (panel A), and to a sinusoid plus a 5 kpps conditioner (panel B) (adopted from Rubinstein et al., (1998). By itself, the conditioner evokes "pseudospontaneous" activity with discharge rate of 125 spikes/s.

the fine temporal structure of the waveform could be better represented in the neural response. Activity evoked by the DPT can be considered to be such noise. This is because the DPT will cause different fibers to be in different refractory or sensitized states and therefore have different instantaneous thresholds to electric stimulation.

To demonstrate that these effects may in fact occur, Rubinstein *et al.* simulated responses of a biophysical model of electrically-stimulated ANFs to 5 kpps pulse trains superimposed upon electric sinusoids (Rubinstein et al., 1998). A period histogram of model responses to a sinusoid with and without the DPT is shown in Fig. 5. Without the DPT, the response histogram is highly synchronized, and poorly represents the sinusoid (A). Alternation of high and weak responses to successive cycles is also apparent. With the DPT (B), the response better represents the sinusoidal waveform. In addition, alternation is not apparent in these responses.

Rubinstein further argued that responses to the DPT might mimic spontaneous activity that is present in a healthy ear, but is usually absent in a deaf one (Rubinstein et al., 1999b). Kiang *et al.* (1970a) suggested that some forms of tinnitus (ringing in the ear) are related to the absence of spontaneous activity. By restoring spontaneous activity through electric stimulation, it may be possible to reduce tinnitus in these patients.

While Rubinstein's idea involves sustained stimulation with a high-rate electric pulse train, only a few studies have investigated responses of ANFs to such stimuli. These studies are reviewed in the Introduction of Chapter 3. To the best of our knowledge, there are no published single-unit studies of ANFs responses to either 5 kpps pulse trains lasting longer than a minute, or to modulated pulse trains. The effect of a DPT on the coding of modulated pulse trains has been investigated by measuring the ECAPs. Wilson and coworkers measured the ECAPs evoked by modulated stimuli with and without a DPT (Rubinstein et al., 1999b; Wilson et al., 1998). They showed a greater similarity between the stimulus waveforms and ECAP responses when a 5 kpps DPT was presented in addition to the modulated waveform. A limitation of these studies is their reliance on ECAP measures of neural activity, which only provide indirect evidence for the events that occur at the single-unit level. Although these studies do suggest that the DPT may

desynchronize ANFs, it is unclear whether the underlying ANF responses resemble ANF responses to sound in a healthy ear. In addition, these studies investigated very short (30 ms) desynchronizers. Short pulse trains are a poor model of the sustained DPT proposed by Rubinstein et al (1999b).

V. Overview of the thesis document

In this thesis, we directly test Rubinstein's idea by recording from single fibers in the electrically-stimulated auditory nerve of deaf, anaesthetized cats. The following four chapters of the thesis are self-contained papers. Chapter 2, which is published (Litvak et al., 2001), describes our initial experiments with responses to high-rate pulse trains. To simplify our task, we recorded responses to relatively short (400 ms) electric bursts. When the level of the pulse train was adjusted for each ANF, we found that responses possessed many of the properties hypothesized for DPT stimulation.

Encouraged by the results of the first study, we undertook the more challenging task of recording activity to longer (10 minute) stimuli. Chapters 3-5 analyze the responses of ANFs to a 10 minute, 5 kpps DPT, which better represents the DPT concept. In the Chapter 3, we compare the activity evoked by the unmodulated DPT to spontaneous activity in a healthy ear. In Chapter 4, we measure responses of ANFs to sinusoidal modulations of the DPT, and compare those to ANF responses to tones in a healthy ear. In Chapter 5, we examine responses to the DPT modulated by complex, vowel-like waveforms.

The last chapter of this thesis discusses the implications of the results of this research for the design of new sound-processing strategies for cochlear implants that may produce ANF responses mimicking sound-evoked responses in a healthy ear. Connections are also made between this research and several fundamental issues in the physiology and psychophysics of electric hearing, such as the relationship between neural dynamics and physiological responses, neural mechanisms of psychophysical modulation detection, and the role of temporal information in auditory perception.

Chapter 2

Auditory-nerve fiber responses to electric stimulation: modulated and unmodulated pulse trains

I. Introduction

In the continuous interleaved sampling (CIS) strategies used in many modern cochlear implant processors, temporal information about incoming sounds is encoded in the amplitude modulations of pulse trains (Wilson et al., 1991). Proper representation of modulation in temporal discharge patterns of the auditory nerve is an important goal in these strategies.

Despite the popularity of CIS schemes, the responses of auditory-nerve fibers to a sinusoidal modulation of an electric pulse train can be very different from responses to a tone in a healthy ear. For modulation frequencies below 500 Hz, virtually every stimulated neuron is likely to entrain to the modulator (i.e. produce a spike discharge for every modulator cycle). In contrast, neurons of a healthy ear responding to an acoustic tone fire at random multiples of the stimulus period. For example, there may be 1, 2, 3 or more cycles between successive spikes (Rose et al., 1967). The situation is even worse at higher frequencies because, with electric stimulation, neurons may fire on every other cycle or even higher multiples of the modulation period. If most stimulated neurons fire together, then the population of auditory neurons would code a submultiple of the modulator frequency rather than the actual frequency (Wilson et al., 1997).

Rubinstein et al. (Rubinstein et al., 1999b) proposed that the coding of modulation waveforms might be improved by introducing a sustained, high-frequency, "desynchronizing" pulse train (DPT) in addition to the modulated pulse train. The

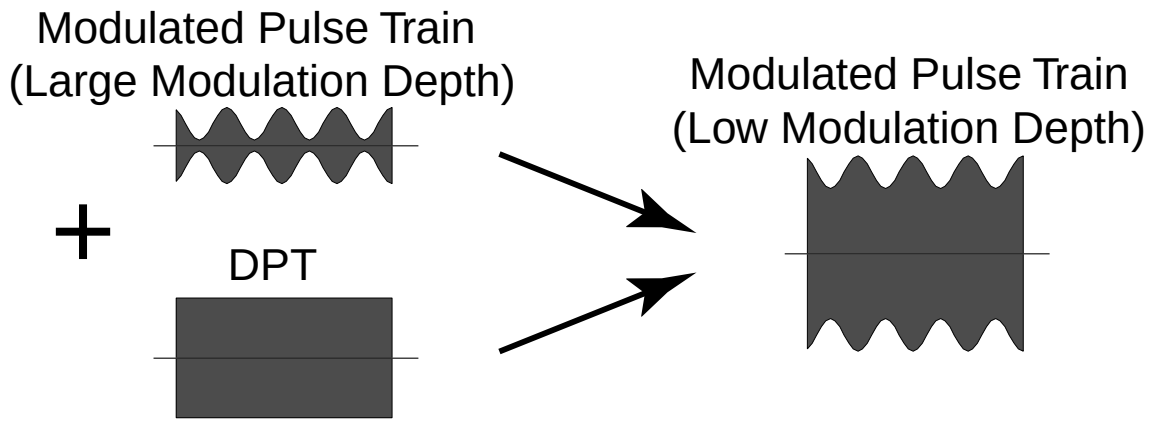


Figure 1. The left panel shows two stimuli generated by a cochlear implant to implement a DPT protocol, as suggested by Rubinstein et al. (1999). This stimulus is composed of two pulse trains: one, strongly modulated, is the CIS signal, and the other, unmodulated, is the DPT. The carrier for the two signals need not be identical. We assume that the DPT-enhanced stimulus can be modeled by a carrier of the same frequency as the DPT that is more weakly modulated than the original CIS stimulus (right). This approximation is exact if the CIS and the DPT have identical carriers.

rationale for the DPT is that across-fiber differences in refractory, sensitivity and other properties, as well as noise present in the neural membrane will result in the responses across fibers being desynchronized after the first few hundred milliseconds of DPT stimulation. Such desynchronization would lead to an improved representation of the modulator in temporal discharge patterns by allowing an ensemble of neurons to encode the true modulator frequency rather than a submultiple.

In this paper, we report the responses of auditory-nerve fibers to both modulated and unmodulated electric pulse trains that were recorded to test the ideas underlying the DPT. We focused on two specific questions:

1. Do the responses to a sustained high-frequency pulse train resemble spontaneous activity? Specifically, we characterized interval histograms (IH) for pulse trains and compared them to the nearly exponential histograms observed for spontaneous activity in an intact ear (Kiang et al., 1965). We also quantified the variability in the spike count from presentation to presentation, and compared it to the variability expected for normal spontaneous activity.

2. Does use of a high-frequency DPT result in a better representation of modulation frequency? We used modulated high-frequency pulse trains with low modulation depths (≤ 10%) to imitate the effect of a DPT. We assumed that neural responses to a high-frequency pulse train with a low modulation depth are similar to those elicited by a stimulus that is a sum of a sustained DPT and a highly modulated pulse train (Fig. 1). This assumption may hold if the membrane time constant is large compared to the intervals between pulses. We compared period and interval histograms of responses to electric, modulated pulse trains with acoustic responses to tones.

II. Methods

A. Animal preparation

Cats were first anesthetized using dial in urethane (75 mg/kg). Co-administration of kanamycin (subcutaneous, 300 mg/kg) and ethacrinic acid (intravenous, 25 mg/kg) was then used to deafen the animals (Xu et al., 1993). The bulla was opened to expose the round window. An intracochlear stimulating electrode was inserted about 10 mm into the cochlea through the round window. The electrode was a 400 um Pt/Ir ball. A similar electrode was inserted into the base of the cochlea for compound auditory potential (CAP) recordings. The round window was then sealed using connective tissue. The ear bar was used as a reference electrode for both stimulation and recording.

In order to verify that the animal was deafened, we measured a CAP in response to acoustic clicks (condensation, 100 μs). The CAP was measured in the implanted ear. In all cases, no CAP was noted for the highest click levels (~90 dB SPL) investigated.

B. Stimuli

Stimuli were delivered through an isolated current source and were either (1) unmodulated pulse trains (150 ms or 250 ms duration) with pulse rates of 1.2, 2.4, 4.8 or 24 kpps (kilo pulses per second) or (2) "modulated" 4.8 kpps pulse trains (first 50 ms or 150 ms unmodulated; last 100 ms modulated; modulation frequency: 400 Hz) of varying modulation depth. In all cases, pulse trains consisted of cathodic-anodic (CA) biphasic pulses (20.8 μs per phase). Modulated stimuli were modulated "down" such that the peak amplitude was equal in the modulated and the unmodulated portion of the stimulus.

Stimuli were presented at a repetition rate of 1 per second. Stimulus level was adjusted to obtain discharge rates of 50 to 400 spikes/s. All stimulus levels reported in this study are peak currents.

C. Recording techniques

Standard techniques were used to expose the auditory nerve via a dorsal approach (Kiang et al., 1965). We measured from single units in the auditory nerve using glass micropipettes filled with 3M KCl. A digital signal processor (DSP) was used to separate neural responses from the stimulus artifact (voltage excursions recorded at the micropipette that are not due to neural discharges). First, we recorded the "artifact" at a subthreshold stimulus level. Then, a scaled version of the recorded "artifact" was subtracted from the incoming waveform in real time. The gain applied to the recorded waveform was adjusted to optimally match the recorded and the incoming waveforms. The operation of recording the artifact was repeated for each neuron and for each stimulus studied. An important assumption of this technique is that the artifact grows linearly with stimulus level. Consequently, non-linearities in the conducting medium, stimulation system or the recording equipment decrease the effectiveness of the cancellation for levels that are significantly above the neural threshold. In a saline solution, the stimulus artifact could be cancelled effectively at up to 6 dB above the recorded level. In an actual experiment, however, time constraints in finding the highest level at which there are no spikes, instability of the artifact waveform, the non-linearity of biological tissue and contributions of non-linear gross evoked responses limited the levels that could be investigated to no more than 2.5 dB above fiber threshold.

Times of the spike peaks were measured with 1 μs precision, and recorded in computer files for both on-line and off-line analysis.

D. Unit selection criteria

Possible hair-cell mediated activity ("electrophonic hearing") complicates the interpretation of responses of ANFs to electric stimulation. Hair cells might not be completely eliminated by the acute deafening protocol used in this study. To minimize the effect of any remaining hair cells, only units that (1) had no spontaneous activity, (2)

did not respond to an acoustic click at 90 dB SPL and (3) had unimodal, short-latency (<1 ms) PSTs in response to an electric biphasic pulse (CA, 20 μs per phase) were included for further analysis. The last criterion is based on the observation that when the hair cells are intact, responses to single pulses may contain late components that are hair cell mediated (Moxon, 1967; Moxon, 1971; van den Honert and Stypulkowski, 1984; Javel et al., 1987).

Electrically-stimulated ANFs can exhibit long-term adaptation with the time scale of seconds (Moxon, 1967). This implies that for relatively short repetition times, the responses can undergo a steady change in discharge rate from one run to the next. Although we used reasonably long (750 ms) pauses between successive presentations, some units still showed statistically significant adaptation in discharge rate throughout the measurement. Records were included in the analysis only if the correlation between the number of spikes per presentation and presentation time was not significantly different from 0 ($P > 0.01$).

Electrically stimulated ANFs also adapt on a shorter time scale. Responses of ANFs to sustained electric stimuli typically show a strong initial response followed by a gradual decrease in discharge rate over the course of 30-100 ms (Moxon, 1967; Killian, 1994). For analysis of interspike intervals, we selected a window in which the discharge rate is nearly constant. We developed a recursive algorithm for selecting such a window automatically. The algorithm begins by selecting a 10 ms window centered 135 ms after the onset of the 150 or 250 ms stimulus. For each step, the mean and variance of the spike count in the current window is compared to the mean and the variance of the spike count in the adjacent 10 msecs windows using the permutation test (Efron and Tibshirani, 1993). If the two measures are not significantly different in the two windows (two-sided achieved significance level (ASL) > 0.03), the adjacent window is appended to the steady-state window, and the procedure is repeated. The algorithm stops when both adjacent windows are rejected, or when the end of the stimulus is reached. Responses to electric pulse trains in the computed window were used to compute histograms that were compared with those for spontaneous activity in a healthy ear.

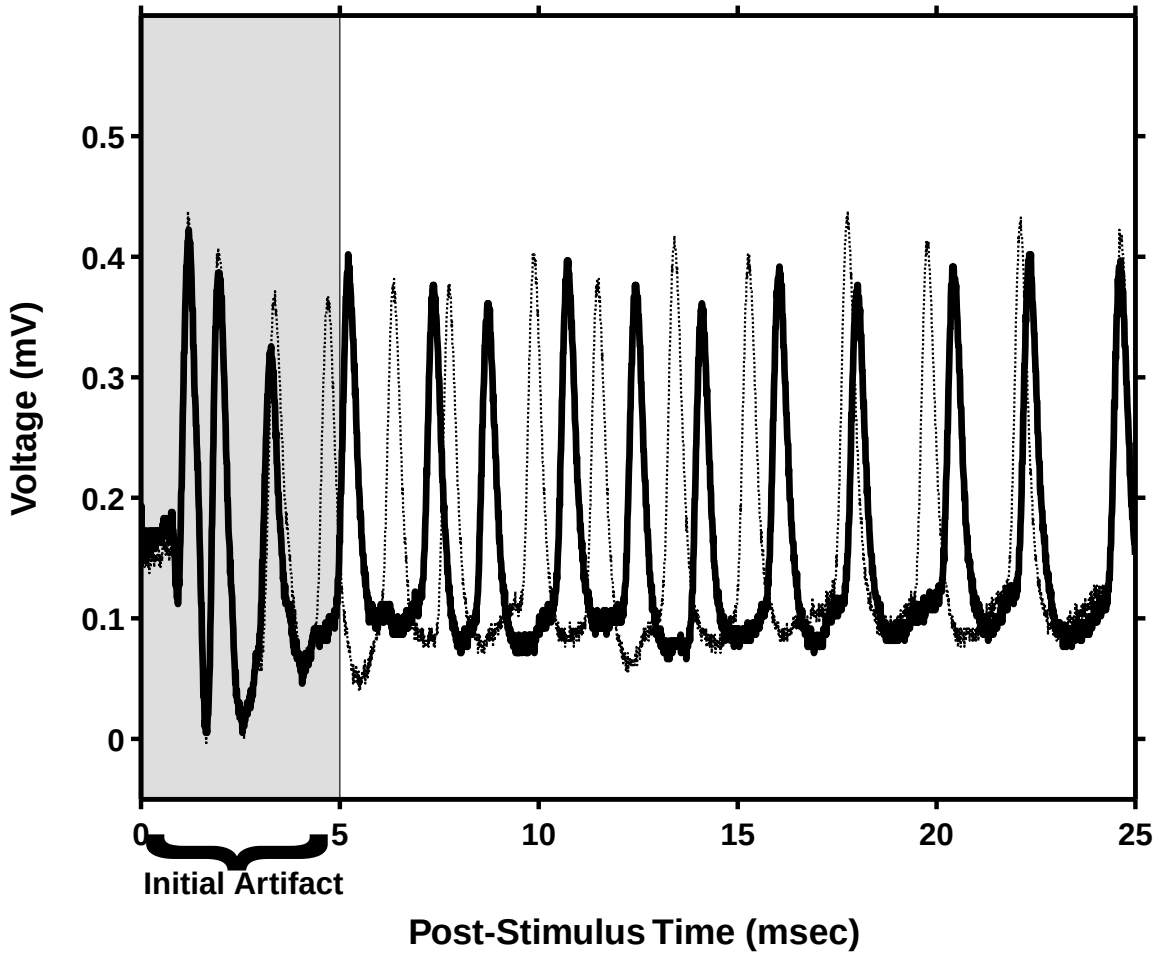


Figure 2. Example of the raw spike waveforms derived after applying the cancellation strategy to remove the artifact. Responses to two presentations of the stimulus are shown (dark and light lines). The artifact template was learned at approximately 1 dB below the stimulus level used to generate the plotted responses. The gray area shows the portion of the initial response where the artifact is typically large. Both the initial peak, as well as the change in the envelope of the response during the first 5 ms are related to the residual artifact.

III. Results

We recorded from 106 single units in 5 acutely deafened, anaesthetized cats. Spontaneous activity was present in 13 units. Spontaneous activity can be detected in deafened preparations with no remaining hair cells (Shepherd and Javel, 1997). Nevertheless, units with spontaneous activity were not included in the consequent analysis. No unit responded to acoustic stimulation or exhibited a multi-modal response

to a single biphasic pulse. These results suggest that the deafening protocol was largely effective in eliminating hair cell mediated activity.

Artifact cancellation was applied in real time to the neural recordings. An example of the output of the cancellation is shown in Fig. 2. For the data presented in this paper, the artifact was at most 5% of the spike height. Because the cancellation technique was not effective in canceling the artifact for the earliest responses to high-rate pulse trains, responses that occurred within 5 ms of the stimulus onset were not analyzed.

The presentation-to-presentation stability test (Methods, Section D) was applied to 430 spike records. The test rejected nearly 40% of the records. Thus, despite our relatively long inter stimulus times (greater than 750 ms for 250 ms stimulus) adaptation of single units from presentation to presentation can still be significant. Responses to longer stimuli (250 ms) were less stable than responses to shorter stimuli (150 ms), as were responses at higher discharge rates.

For the records rejected, we frequently observed that the response rate in the first 10-20 presentations was significantly different from the later responses. For these records, the stability analysis was repeated for the late responses only. Using this less stringent criterion, we were able to include the late responses for an additional 7% of the records.

It should be pointed out that although the results in this paper are based only on analysis of the records that passed the latter stability test (67% of the data), the conclusions are not changed if the rejected records are included in the analysis. The primary purpose for excluding the unstable records is to demonstrate that instability in the recordings cannot account for the results reported in this study.

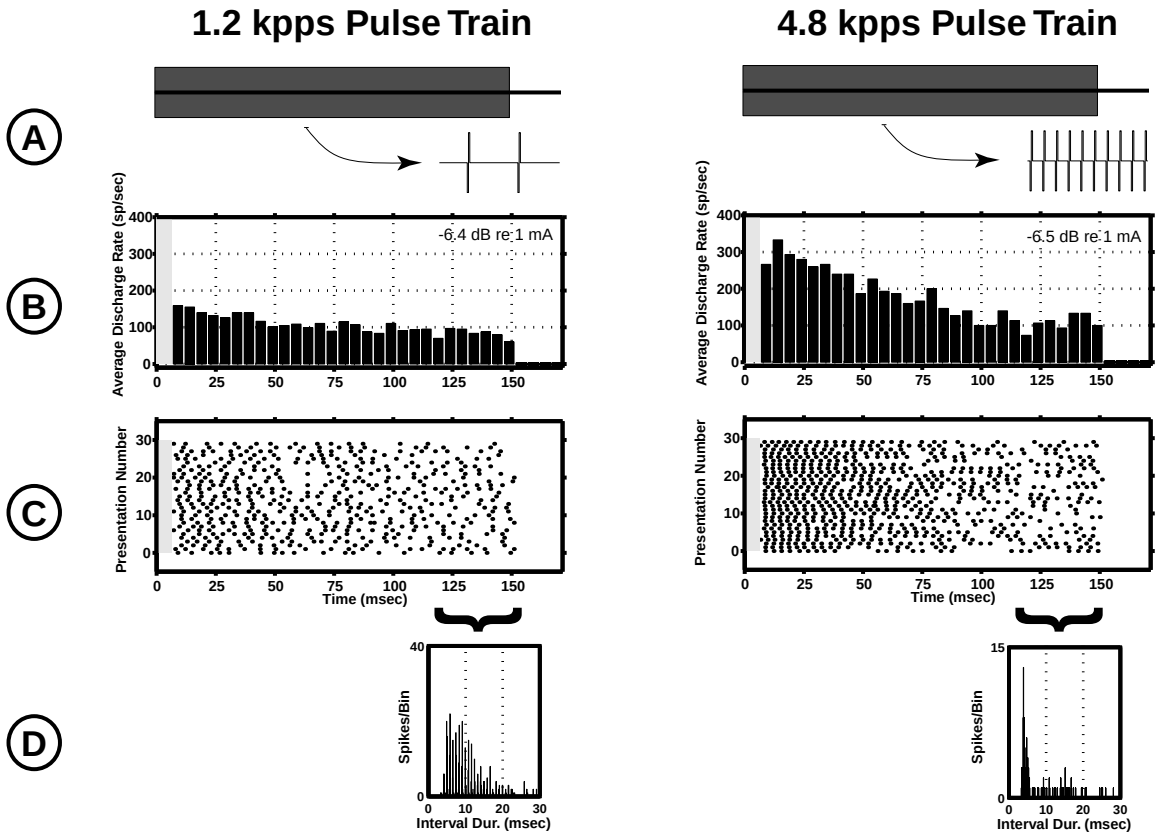


Figure 3 Responses of a fiber to two unmodulated pulse trains of similar levels with pulse rates of 1.2 kpps (left) and 4.8 kpps (right). Panel A shows each stimulus with an expanded time segment of 2 ms. Panel B shows the PST histogram of the unit's responses. Gray bars indicate areas where spike data were discarded due to a large stimulus artifact. The response to the 4.8 kpps pulse train shows greater adaptation during the 150 ms window than the response to the 1.2 kpps stimulus. Panel C shows the dot raster plots of the responses. The desynchronization of the responses towards the later part of the stimulus is apparent. Finally, the last panel shows the interval histogram computed from responses that occurred during the last 30 ms of each stimulus. While the envelope of the IH computed from responses to the 1.2 kpps pulse train is similar to Poisson, the envelope of the IH computed from the responses to the 4.8 kpps pulse train is very different, showing a strong mode at 5 ms that is followed by a longer tail.

A. Unmodulated pulse trains

Fig. 3 shows responses of a fiber to two unmodulated pulse trains with pulse rates of 1.2 kpps (A, left) and 4.8 kpps (A, right). At this level, both stimuli evoke sustained responses from the unit. However, there is more adaptation in response to the higher-rate stimulus (B). This is a common finding in our data. For both pulse rates, responses are initially highly synchronized across trials, and become desynchronized over the course of the stimulus. This can be seen in the scatter of the response times from trial to trial in the dot raster plots (C). The interval histogram (IH, panel D) for the 1.2 kpps pulse train exhibits phase locking to the pulses, and a roughly exponential envelope. In contrast, the IH for the 4.8 kpps pulse train has a non-exponential envelope, with a pronounced mode

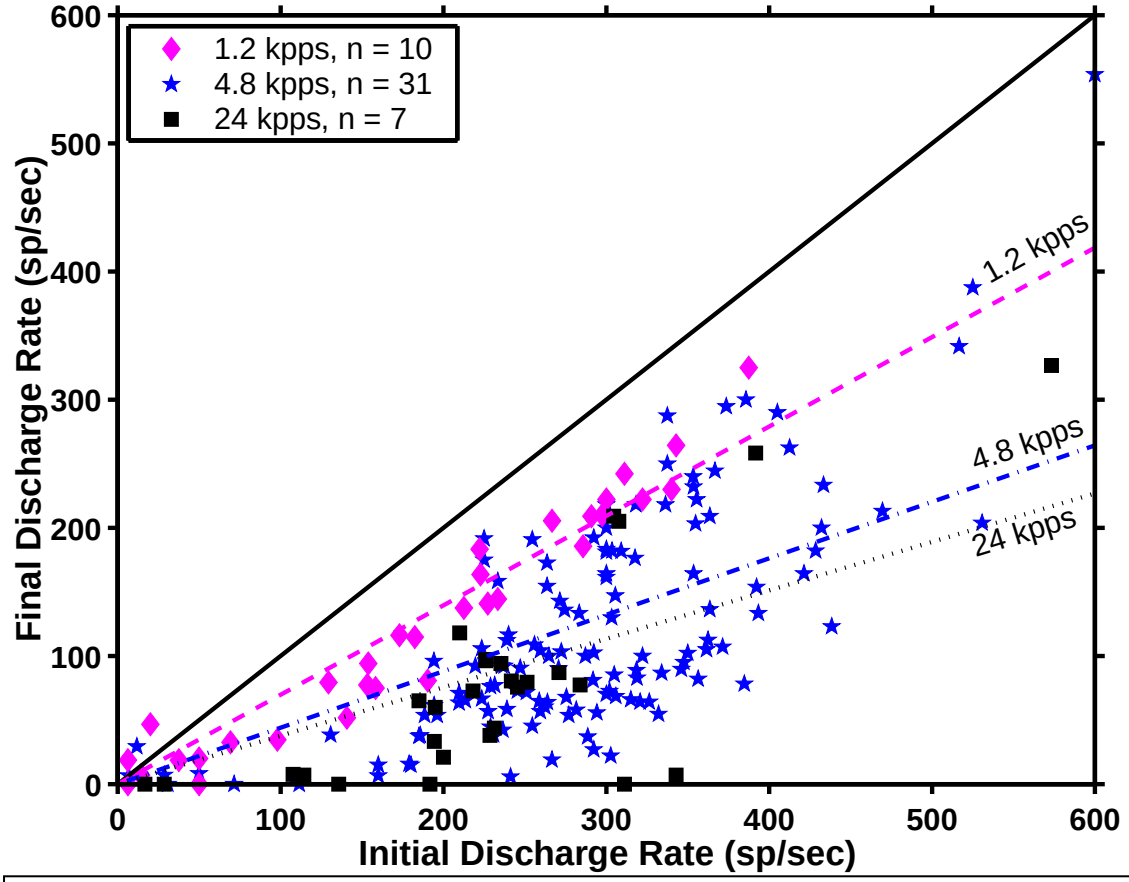


Figure 4. Adaptation of an auditory nerve fiber response to an electric pulse train is a function of pulse rate. Adaptation is represented by the fact that the final discharge rate (rate at 140-150 ms) is lower than the initial discharge rate (rate at 10-20 ms). Each point is based on an average response over 20-40 stimulus presentations for one unit and stimulus level. In the legend, "n" is the number of units for which at least one record was included. Different symbols represent responses to different pulse rates of the electric stimulus. Linear regressions (represented by broken lines) were computed for each pulse rate and were constrained to include the origin. The 1.2 kpps data show significantly less adaptation than the responses to 2.4 or 4.8 kpps. The difference between the 4.8 and 24 kpps data is not statistically significant.

at 5 ms. This mode is not related to the stimulus period, but is inversely related to the average discharge rate. At this coarse bin width (0.208 ms), phase locking to the pulse train is not apparent.

1. Adaptation

As used here, *adaptation* refers to a slow (on the order of 30 to 100 ms) change in the response discharge rate over the course of the stimulus. We found that adaptation is a function of pulse rate. Fig. 4 shows the final rate (the discharge rate in the 10 ms window centered at 145 ms after the onset of the stimulus) versus the initial rate (the rate in a 10 ms window centered at 15 ms from stimulus onset). The responses during the first 10 ms

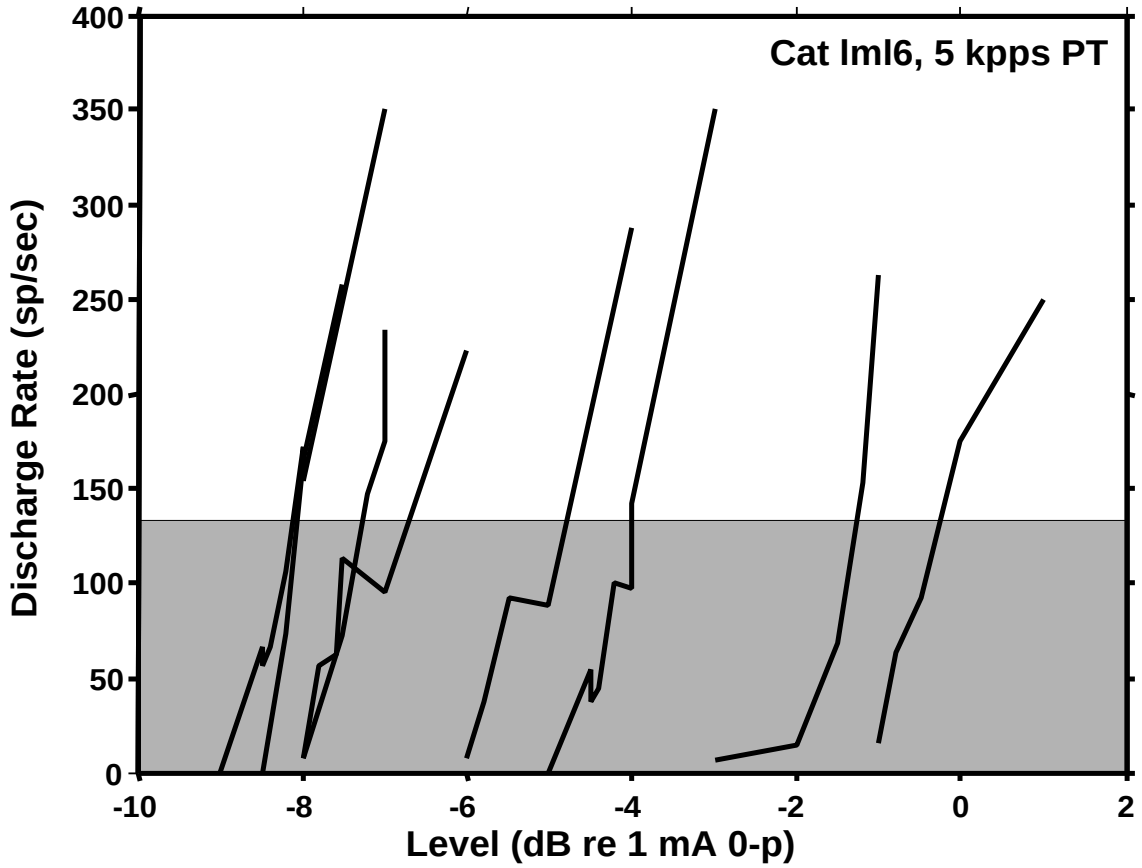


Figure 5. Rate-level functions for 8 fibers recorded from one animal as a function of stimulus current. Discharge rate is computed in a 10 ms window centered 145 ms after stimulus onset. The gray area indicates the range of rates that have been reported for spontaneously responding ANFs in an intact ear (Liberman, 1978). For each fiber, only a narrow (1-2 dB) range of stimulus levels result in a discharge rate that is appropriate for spontaneous activity.

were not included in the analysis. The solid black line indicates where the initial and the final rates are equal. Records falling on this line would show no adaptation.

The scatter in the points indicates that adaptation varies greatly across units, as reported in previous studies (Dynes and Delgutte, 1992; Killian, 1994). This variability is large even when the stimulus evoked comparable initial discharge rates.

The dashed lines represent linear regressions for 1.2, 4.8, and 24 kpps stimuli. The slope of the 1.2 kpps regression line is significantly steeper than those for the 4.8 and 24 kpps ($p<0.001$, permutation test for 4.8 and 24 kpps). Thus, for stimuli that evoke similar initial discharge rates, the response tends to adapt less for pulse trains of lower pulse rate.

2. Dynamic range

In this study, level of the stimulus was adjusted for each fiber. While future cochlear implants might stimulate the auditory nerve more selectively, in the current designs a single stimulating electrode excites many fibers. If such an electrode is used to present DPT stimulation, differences in the responses of the stimulated fibers must be considered. Fig. 5 plots the response rate in a 10 ms window centered 145 ms after stimulus onset as a function of level for 8 fibers from the same animal. Each fiber responds at a rate appropriate for the range of spontaneous activity (gray area in the plot) over a limited range of levels (about 2 dB). Because the range of threshold across fibers for electric stimulation is 10-15 dB, a single electrode stimulus will, at best, result in only a small fraction (~20%) of the available ANFs responding at a rate appropriate for spontaneous activity.

3. Variability

When stimulated acoustically, auditory-nerve responses show pronounced variability in the number of spikes elicited from trial to trial. Similarly, spontaneous activity shows variability in the spike count from one time interval to another (Kelly et al., 1993; Teich and Khanna, 1985). The variability can be quantified by the Fano Factor:

$FF \equiv \frac{Variance[N_i]}{Mean[N_i]}$, where Ni represents the number of spikes on trial i.

For short (<100 ms) time intervals, the FF for normal spontaneous activity is consistent with a Poisson model with dead time of 2 ms (Kelly et al., 1996). Fig. 6 plots the FF for responses to unmodulated pulse trains of 1.2, 4.8 and 24 kpps as a function of average discharge rate in the "steady-state" window. For rates below 180 spikes/s, most points fall within the 99% confidence interval for a Poisson model with a dead time of 2 ms (shaded area). For higher discharge rates, the data tend to fall below the predicted range, indicating that there is less variability than expected for a Poisson model. In any case, for low and moderate discharge rates, variability in spike count from trial to trial is comparable with that for spontaneous activity.

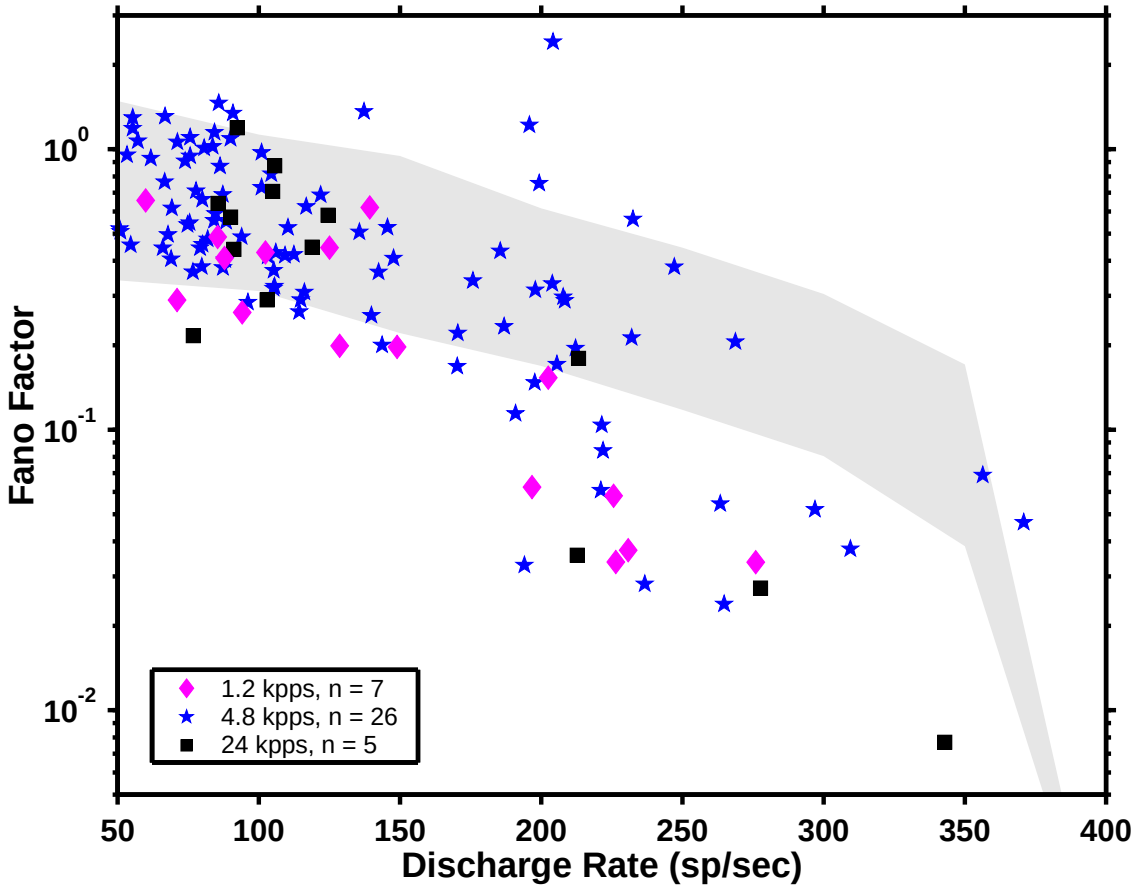


Figure 6. Fano Factor (which characterizes variability in the stimulus from presentation to presentation) as a function of discharge rate. Measurements were made in the analysis window determined as described in the Methods section. The gray area represents the 99% confidence interval for the distribution expected from a Poisson process.

4. Interval histograms

As indicated in Fig. 3, the shape of the interval histogram (IH) can depend on pulse rate. For the lowest pulse rate (1.2 kpps), IHs have an exponential envelope (Fig. 7, upper inset). An exponential shape is expected for Poisson discharges and is approximately consistent with the IHs for spontaneous activity in an intact ear. For high pulse rates, some but not all IH envelopes are clearly non-exponential, showing a sharp mode followed by a long tail (Fig. 7, left inset). To quantify the shape of the interval histogram, we fit the interval histogram with both a single exponential (dashed line in the insets) and piecewise, with three exponentials (solid line in the insets). The numerical procedures used in fitting the data are described in the Appendix. We measured the root mean squared error of each fit to the data, and defined an IH exponential shape factor (IH-

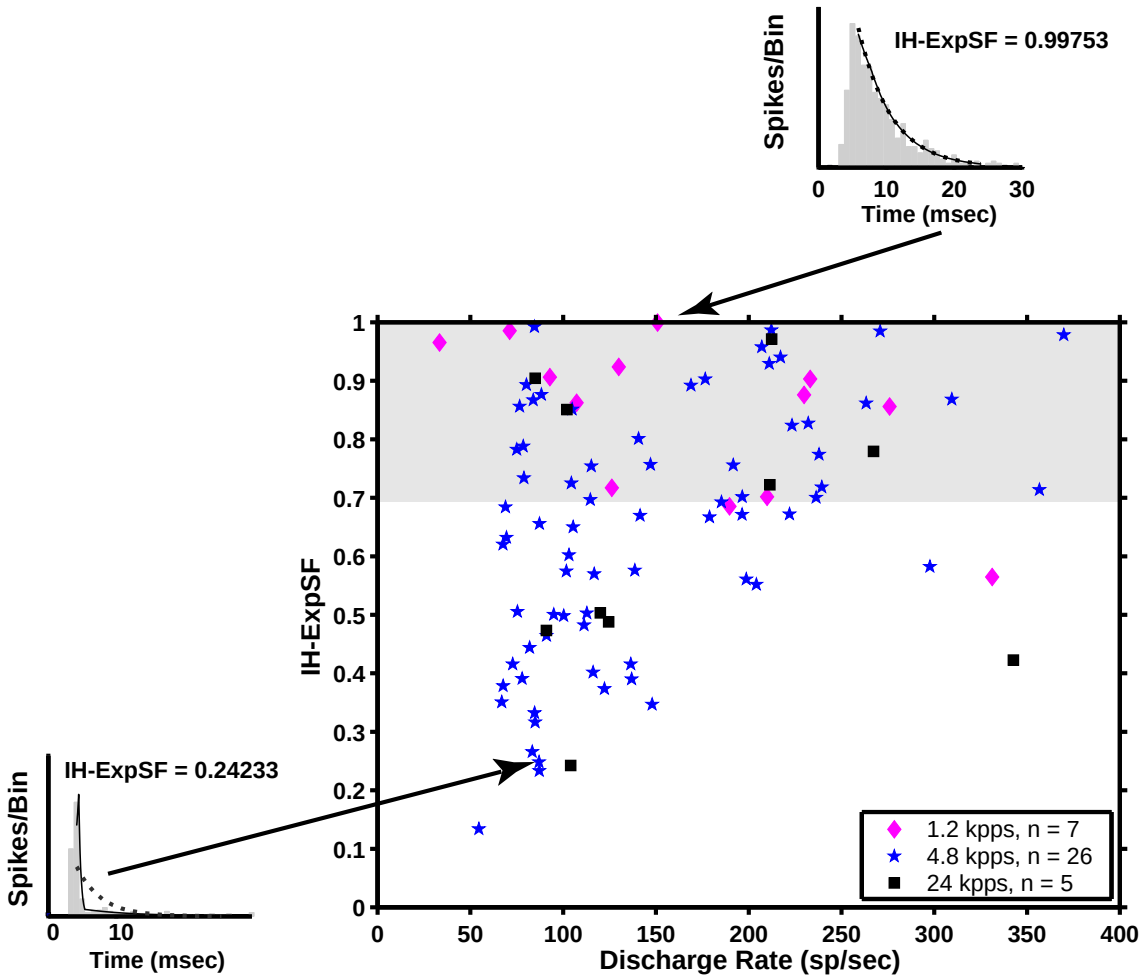


Figure 7. Distribution of the IH-ExpSF as a function of the average discharge rate in the steady-state portion of the response. The filled area shows the region that would contain 99% of the units if their responses could be described by the Poisson process (see appendix).

ExpSF) as the ratio of the error for the piecewise fit to that for the single exponential fit. The IH-ExpSF for samples from a Poisson process is approximately 1.

The scatter plot of Fig. 7 shows IH-ExpSF versus average discharge rate for measurements made with pulse rates of 1.2, 4.8 and 24 kpps. For a pulse rate of 1.2 kpps, most points fall in the region expected for a Poisson model (shaded area, Appendix 1). For higher pulse rates, 50% of the data points are outside of the range expected for a Poisson process. Thus, only lower pulse rates consistently produce IHs that resemble spontaneous activity in intact ears.

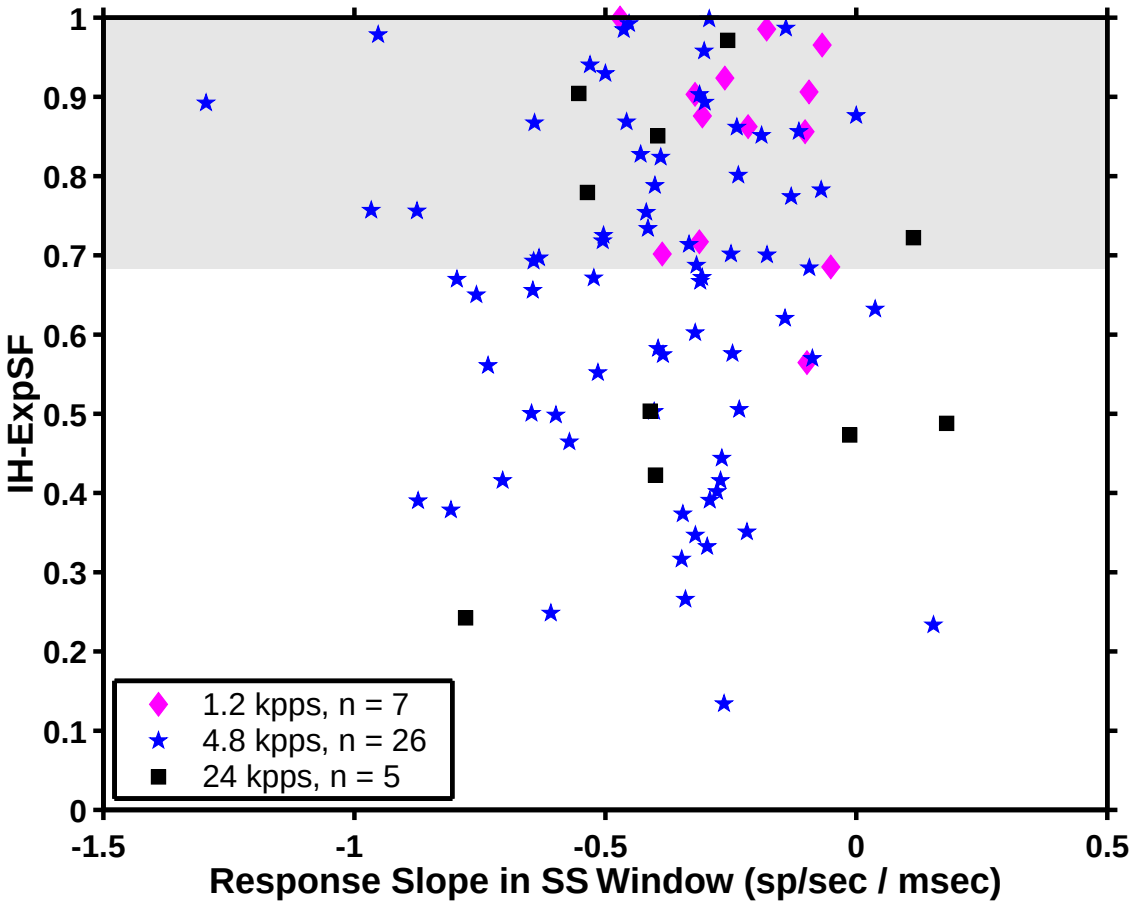


Figure 8. Relationship between the measured IH-ExpSF and the adaptation that occurs within the analysis window. Adaptation is measured as the slope of the regression line to the PST histogram computed in the analysis window (bin width of PST computation 10 ms). The analysis window (referred in the legend as the SS window) is a window over which the discharge rate is considered approximately constant.

One possible explanation for the difference between the 1.2 kpps and higher rate responses is the difference in adaptation during the analysis window (the window over which the rate is assumed to be approximately steady-state). To test this possibility, Fig. 8 plots the IH-ExpSF versus adaptation in the analysis window. Adaptation is defined here as the average decrease of discharge rate. There is no obvious trend, suggesting that adaptation in the analysis window does not significantly alter the IH.

IV. Modulated pulse trains

Fig. 9 shows responses from a single unit to a sinusoidally amplitude modulated (400 Hz) pulse train (4.8 kpps) for modulation depths of 1% (left) and 10% (middle). These modulation depths might be representative of the modulation depths that would be used

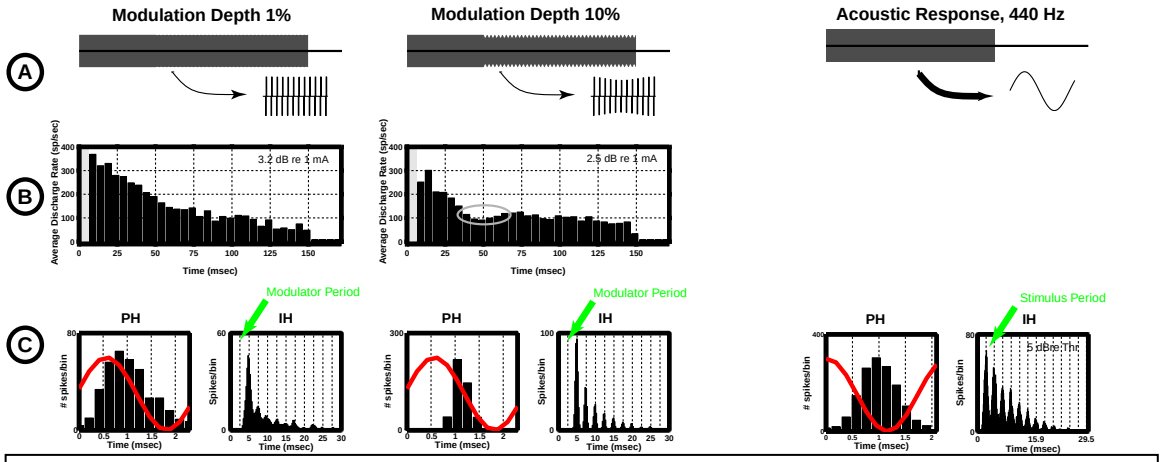


Figure 9. The first two columns show the response of a neuron to a modulated pulse train for two modulation depths (1% in the left and 10% in the middle). For comparison, the right panel illustrates a response of a neuron (CF near 400 Hz) to an acoustic stimulus. Row A shows the stimulus waveforms. In the case of the left and the middle columns, the stimulus is unmodulated for the first 50 ms and is down modulated during the last 100 ms. Row B shows period histograms of the responses to the two electric stimuli. For the 10% modulation depth, the response actually increases after the onset of the modulation (gray oval). Finally, row C shows the period and the interval histograms computed during the modulated portion of the stimulus for the two electric stimuli. These histograms are also plotted for the responses of a fiber to the acoustic tone. Both electric responses are broadly similar to the acoustic response in their temporal properties.

in a DPT-enhanced strategy. The smaller modulation depth is comparable to the psychophysical threshold to modulation in cochlear implant patients (Shannon, 1992). Pulse trains were modulated during the last 100 ms of the 150 ms train duration (row A). The levels of the two stimuli were adjusted to produce similar response rates during their modulated segment. For both stimuli, the response adapts over the first 50 ms while the pulse train is unmodulated. At the onset of modulation, average discharge rate increases for the modulation depth of 10% (row B, middle column; also, Fig. 10). This increase in rate is interesting because the RMS current actually decreases (by 0.9 dB for 10% modulation) when modulation begins since the peak amplitude remains constant. Row C shows period and interval histograms computed from the responses measured during the modulated portion of the stimulus. For comparison, the right column shows both the interval and the period histogram computed from responses to a 440 Hz tone at a moderate level in a normal ear (from (McKinney, 1998)). For both modulation depths, the period histograms show pronounced modulation, although spikes are more precisely phase locked for the higher modulation depth. Even when the modulation depth is only 1%, the response is nearly fully modulated. Furthermore, for this modulation depth, the

period histogram is nearly sinusoidal in shape. This suggests that the responses may be representing the details of the sinusoidal modulator waveform.

Phase locking can be seen in the interval histogram as the clustering of modes around multiples of the stimulus period, which are shown here as dashed lines. For both modulation depths, the mode distribution is broadly similar to that for tones. Close examination reveals several differences. First, a pronounced mode at the modulation period is absent in the electrical case (three arrows). Secondly, the mode at twice the modulation period is strongly exaggerated for the smaller modulation depth. This exaggeration may be related to the preferred interval near 5 ms found for unmodulated pulse trains.

1. Large modulation depth: entrainment

Fig. 10 shows the response of another unit to two different levels of a pulse train modulated at a depth of 10%. The pulse train was 250 ms long, and was modulated only in the last 100 ms. The increase in rate at the onset of modulation is more pronounced for this unit than for the unit in Fig. 7. For this modulation depth, we observed increases in rate at the onset of modulation in 80% of the units studied.

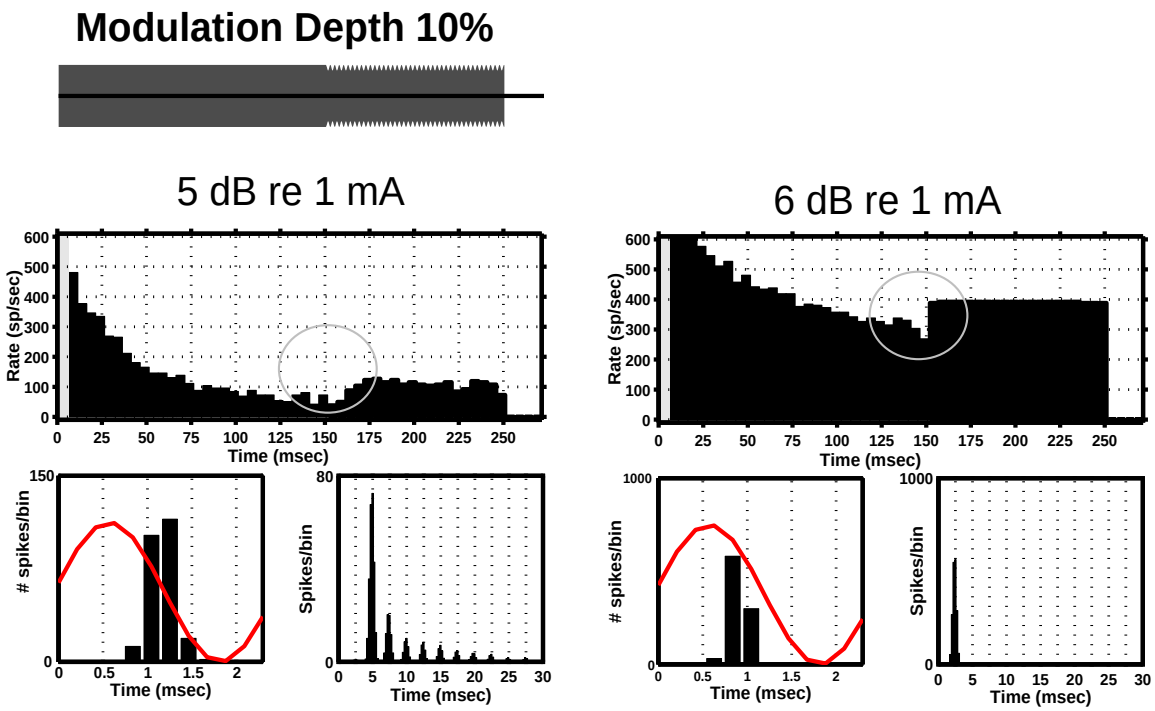


Figure 10. Response of a fiber to a stimulus with a large (10%) modulation depth at supra-threshold (left panels) and saturation (right panels) levels. The top-most plot shows the stimulus. The middle left and right plots show responses at the lower (left) and at the higher (right) levels. Note that in both cases the rate increases after the modulation onset. The bottom row of plots shows the period (left) and the interval (right) histograms at each level for the responses measured during the modulated portion of the stimulus. Note that at the higher level, the neural response is entrained to the stimulus.

At the lower level, the distribution of the modes in the interval histogram is similar to that for the acoustic responses to tone shown in Fig. 9. However, the mode at the stimulus period is again missing in the response to the electric stimulus. At the higher level, responses entrain to the modulator frequency, as indicated by a single mode at the modulation period in the IH.

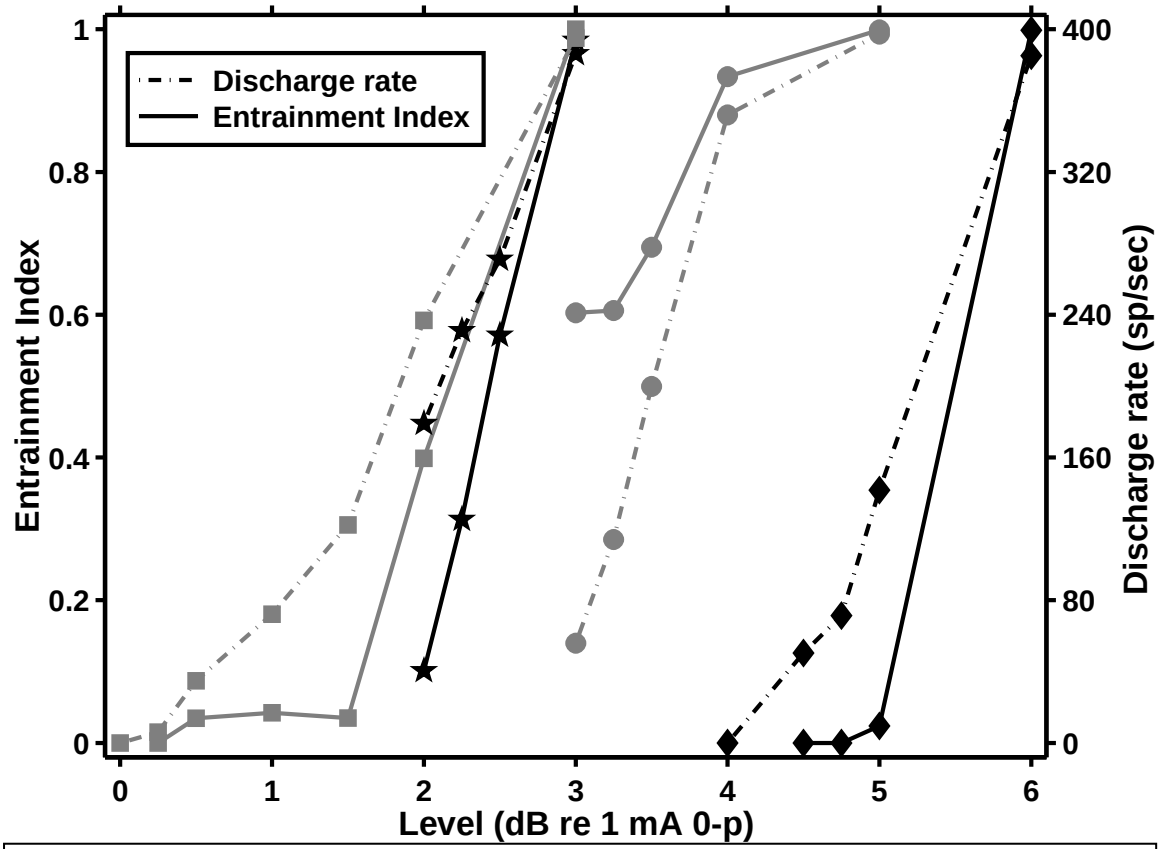


Figure 11. Entrainment index (solid line) and discharge rate (dashed line) as a function of stimulus level for four units from the same animal. Entrainment index is defined as the ratio of intervals that fall between 1/2 and 3/2 of the stimulus period to the total number of intervals. An entrainment index of 1 represents perfect entrainment. At the highest levels, all four fibers entrain to the stimulus. This entrainment does not occur in a healthy ear in response to pure tones. At stimulus levels below those that evoke discharge rates of 200 spikes/s, the entrainment index is 0 for two units, indicating that the mode at the stimulus period is missing in the units' responses. In contrast, responses to pure tones in a healthy ear have entrainment indices between 0.6 and 0.9 (Joris et al., 1994).

To quantify the difference between acoustic and electric interval histograms, we computed an *entrainment index* (Joris et al., 1994). The entrainment index is defined as the ratio of the number of the intervals that are between 1/2 and 3/2 of the stimulus period to the total number of intervals in the interval histogram. Fig. 11 plots the entrainment index (solid line) along with the average discharge rate (dashed line) as a function of stimulus level for four units from one animal. The stimulus was the same as in Fig. 10. At levels that evoked a moderate discharge rate (less than 200 spikes/s), two out of four units had an entrainment index near zero, indicating that the mode at the stimulus period was either very small or entirely missing from the interval histogram of these units. In contrast, the entrainment index computed from responses to tones recorded from ANFs in

a healthy ear is between 0.6 and 0.9 for tone frequency near 400 Hz (Joris et al., 1994). When the stimulus level was at least 2 dB over the level that evoked a discharge rate of 100 spikes/s, units entrained to the modulator (had an entrainment index of 1). Such entrainment is never seen in the responses of ANFs in an intact ear to tones (Joris et al., 1994; Rose et al., 1967). Thus, except for an extremely small range of levels, interval distributions for responses to electric modulated pulse trains can differ from the distributions for responses to pure tones recorded in an intact ear.

2. Small modulation depth

Fig. 12 shows average discharge rate and response modulation depth (which is a measure of modulation of the period histogram, and is twice the synchronization index) for a single unit as a function of level. The stimulus was a 250 ms pulse train (4.8 kpps) that was modulated by a 400 Hz sinusoid in the last 100 ms (modulation depth 1%). Average discharge rate increased 3-fold over the 1.5 dB range of levels. In contrast, response modulation depth was nearly constant at 0.7-0.8 for all levels, indicating robust phase locking. The first mode in the interval histogram shifts to the left as the level increased, reflecting the increase in average rate (lower panels). While the first mode is a multiple of the modulator period at the lowest and at the highest level shown, at the intermediate level of 4.75 dB re 1 mA, it falls nearly halfway between the first and the second multiple of modulator period (arrow). This result suggests that at this level, the first mode is not related to the modulator frequency. We hypothesize that for the higher levels, the first mode is related to the preferred firing periods demonstrated earlier for high-frequency unmodulated stimuli. To test this hypothesis, we computed IHs from responses that immediately precede the modulation onset (gray line). As predicted by the hypothesis, the location of the mode of the IH for these responses roughly matches the mode observed during the modulated segment at levels above 4.75 dB re 1 mA. At the lower level, however, the first mode of the modulated segment differs from the first mode of the pre-modulation responses and is related to the modulation frequency.

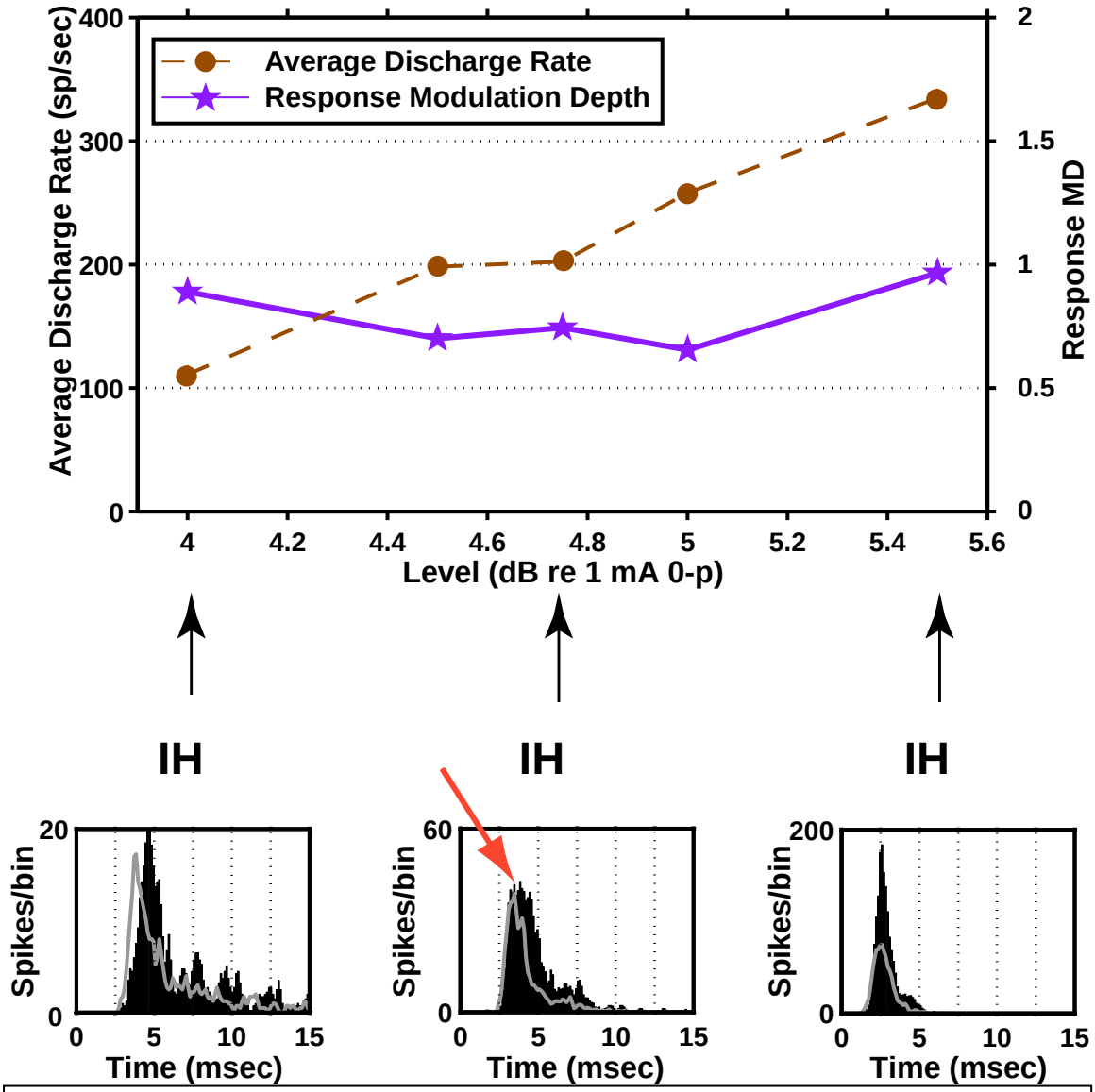


Figure 12. Responses of a unit to a stimulus with modulation depth of 1%, modulation frequency of 400 Hz, at several levels. The top panel shows both the response rate and the modulation depth of the response as a function of level. Both the response rate and the response modulation depth are computed during the modulated portion of the stimulus. Note that the modulation depth of the response is nearly independent of level. The lower panels show the IHs computed from the responses at three different levels. The solid areas represent the IH computed from responses during the modulated portion of the stimulus while the gray lines are the IHs computed from responses preceding the modulation. At the lowest level, the mode of the IH for the modulated portion is related to the modulation frequency. At the higher levels, however, the mode of the IH for the modulated portion is similar to the mode described earlier for the unmodulated responses and not to the modulation frequency.

Thus, for very low modulation depths, interval histograms can differ from IHs evoked by acoustic tones because their IHs can include modes that are not multiples of the modulator period.

V. Discussion

A. Adaptation

The adaptation that we report here is generally consistent with previous studies of electrically stimulated ANFs (Moxon, 1967; van den Honert and Stypulkowski, 1987b; Javel et al., 1987; Dynes and Delgutte, 1992; Killian, 1994). The result that at least short-term (<150 ms) adaptation depends on pulse rate of the electric stimulus is new. This result may have important consequences for models of electrically stimulated ANFs.

Our working hypothesis is that the slow decrease in discharge rate over time is evidence either of depletion of an "excitatory" agent or accumulation of an "inhibitory" agent. For example, intracellular sodium and extracellular potassium might be the relevant agents for the electrically stimulated ANFs. The changes in the concentrations of these substances may occur because of change in the membrane conductance resulting from the increased spike activity induced by the electric stimulus. Alternatively, concentrations may change even without increased spike activity because the electric stimulus can induce voltage changes across the neural membrane that do not lead to spikes. If the increased activity itself were the primary influence on accumulation or depletion of the agent, we would expect that adaptation would depend primarily on the initial response rate and not on the pulse rate of the electric stimulus. We showed that adaptation of ANFs to electric stimulation depends not only on the initial rate, but also on stimulation frequency. This implies that significantly more agent is depleted due to voltage changes across the membrane that do not lead to spikes during the higher frequency stimulus. Whether neurons will be able to readjust to the new homeostatic balance if the high-frequency DPT is presented continuously is a topic for future investigations.

B. Temporal response patterns to unmodulated pulse trains

We found that the detailed response pattern of auditory-nerve fibers to electric pulse trains also depends strongly on stimulation rate. In particular, while the envelope of the IH of responses to a 1.2 kpps pulse train is nearly exponential, the envelope can strongly

differ from an exponential for pulse rates above 4.8 kpps. This conclusion is consistent with published data. Van den Honert and Stypulkowski (1987b) described ANF responses to moderate-level electric sinusoids with frequencies between 100 and 1000 Hz. The interval histograms in these data appear exponential. IHs reported for responses to low-rate (<2000 pps) pulse trains appear exponential as well (Javel et al., 1987; Javel, 1989). Dynes and Delgutte found Gaussian-like interval histograms of responses of ANFs to electric sinusoids for frequencies above 4 kHz (Dynes and Delgutte, 1992). The mode of the their IHs is at about 5-6 ms, as in our data. Unfortunately, they do not report IHs for responses with discharge rates below 200 spikes/second, where the tail may be more apparent.

C. Does the response to a desynchronizing pulse train (DPT) resemble spontaneous activity?

Rubinstein et al. (1999b) suggested introducing a continuous pulse train into CI strategies to produce neural responses resembling normal spontaneous activity. Our results indicate that if the level of the DPT can be adjusted for each fiber, the responses to sustained high-frequency pulse trains resemble spontaneous activity in some, but not all respects. The variability across stimulus presentations is comparable with that expected for spontaneous activity. For a relatively low pulse rate (1.2 kpps) the envelope of interval histograms resembles those for spontaneous activity for most units. However, for higher rates (4.8 kpps and above) interval histograms can clearly deviate from those for spontaneous activity, showing a sharp mode that is followed by a long tail. On the other hand, a DPT with a low pulse rate may introduce psychophysically significant periodicity in neural responses that is related to the DPT period. Thus, intermediate pulse rates (above 1.2 and below 4.8 kpps) may be optimal for imitating spontaneous activity.

A further difficulty with the DPT idea is that electrically-stimulated fibers respond with discharge rates that are appropriate for spontaneously responding neurons in a healthy ear only over a narrow range of stimulus levels. The small dynamic range reported in this study is consistent with earlier reports of responses of ANFs to electric pulse trains (Javel et al., 1987; Parkins, 1989). Fibers can differ in threshold by 10 dB or more at a single cochlear location near the stimulating electrode (van den Honert and

Stypulkowski, 1987a). If a DPT is presented at a level that stimulates a large percentage of the fibers, most fibers will respond at rates that are higher than the rates appropriate for producing spontaneous-like responses.

While the differences in DPT sensitivity across fibers are large compared to the level range producing spontaneous-like responses, the DPT idea should not necessarily be ruled out based on the results of this study of short-term (less than 250 ms) stimulation. Electric responses are known to adapt over a course of minutes (Moxon, 1967; Killian, 1994). This longer adaptation may selectively decrease the response discharge rates of fibers with low thresholds to the DPT.

D. Does a high-frequency DPT help encode modulation frequency?

We showed that a modulated pulse train with low (<10%) modulation depth can produce interspike and period histograms resembling responses to tones in intact ears. If we interpret those stimuli as consisting of a DPT plus a highly-modulated signal, this result suggests that realistic responses to sinusoidal stimuli might be obtained with a DPT. However, the realistic, tone-like responses are only observed over a narrow range of stimulus levels. Furthermore, for very low (1%) modulation depths, there can be preferred intervals unrelated to the modulation frequency which may be confusing to the central processor. Interestingly, these intervals may co-exist with a high level of phase locking to the modulator frequency.

E. Implications for Mechanisms

Our results with unmodulated and modulated stimuli also pose important questions for models of ANF responses to electric stimulation. Published reports based on biophysical models have not described the complex shape of interval histograms (a pronounced mode with a long tail) we observed in responses to high-frequency pulse trains (e.g. (Rubinstein et al., 1999b)). In evaluating mechanisms that may account for this discrepancy it is important to determine whether the relevant processes occur on an interval-by-interval basis or on a slower time scale (e.g. bursting). In a limited number of units for which sufficient data are available, we failed to detect a correlation between consecutive intervals. Thus, the relevant processes appear to occur on an interval-by-interval basis.

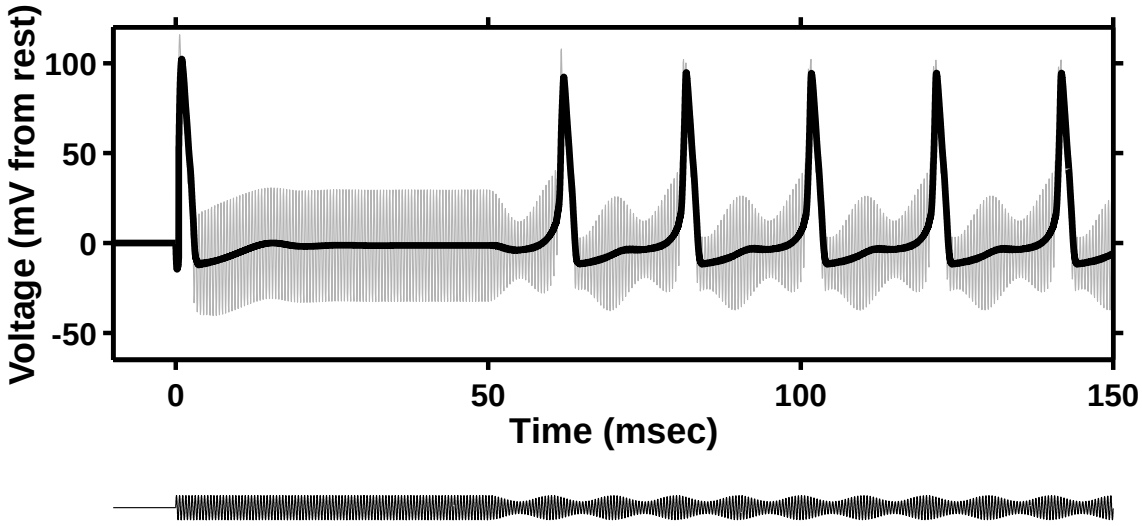


Figure 13. Response of a current-clamped Hodgkin-Huxley model to a high-frequency sinusoid that is modulated in the last 50 ms of the 100 ms stimulus (bottom). Parameters for this model are summarized in Weiss (1996), p. 191. Simulation temperature was 6.3ºC. The model was implemented in MATLAB and simulated numerically using a 5 μs step. The stimulus was presented as intracellular current injections. The carrier was a 2 kHz sinusoid (400 uA/cm^2 0-peak) that was modulated "down" (sinusoidal 50% modulation, 100 Hz modulation frequency). This stimulus was chosen to represent the modulated pulse trains that were used in the physiological study. However, both the shape of the "pulse" and the carrier frequency were adjusted to account for the slower dynamics of the Hodgkin-Huxley model. The time course of the membrane voltage is shown by the gray line. The membrane voltage is shown relative to its resting state. We also show the time course of the membrane voltage averaged over a period of the sinusoidal carrier (dark line).

Responses that are qualitatively similar to the ones that we report for cat ANFs, in which rapid firing randomly alternates with longer pauses, have also been observed in the squid giant axon that is stimulated by a DC current injection (Guttman and Barnhill, 1970; Guttman et al., 1980). The mechanism for these responses may be similar to the mechanism responsible for the non-exponential responses reported here. A biophysical model that has Hodgkin-Huxley channel dynamics, and that explicitly models random channel noise can account for the observed responses in that preparation (Schneidman et al., 1998). Therefore, it is possible that a biophysical model of ANFs that correctly captures ANF channel dynamics and that explicitly models channel noise may predict the responses that we observed as well.

Although high-frequency pulses used in this paper differ from DC injections used in the study of the squid giant axon, some parallels may exist between the responses evoked by both. In particular, rectification and low pass filtering, both of which are known to

occur in neural membrane, may transform the extracellular pulse train into an intracellular stimulus with a significant DC component. Intracellular recordings near the site of excitation are necessary to determine mechanisms underlying the observed responses.

Another intriguing observation is the increase in average discharge rate after the "down" modulation is turned on. The increase in average discharge rate is surprising because the RMS stimulus current actually decreases at the onset of the modulator. To our knowledge, such an increased response to modulation has not been reported for biophysical models. Our simulations using the Hodgkin-Huxley (HH) model indicate that an increased response rate can be observed at the onset of the “down” modulation of a high-frequency carrier (Fig. 13), although the model does not mimic the data quantitatively.

One difference between the HH model and the data is that the HH model exhibits an onset-only response to the sustained high-frequency stimulus. It is unknown whether introducing channel noise into the model would improve the fit between the model and the data. While Frankenhaeuser-Huxley model (Frankenhaeuser and Huxley, 1964) of the frog sciatic nerve predicts the sustained response during the high-frequency stimulus, it does not show an increase of firing rate at the onset of the down-modulation (personal observations). Future modeling studies would also be needed to determine whether mammalian-based models produce responses to modulated stimuli that more closely resemble our data for ANFs.

F. Implications for the cochlear implant processor

The purpose for introducing a DPT is to improve the coding of modulation in auditory nerve responses. Some aspects of our data appear promising in that respect. A DPT can produce responses that are desynchronized across trials, suggesting that the responses may also be desynchronized across different fibers, as occurs for normal spontaneous activity (Johnson and Kiang, 1976). Desynchronization of auditory nerve responses may lead to an improved temporal coding of stimuli with rapid onsets and high frequencies. For low pulse rates of 1.2 kpps, responses to a DPT imitate some characteristics of

spontaneous activity. Over a range of levels and modulation depths, a DPT can help provide temporal discharge patterns for amplitude-modulated stimuli that resemble responses to tones. Other aspects of the data are less promising. For example, we found that DPTs that accurately encode sinusoidal waveforms in the period histogram can show modes in the interval histogram that are not related to the modulator frequency. In addition, for some modulation frequencies and levels, the mode of the interspike distribution that corresponds to the stimulus period can be entirely absent in the interval histogram of the response. It appears, therefore, that successful use of a DPT depends on what exact aspects of the temporal discharge pattern (e.g. intervals versus cross-fiber synchrony) are most important to the central processor for extracting information about the stimulus. Because the exact mechanisms used by the central processor are not known, it is difficult to predict whether adding the DPT will result in an overall improvement in the reception of speech by cochlear implant users.

VI. Appendix

The interval histogram exponential shape factor (IH-ExpSF) is defined as the ratio of the error in the fit of the IH with the piecewise, three exponential, function to the error of the fit with a single exponential. First, a number drawn from a uniform distribution from 0 to 0.832 ms was added to each interval to eliminate the effect of locking to the 1.28 kHz pulse frequency. Next, the mode of the interval distribution was estimated. Intervals that were longer than the mode were considered to constitute the tail and were included in the analysis; the remaining intervals were discarded.

Next, the IH was recomputed based on the remaining intervals. Instead of computing the interval histograms with a fixed bin size, however, we changed the bin size from bin to bin, such that exactly 5 intervals fell into each bin. This normalization is convenient because the coefficient of variation of the interval histogram for any given bin is proportional to the number of values that fall into that bin. (Johnson, 1996) By normalizing the interval histogram in this way, one achieves equal coefficient of variation in each of the bins. Finally, a logarithmic transformation was applied to the IH so that straight-line fits generated exponential fitting function.

The resulting logarithmic histogram was fit by a linear and by a piecewise linear function by minimizing the least squares error. The piecewise linear function was constrained to be continuous. The times at which the function changed its slope were chosen such that each segment covered an equal number of bins.

To estimate IH-ExpSF distribution expected from a Poisson-with-dead time model, the IH-ExpSF was computed for 1000 interval histograms, each histogram composed of 200 independent “intervals” drawn from an exponential with dead time distribution (dead time of 3 ms). 99% of the generated data fell between 0.68 and 1. This boundary was nearly independent of the event rates of the simulated distribution for rates of 50 to 250 events/s.

Chapter 3

Responses of auditory-nerve fibers to sustained, high-frequency electric stimulation. I. Unmodulated pulse trains

I. Introduction

In continuous interleaved sampling (CIS) sound-processing strategies for cochlear implants, temporal information about incoming sounds is encoded in the amplitude modulations of pulse trains (Wilson et al., 1991). Proper representation of these modulations in the temporal discharge patterns of the auditory nerve is therefore an important goal of these strategies.

Despite the popularity of CIS schemes, the responses of auditory-nerve fibers to a modulation of an electric pulse train can be very different from responses in an acoustically stimulated ear. For example, sinusoidally modulated electric pulse trains with modulation frequencies below 500 Hz may elicit a response on every modulator cycle (van den Honert and Stypulkowski, 1987b). In contrast, in response to an acoustic pure tone, neurons fire at random multiples of the stimulus period, in that there may be 1, 2, 3 or more cycles between successive spikes (Rose et al., 1967). The situation is even worse at higher frequencies, because, with electric stimulation, neurons may fire exactly on every other cycle or even higher multiples of the modulation period. If most stimulated neurons fire together, then the population of auditory neurons would code a submultiple of the modulator frequency rather than the actual frequency. Recent measurements of auditory-evoked potentials suggest that this may be the case (Wilson et al., 1997).

Rubinstein et al. (1999b) proposed that the coding of modulation waveforms might be improved by introducing a sustained, high-frequency, "desynchronizing" pulse train

(DPT) in addition to the basic processor output. The rationale for the DPT is that across-fiber differences in refractory, sensitivity and other properties, as well as noise present in the neural membrane will result in the responses across fibers being desynchronized after the first few hundred milliseconds of DPT stimulation. Such desynchronization would lead to an improved representation of the modulator in temporal discharge patterns by allowing an ensemble of neurons to encode the true modulator frequency rather than a submultiple.

Rubinstein further argued that responses to the DPT might mimic the spontaneous activity recorded in a healthy ear (Rubinstein et al., 1999b). Because fibers with higher spontaneous activity typically have lower thresholds, spontaneous activity is thought to enhance detection of weak periodic signals. The mechanism for enhancement may be related to the phenomenon of stochastic resonance (Wiesenfeld and Moss, 1995). If spontaneous activity could be restored by the DPT, then similar benefits could be provided to cochlear implantees. Finally, Kiang and Levine (1970a) suggested that some forms of tinnitus (sounds in the ear without external stimulation) are related to the absence of spontaneous activity in a portion of the cochlea. By restoring spontaneous activity using electric stimulation, it may be possible to reduce tinnitus in these patients.

While Rubinstein's proposal involves sustained stimulation with a high rate, electric pulse train, only a few studies have investigated the responses of auditory-nerve fibers (ANFs) to such stimuli. Moxon (1967) investigated responses of ANFs of cats to sustained (200 second long), 900 pps electric pulse trains. He showed that fibers' discharge rates were large initially, but decreased markedly throughout the stimulus. The responses were stochastic and skipped one, two or more stimulus pulses in an irregular sequence. In their stochastic nature, these responses were similar to normal spontaneous activity (Kiang et al., 1965). However, discharge rates could exceed 500 spikes/s, which is much higher than the rates reported for spontaneous activity in a healthy ear, which are always below 150 spikes/s (Liberman, 1978). It is unclear whether the higher pulse-rate (5 kpps) DPT proposed by Rubinstein (1999b) would lead to greater adaptation of the neural response, and therefore to discharge rates more consistent with spontaneous activity.

In the previous report (Litvak et al., 2001), we investigated responses to short (200 ms) DPTs. We found that ANF responses to a DPT were stochastic, resembling spontaneous activity, but that discharge rates below 150 spikes/s could only be achieved over a narrow (2-3 dB) range of levels. However, fiber threshold (defined as the level of the DPT at which the discharge rate was near 100 spikes/s) differ by as much as 10 dB across fibers. It is unclear whether the sustained discharge rate evoked by a long duration unmodulated DPT would be less dependent on the DPT level relative to the fiber's threshold. In addition, safety of sustained DPT stimulation was not addressed in our previous study.

In the first part of this paper, we report the results of experiments designed to test whether responses of single ANFs to a sustained DPT resemble spontaneous activity in a healthy ear. In particular, (1) the discharge rates observed after 5-10 minutes of DPT stimulation are compared to the distribution of discharge rates for spontaneously responding ANFs (Liberman, 1978), (2) interval histograms (IH) computed from DPT responses are compared to the histograms reported for spontaneous activity (Kiang et al., 1965), and (3) cross-fiber correlations computed from pairs of simultaneously recorded DPT responses are compared to those reported for spontaneous activity (Johnson and Kiang, 1976).

In the second part of this paper we report electrically-evoked compound action potentials (ECAPs) for a single-pulse probe following sustained DPT stimulation. Several studies have reported changes in excitability of the auditory nerve following sustained high-frequency electric stimulation (Duckert and Miller, 1982; Shepherd, 1986; Killian, 1994; Tykocinski et al., 1995b; 1997). While a truly sustained DPT may be unsafe, a DPT that turns on and off every 10 minutes may be less traumatic to ANFs. To test this hypothesis, we recorded ECAPs following a 10-minute DPT. We compared these ECAPs to ECAPs after 1-hour of DPT stimulation. We also extend the previous findings to stimulus frequencies that are appropriate for a DPT (greater than 2 kpps).

In a companion paper, we address the question of whether responses of ANFs to modulated electric pulse trains presented during a sustained DPT resemble ANF responses to corresponding acoustic stimuli.

II. Methods

A. Animal preparation

The animal preparation used in this study is similar to that described by Litvak et al. (2001). Cats were anesthetized using dial in urethane (75 mg/kg). The bulla was opened to expose the round window. An intracochlear stimulating electrode was inserted about 8 mm into the cochlea through the round window. The electrode was a Pt/Ir ball (400 μm in diameter). A similar ball electrode was inserted just inside the round window and used as the return electrode in most animals. In other animals, the ear bar was used as the return electrode. The round window was then sealed using connective tissue.

Co-administration of kanamycin (subcutaneous, 300 mg/kg) followed by slow intravenous injection of ethacrinic acid (1 mg/min) was used to deafen the animals (Xu et al., 1993). A 100 μs condensation click was used to measure compound action potentials (CAPs), which were continuously monitored in the non-implanted ear while ethacrinic acid was administered. CAP threshold, which was defined as the lowest-level of the acoustic click evoking a visually-detectable CAP response, exhibited a rapid drop within 1 hour of initiation of ethacrinic acid injection. However, unlike Xu et al. (1993), we failed to achieve complete deafening in all cases, even with the maximum safe dose of 40 mg/kg. Residual CAP thresholds ranged from 80 to >120 dB peak SPL, but often dropped slowly during the course of the experiment (which lasted for two to three days).

B. Stimuli

In the initial experiments, we recorded responses from single ANFs to unmodulated, 10 minute, 5 kpps electric pulse trains. Pulse trains were composed of biphasic (cathodic/anodic) pulses (25 μs per phase). In later experiments, the unmodulated pulse trains were replaced by pulse trains that were modulated every second for 200-400 ms. Modulation depth never exceeded 15%. Details of the modulation waveforms are described in the companion paper (Litvak et al., 2002a). For the purpose of this paper, only responses recorded during the unmodulated portions were analyzed. These responses did not differ systematically in their distribution of discharge rates from the responses to the unmodulated pulse train. We will refer to the 5 kpps electric pulse train as a “desynchronizing pulse train” or DPT.

Stimuli were synthesized digitally using a 40 kHz, 16 bit D/A converter, and presented through a custom-built high-bandwidth stimulus isolator. All stimulus levels reported in this study are peak currents.

A single level of the DPT was used for each animal. To maximize the number of units stimulated by the DPT, its level was selected to be sufficiently large to stimulate a high percentage of the fibers. To this purpose, electrically-evoked compound action potentials (ECAPs) were used to assess the sensitivity of auditory nerve to electric stimulation for each animal prior to the beginning of the single-unit recordings. The ECAP electrode (a 400 μm diameter Pt/Ir ball) was placed on the auditory nerve near the internal auditory meatus. The stimulus used to evoke the ECAP was a biphasic electric pulse (25 μs per phase) presented at a repetition rate of 15/second. The initial polarity of the pulse alternated from cathodic to anodic with each presentation. Because the stimulus artifact is reversed by a polarity change, while the neural response remains of the same polarity, most of the stimulus artifact is cancelled by averaging the responses to the two polarities. To further reduce noise, the response was computed by averaging 30 single-pulse responses.

C. Single-unit recordings

Standard techniques were used to expose the auditory nerve via a dorsal approach (Kiang et al., 1965). We recorded from single units in the auditory nerve using glass micropipettes filled with 3M KCl. A digital signal processor (DSP) was used to separate neural responses from the *stimulus artifact* (stimulus-related voltage excursions recorded at the micropipette that are not due to neural discharges) during the unmodulated stimulus segments in real time. For 5 kpps unmodulated pulse trains, a digital boxcar filter with duration equal to the period of the pulse train removed most of the stimulus artifact while preserving the spike shape. Fig. 1 shows an example of applying artifact cancellation to the recorded signal during 5 kpps pulse train stimulation. Response waveforms were also streamed to disk for off-line analysis.

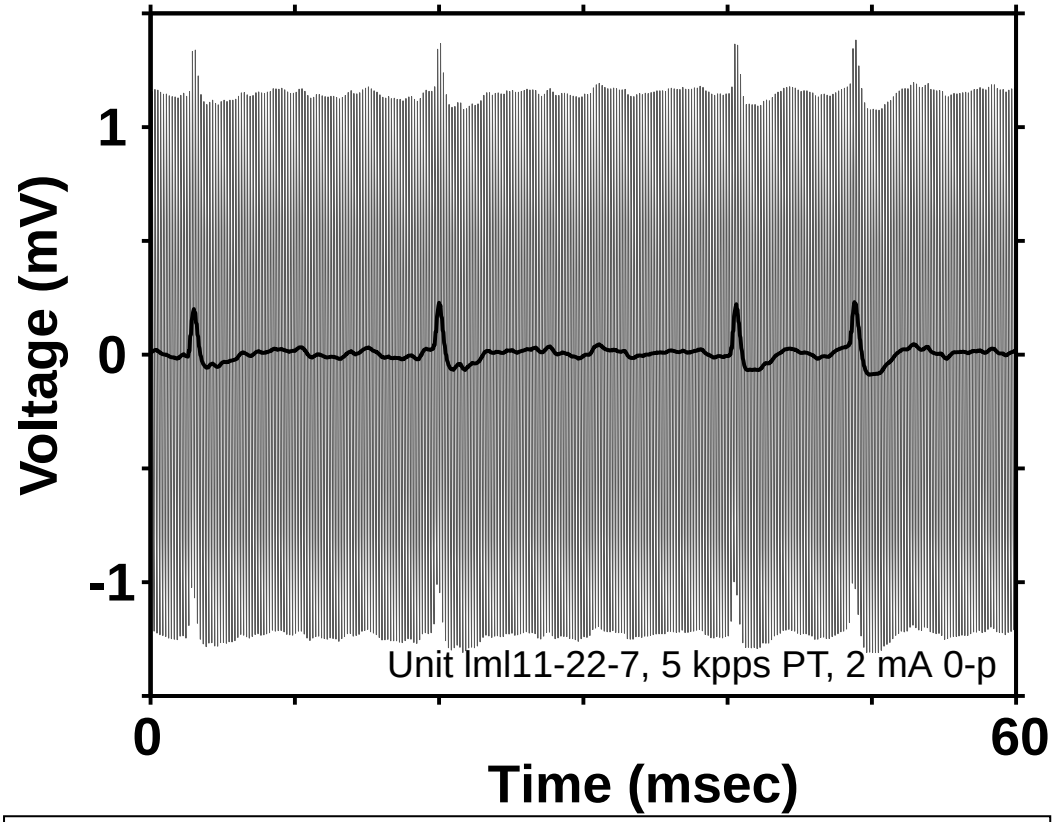


Figure 1. The gray trace is the signal recorded from the micropipette. This waveform contains spikes, as well as the 5 kpps stimulus artifact. The black line is the neural recording after artifact cancellation using an averaging filter. The artifact has been entirely removed, while spikes are preserved.

Threshold of each fiber to electric stimulation was routinely measured prior to DPT stimulation using a tracking procedure (Taylor and Creelman, 1967). Threshold was defined as the level of a 25 ms, 5 kpps electric pulse train that evoked a single spike during the last 10 ms of the stimulus for 50% of the presentations.

Both the auditory nerve and the vestibular nerve pass through the internal auditory meatus (Liberman and Kiang, 1978). The border between the vestibular and the auditory nerve can usually be visualized from our dorsal view. Stereotaxic and neuro-physiological criteria have been developed to identify vestibular nerve fibers (Liberman and Kiang, 1978) (see below). These criteria were only adopted in later experiments. Consequently, some of the recordings from the earlier experiments may be from vestibular fibers. This issue is discussed in Section III-E.

To quantitatively assess the possible influence of vestibular recordings on results of the earlier experiments, we deliberately recorded from both vestibular and auditory-nerve

fibers in some experiments. The recording microelectrode was oriented 45° from the horizontal plane and 10-15° lateral from the sagittal plane. We labeled a fiber as vestibular if (1) the point of entry of the electrode was anterior to the border between the vestibular nerve and the auditory nerve, and (2) the depth of the electrode was less than 1 mm from the nerve surface, and (3) a unit with regularly firing spontaneous activity was found deeper in the nerve. Since only vestibular units have regular spontaneous activity, detection of a unit with regular spontaneous activity is an unambiguous indication that the electrode is in the vestibular nerve, and not in the auditory nerve beneath (Walsh et al., 1972). A fiber was labeled as acoustic if (1) the point of entry of the electrode was posterior to the auditory/vestibular nerve border, or (2) the fiber was deeper than 1 mm from the surface of the nerve and no regular vestibular unit was found deeper than the fiber (we always looked at least 0.3 mm deeper). Neurons that could not be classified as either vestibular or acoustic were excluded from analysis.

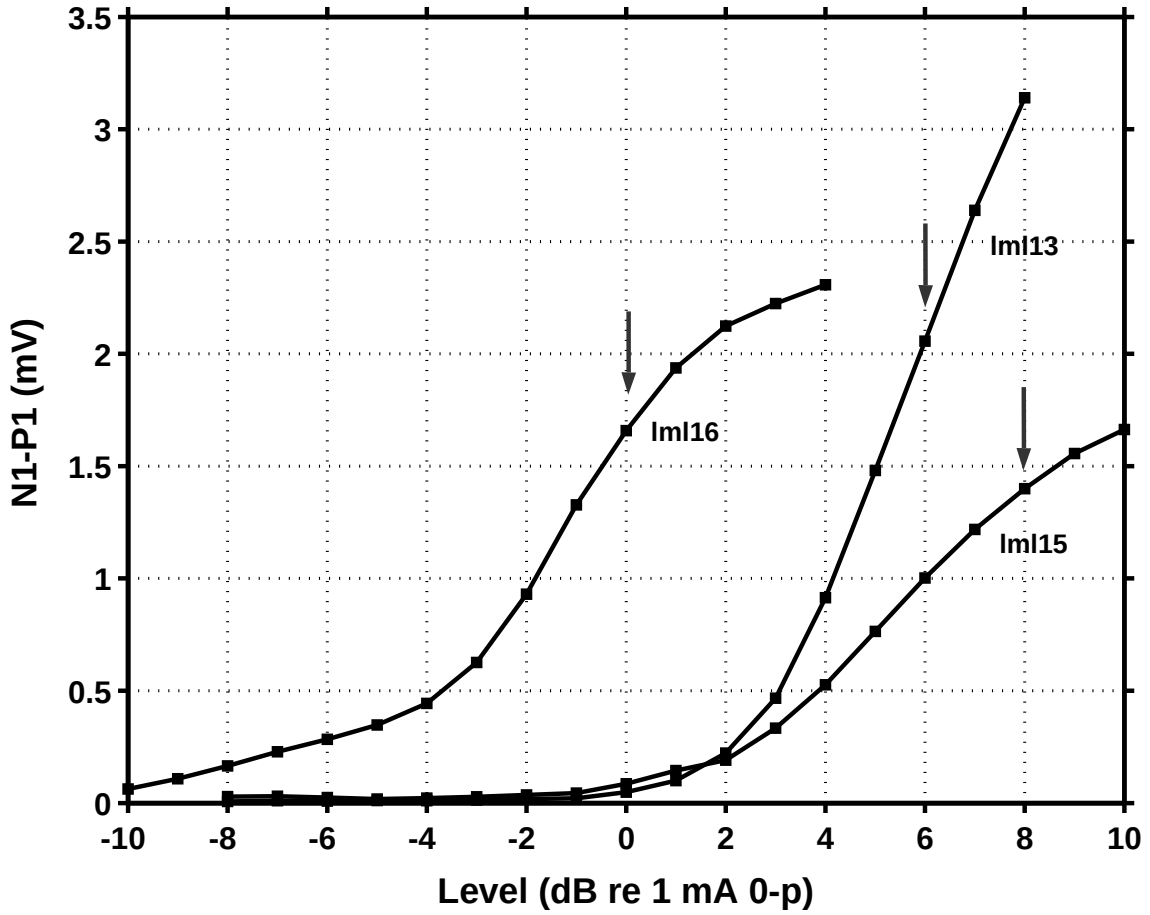


Figure 2. ECAP growth curve and DPT level for three animals that span the range of thresholds in our data set. Each point in represents the peak-to-peak ECAP response evoked by the alternating, biphasic pulse. The evoked response first grows slowly, then rapidly with stimulus level, and then slowly again. In most cases, the rate of growth then decreases. To normalize across animals, we chose the DPT level to be roughly 1 dB below the "knee" (arrow). In animals at which the knee was not apparent (e.g. lml13), DPT level was set at the highest stimulus level that did not evoke a twitch response.

Possible hair-cell mediated activity ("electrophonic hearing") may also complicate the interpretation of responses of ANFs to electric stimulation. Hair cells might not be completely eliminated by the acute deafening protocol used in this study. To minimize the effect of any remaining hair cells, only units without sustained spontaneous activity were included in this report. In about 10% of the units, we tested for a response to an acoustic click presented to the implanted ear at 90 dB SPL. Units that responded to this click were also excluded from the analysis.

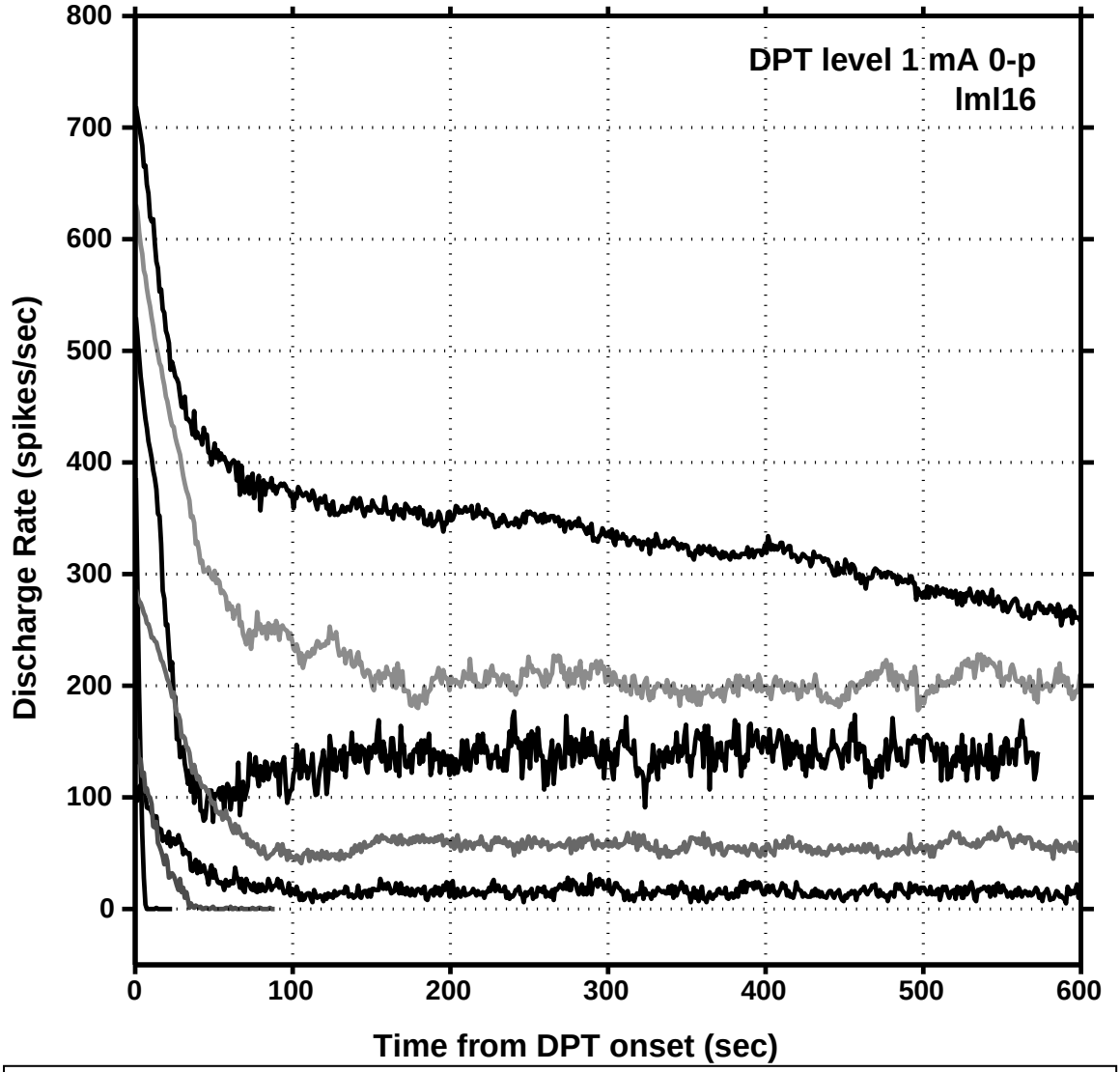


Figure 3. Discharge rate as a function of time in 6 units recorded from a single animal (lml16). The stimulus was a 10 minute, 5 kpps electric train of biphasic pulses presented at 1 mA 0-p. Rate was estimated over 1 s windows. Units adapted rapidly over the first two minutesThe rapid adaptation was followed by slower adaptation or a steady response. Two units stopped responding within a minute after the onset of the DPT.

III. Results

A. DPT level selection

To assess the sensitivity of the auditory nerve to electric stimulation, ECAPs were measured by averaging responses to a single, alternating-phase, biphasic pulse (25 μs/phase). Fig. 2 plots the peak-to-peak amplitude of the ECAP response as a function of pulse level for three animals. These curves were chosen to illustrate animal-to-animal

variability. All animals showed a segment of rapid ECAP growth. As the stimulus level increased further, the rate of growth typically slowed (lml16 and lml15). Slower growth in the ECAP suggests that fewer fibers are recruited by similar increases in pulse level. Assuming that the distribution of fiber thresholds is nearly Gaussian, levels near the region where the ECAP growth rate slows probably stimulate a large percentage of the fibers. Whenever possible, these levels were chosen as the DPT level (arrows). In some animals, levels near ECAP saturation evoked a facial twitch. In these cases, a DPT level just below the twitch threshold was selected.

B. Discharge rate distribution

We recorded responses to the DPT from 96 fibers in 8 cats for at least 3 minutes. These fibers had no spontaneous activity, and did not respond to sound. Depending on the animal, units with spontaneous activity or responses to sound constituted from 0 to 30% of the total number of units contacted.

Fig. 3 shows the discharge rate as a function of time from the onset of the DPT for 7 auditory-nerve fibers from one animal. The DPT level was 1 mA 0-p for all 7 units. While initially the discharge rates to the DPT could be extremely large (exceeding 700 spikes/s for some units), discharge rates decreased over several minutes of DPT stimulation. Most units showed rapid initial adaptation over the first 100 seconds of DPT stimulation, followed by either slower adaptation, or a steady response. In addition, 2 units stopped discharging after 1-2 minutes of DPT stimulation.

Fig. 4 plots the discharge rate distribution after 50 seconds of 5 kpps DPT stimulation for 190 fibers that initially responded to the DPT. In this figure, we pool together responses to the 5 kpps pulse trains, and the responses during the unmodulated parts of the partially modulated pulse trains. The dark line is the cumulative distribution of the discharge rates evoked by the DPT. The lighter dashed line is the cumulative distribution of spontaneous discharge rates in a healthy ear (Liberman, 1978). The distribution of rates evoked by the DPT differs from that for spontaneously responding neurons in that (1) a larger proportion (45%) of units responded with discharge rates of less than 5 spikes/s, including the 40% that completely stopped firing after 1 minute of DPT stimulation and (2) about 20% of the units responded with discharge rates that were above those for any spontaneously responding units. However, only 10% of the fibers

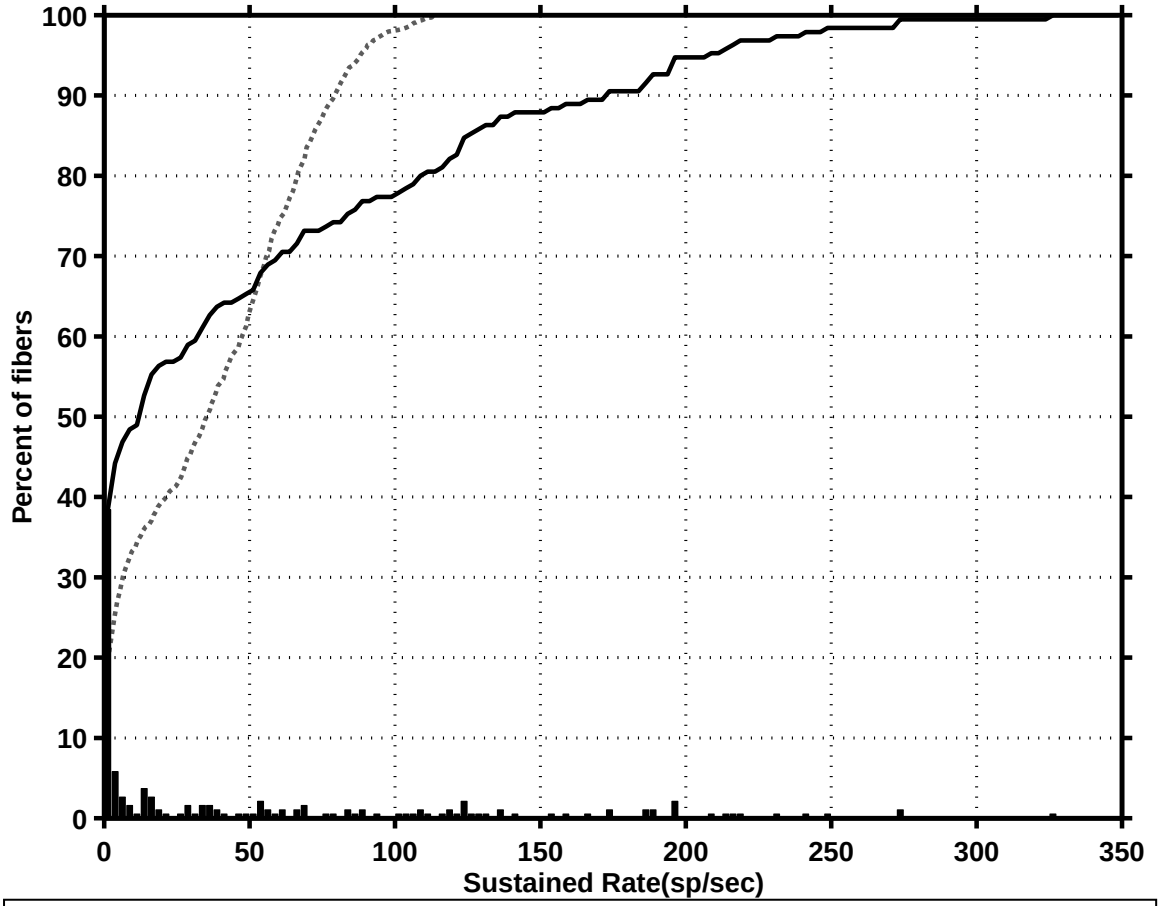


Figure 4. The black vertical bars represent a histogram of discharge rates evoked in response to the DPT (bin width 2.5 spikes/s). For each unit, the discharge rate is computed based on responses that occurred from 50 s to 10 minutes after stimulus onset. The black line shows the cumulative distribution of discharge rates. The dashed line shows the cumulative distribution of discharge rates of spontaneously responding fibers from a healthy ear (Liberman and Kiang, 1978).

exceeded the maximum spontaneous discharge rate of 120 spikes/s by more than 50 spikes/s. Overall, given the large differences in threshold and the narrow dynamic ranges reported for electrically stimulated ANFs to short stimuli (van den Honert and Stypulkowski, 1984; Javel et al., 1987; Miller et al., 1999), the sustained discharge rate distribution is surprisingly similar to that for spontaneous activity.

As can be seen in Fig. 5, nearly all units with threshold within 2 dB of the DPT level either stop responding to the DPT, or respond to the DPT with a very low discharge rate (<20 spikes/s). In contrast, the discharge rate of units with thresholds more than 2 dB below the DPT level varies widely from 0 to 300 spikes/s. The mean discharge rate correlates negatively with threshold. However, there is a great deal of variability in discharge rate even for units that have similar thresholds. For example, neurons with

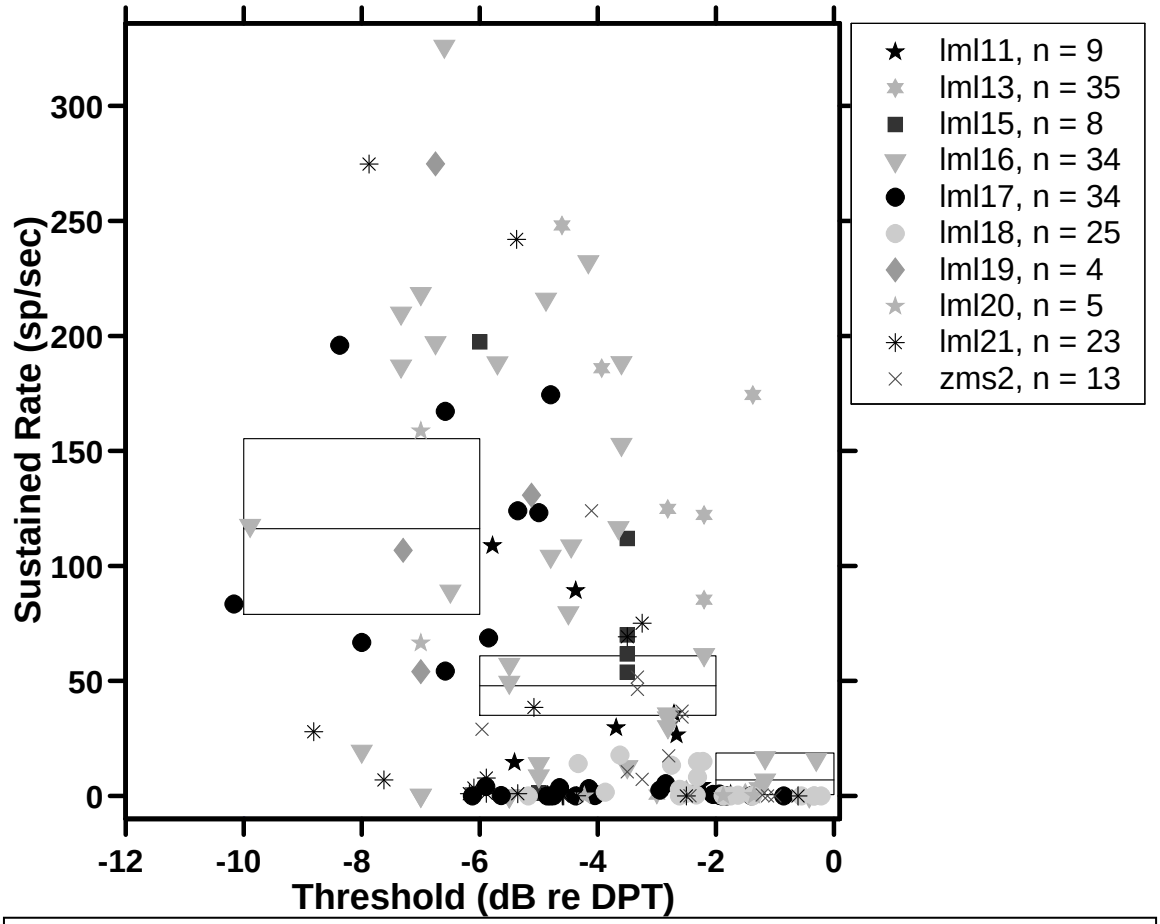


Figure 5. Average discharge rate measured in response to the DPT after 50 seconds from stimulus onset as a function of the fiber threshold. Results from 10 cats are coded by different symbols. In order to compare between different experiments and DPT levels, threshold is plotted relative to the level of the DPT. The black horizontal lines show the mean discharge rate for units that have thresholds in the range represented by the horizontal extent of each line. Boxes around each line represent the 99% confidence interval for the mean discharge rate estimate. The confidence interval was computed using the bootstrap method (Efron and Tibshirani, 1993)

threshold of -4 dB re the DPT level have sustained discharge rates ranging from 0 to 250 spikes/s. The wide range of discharge rates across units of similar threshold is a consistent finding across animals. Over the range of fibers investigated, when the DPT level exceeds threshold by more than 2 dB, threshold accounts for only 11% of the variance in the sustained discharge rates.

It is possible that the difference in sustained discharge rates from fiber to fiber may be due to the differences in the residual hearing status. Hearing status of the animal was assessed by measuring click-evoked CAP thresholds in the non-implanted ear. For each unit, Fig. 6 shows the discharge rate as a function of the click-evoked CAP threshold in the contralateral ear. To be included in the plot, the click-evoked CAP threshold had to be measured within 5 hours of the single-unit measurement. Because click-evoked CAP

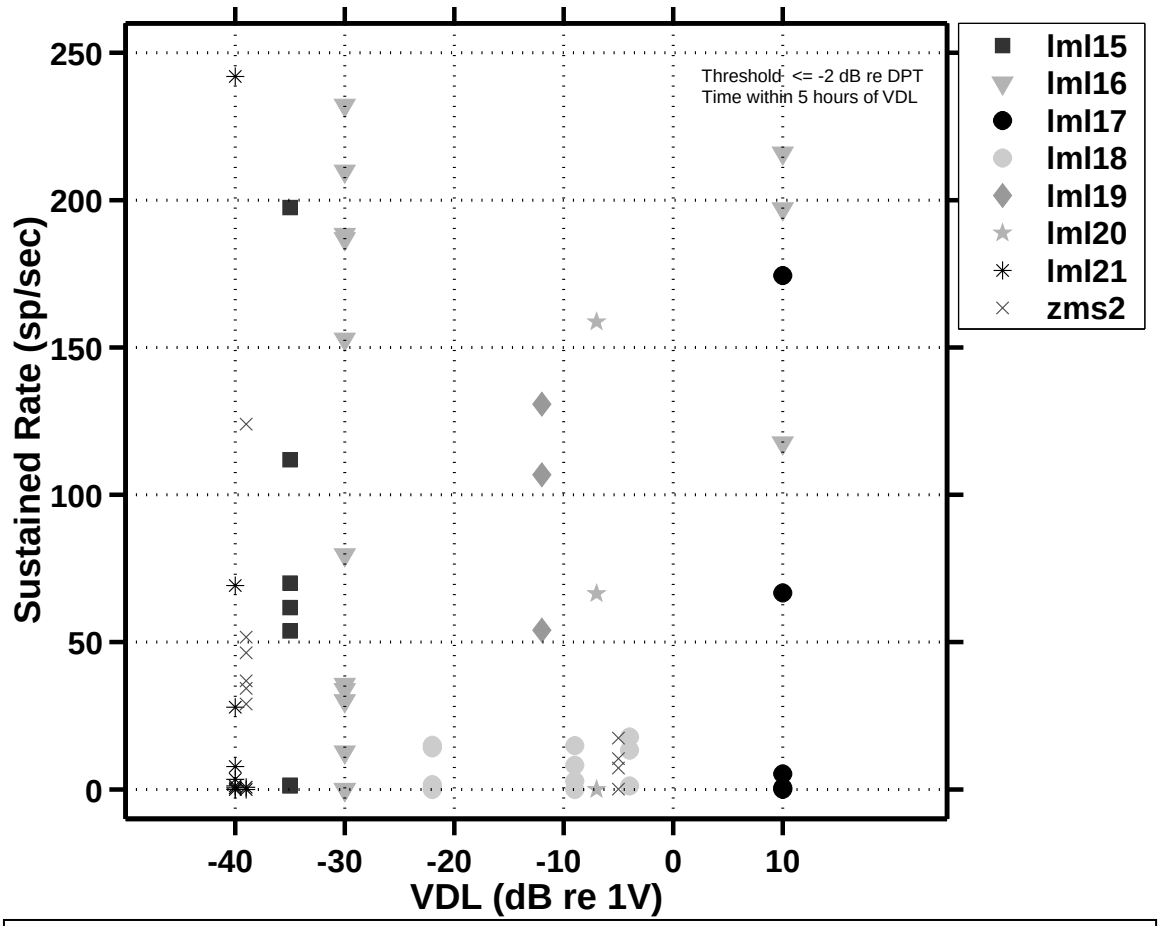


Figure 6. Discharge rate of individual fibers in response to the DPT as a function of the residual hearing status of the animal from which the data were collected. Hearing status was assessed by measuring the click-evoked CAP threshold (Methods). Only units with thresholds below –2 dB re DPT and with click evoked CAP threshold measured within 5 hours of their recordings are included. Animals whose CAP threshold was above 10 dB re 1 V had no detectable residual hearing. Because the click-evoked CAP thresholds changed slowly during the course of experiments, data from those cats may be represented at several levels in this figure.

thresholds were measured sporadically, the number of units in this figure is much smaller than in Fig. 5.

While the paucity of data prevents us from reliably computing the correlation between the average sustained discharge rate and the residual hearing status, the figure does suggest that hearing status alone cannot account for the differences in firing rate from fiber to fiber. In particular, there is a large variability in discharge rates in fibers from nearly deaf cochleas (click-evoked CAP threshold of greater than 10 dB re 1V or about 120 dB peak SPL).

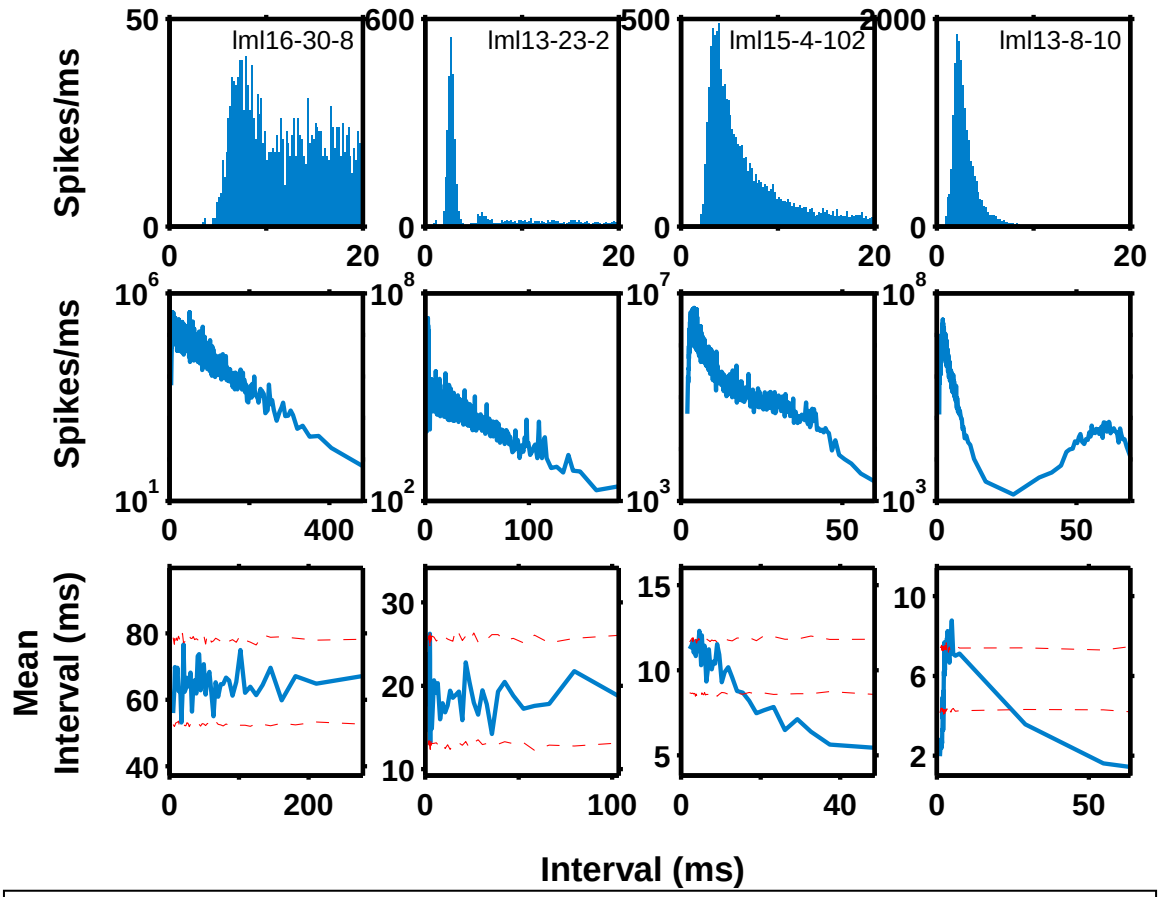


Figure 7. The top row shows the first 20 ms of the interval histogram (bin width of 200 μs) for 4 fibers in response to the DPT. The second row shows the entire interval histogram. To accommodate different average discharge rates, a variable bin width was used such that the same number of intervals fell into each bin. The lowest row plots the mean interval length conditioned by the preceding interval on the horizontal axis (solid line). The plot also shows the 99% confidence intervals for the conditional mean interval. Confidence intervals were computed based on the assumption that the intervals show no serial dependence.

C. Temporal discharge patterns

The temporal discharge patterns were studied during the "steady-state" portion of the response, and compared to the temporal patterns of activity recorded from the spontaneously responding neurons in a healthy ear. Determining a "stationarity" window is, in a formal sense, an ill-posed problem (Kelly et al., 1996). One approach is to select an *ad-hoc* criterion for choosing a window over which the response discharge rate is nearly constant. We accepted a portion of the response as nearly stationary if the slope of the regression line to the PST histogram for that portion of the response was "well within" the distribution of slopes computed from a random sequence of the same length and with a probability distribution identical to the distribution of rates in the PST-

histogram. We defined a slope as “well within” the slope distribution expected for a random sequence if at least 30% of the slopes computed from random sequences were larger than the observed slope in absolute value.

For each recorded response to a sustained DPT, the largest time window accepted as stationary was computed. Small windows were selected for records that slowly adapted throughout the measurement. Larger windows were selected for units that appeared to respond in a steady fashion after initial adaptation. The average window size was 150 s, but ranged from 5 to 500 s across units.

First-order interval histograms (IH) for spontaneous activity in a healthy ear are nearly exponential for intervals that are larger than the refractory period (Kiang et al., 1965). Responses to the DPT in the steady-state window exhibited a variety of IH shapes, as shown in Fig. 7. The top row plots the interval histogram for intervals between 0 and 20 ms. The middle row shows the same histogram, but on a logarithmic vertical scale. In addition, the time scale is adjusted in each panel to include most of the interval data for that recording.

About 20% of the records exhibited a nearly exponential IH, like spontaneously responding fibers in a healthy cochlea (left column). 40% of the records exhibited one or more modes below 10 ms, followed by a nearly exponential tail (second column from left). These patterns have been reported previously for short (250 ms) stimulation with high-frequency pulse trains (Litvak et al., 2001). A third type of response was composed of one or several modes below 10 ms followed by a non-exponential tail beyond 20 ms (two right-most columns). The last of the interval histograms corresponds to a “bursting” response. For this unit, bursting was regular, with a repetition period near 60 ms reflected in a mode of the IH.

A formal test for interval histogram type was developed based on the interval-histogram exponential shape factor (IH-ExpSF) (Litvak et al., 2001). The IH-ExpSF was computed by first fitting the interval histogram with both a single exponential and piecewise, with three exponentials. We measured the root mean squared error of each fit to the data, and defined the IH-ExpSF as the ratio of the error for the piecewise fit to that for the single exponential fit. The IH-ExpSF for samples from a Poisson process is approximately 1. A portion of a histogram was labeled as non-exponential if its IH-

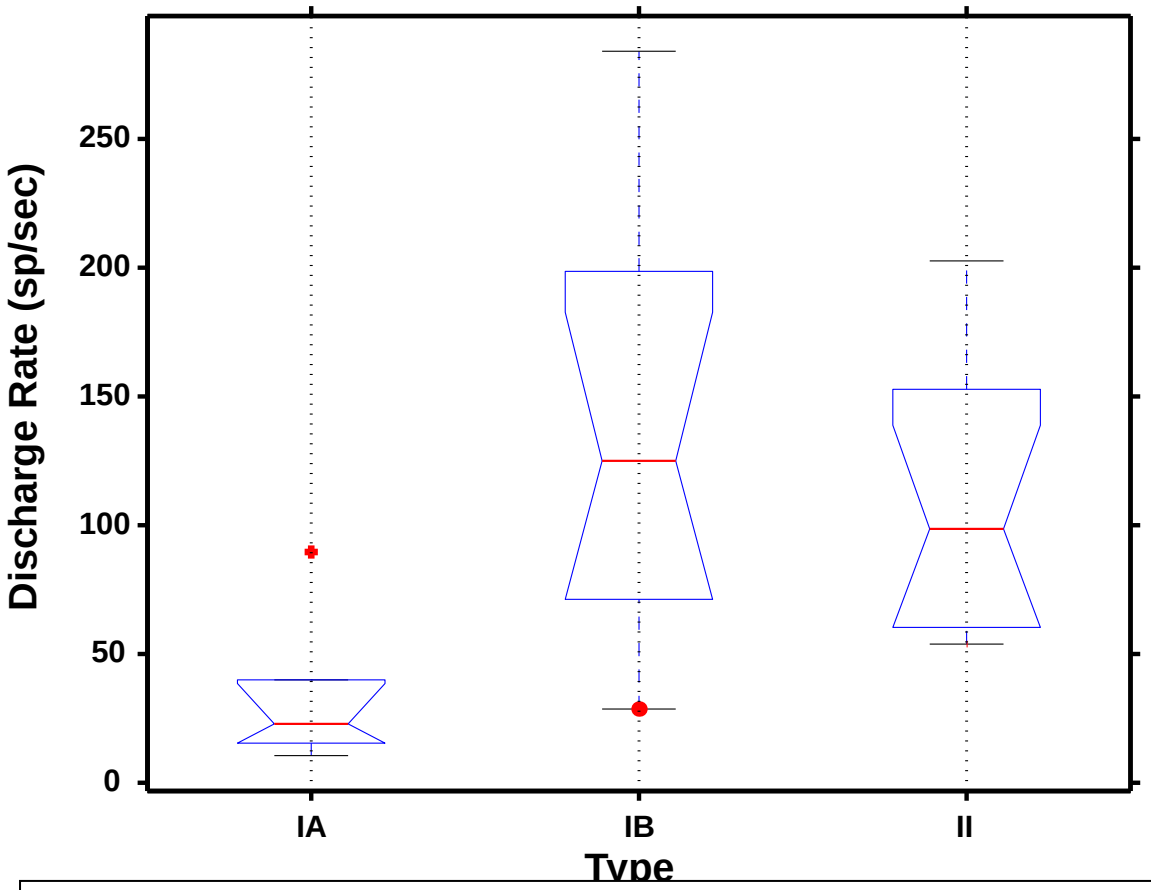


Figure 8. Distribution of average discharge rates for each interval histogram type. For each type, the statistics of the distribution are shown as a box plot (Tukey, 1977). Each box is composed of the median (middle line), upper and lower quartile (upper and lower edges of the box), upper and lower adjacent values (1.5 times the interquartile range, horizontal lines outside the box). Values outside of 1.5 the interquartile range are plotted separately (the +'s and the dots). Notches represent a robust estimate of the uncertainty about the means, and were computed using Statistical Toolbox for Matlab (www.mathworks.com).

ExpSF was larger than 99% of the IH-ExpSFs that are computed from a simulated Poisson process with the same number of spikes.

A Type IA histogram was defined as a histogram that was classified as exponential by the IH-ExpSF test over both the windows 0 to 20 ms and from 20 ms to the longest interval (e.g. Fig.7, left). The 20 ms cutoff was adjusted to be equal to the mean interval for units that had extremely slow discharge rates (<50 spikes/s). Type IB histograms are exponential only beyond 20 ms (or beyond the mean interval if the mean interval exceeded 20 ms), but not exponential in the earlier window (e.g. Fig. 7, second from left). Type II histogram are not exponential in the later window (e.g. Fig. 7, two right-most fibers).

IH Type	Conditional Mean Test	
	# Passed	# Failed
IA	5	1
IB	6	6
II	0	13

Table 1. Conditional mean test results as a function of interval histogram type.

Fig. 8 plots the statistics of the distribution of discharge rates corresponding to each of the interval histogram types using box plots (Tukey, 1977). Responses with histograms of Type IA tend to have lower discharge rates than responses of Type IB and II.

The difference in the mean of the discharge rates between IA histograms and the other responses is highly significant ($P < 0.05\%$, permutation test (Good, 1994)). Differences between the discharge rates of Type IB and Type II responses were not significant.

Joint interval histograms were computed to reveal any serial dependence in the intervals. Spontaneously responding neurons from a healthy ear show little dependence between successive intervals (Kiang et al., 1965; Lowen and Teich, 1992); consequently, they can be approximately described as a renewal process (Cox, 1967). One formal test of serial dependence is the conditional mean test (Kiang et al., 1965; Johnson, 1996). First, the mean of the intervals conditioned by the previous interval is computed. Next, it is determined whether this mean is significantly dependent on the conditioning interval.

The bottom row of Fig. 7 graphically illustrates the conditional mean test for four units. For each unit, the solid line plots the conditional mean for each conditioning interval. The dashed lines indicate the 99% confidence intervals for the mean interval computed based on the assumption that intervals have no serial dependence. These confidence intervals were computed by repeatedly shuffling the intervals and recomputing the conditioned mean interval. A record would fail the conditional mean test ($P<0.01$) if the conditioned mean exceeds the confidence intervals more than 3 times for each 100 points. With this criterion, the two left records pass the test, while the two right records fail.

60% of the DPT responses showed marked dependence on the conditioned interval, with a tendency for shorter intervals to follow longer ones (Fig. 7, bottom row, two right-most plots). The results of the conditional mean test correlated with the interval histogram type (Table 1). Of the 6 units that had a Type IA interval histogram, only one

IH Type	Hearing Status	
	Deaf (Thr>0 dB)	Hearing (Thr ≤ 0 dB)
IA	2	0
IB	4	2
II	5	4

Table 2. The number of neurons (by interval histogram (IH) type and hearing status) for which interval histogram and CAPs were measured. Only neurons for which DPT recording and click-evoked CAP thresholds were measured within 5 hours of each other were included in this table. An animal was classified as hearing if its click-evoked CAP threshold was below 0 dB re 1 V. Otherwise, the animal was classified as deaf.

failed the conditional mean test. Conversely, all of the Type II responders failed the conditional mean test. The results were split for the Type IB responders.

Another test of serial interval dependence that has been applied to spontaneous activity is the serial interevent-interval correlation coefficient (SIICC), which is defined as the correlation coefficient between successive intervals (Lowen and Teich, 1992). In a normal ear, 60% of spontaneously responding fibers had SIICC coefficients that were significantly different from 0 ($P<0.01$) (Lowen and Teich, 1992). In our data, 64% of DPT responders had significant SIICCs using the same criterion. While SIICCs reported for spontaneous activity were small values of either sign (from -0.0037 to 0.0094), SIICC values for responses to the DPT were always negative, and ranged from -0.33 to -0.017. Thus, compared to spontaneous activity, DPT responses have greater interspike interval correlations.

Table 2 shows that Type IB and II responders occurred in deaf preparations as well as in preparations with some residual hearing. Exponential (Type IA) interval histograms were seen exclusively in deaf preparations; however in only two units were measurements of DPT responses that showed IA interval histograms made sufficiently close in time to a click-evoked CAP to make their comparison useful. Thus, presence of residual hearing do not seem to account for the different complex response types.

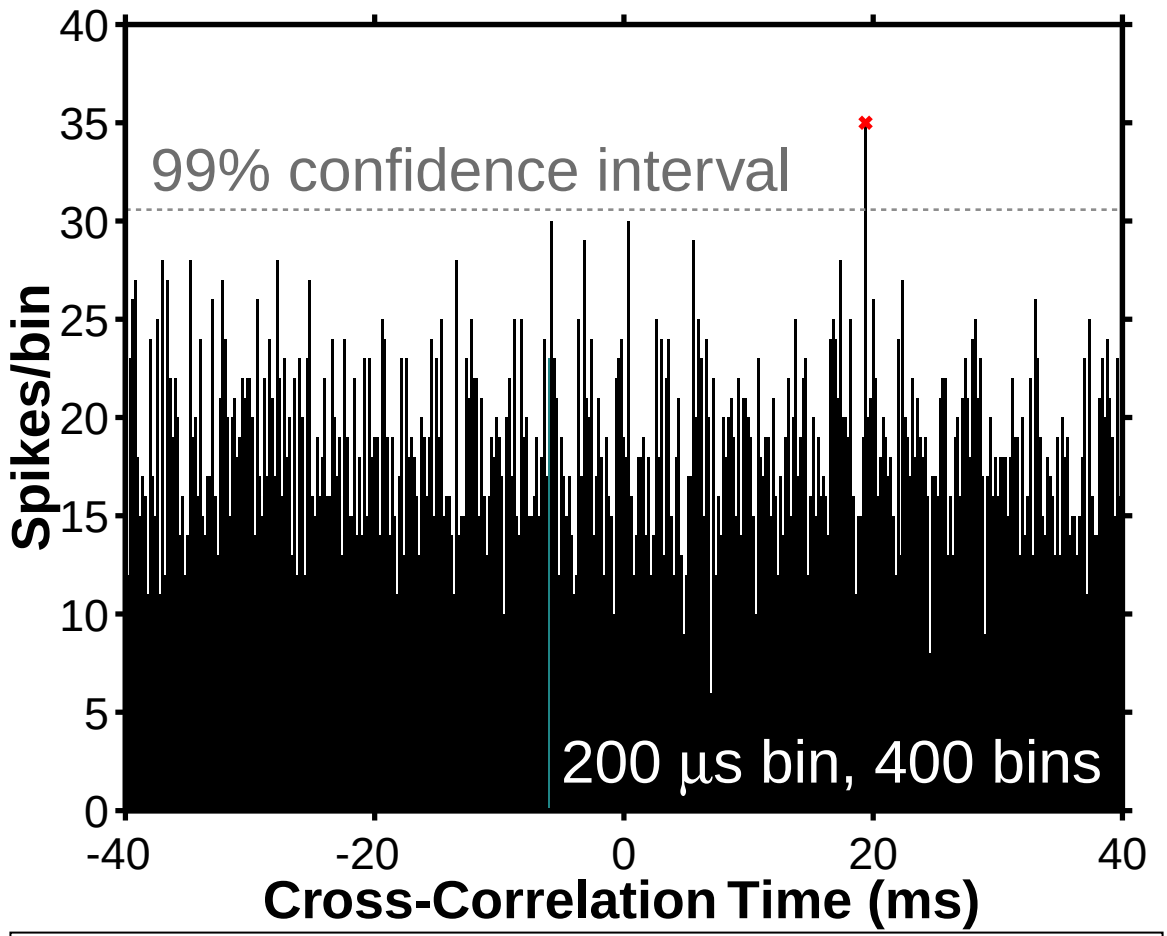


Figure 9. Cross-correlation histogram in response to the DPT from a pair of units recorded simultaneously (bin width 0.2 ms). The dash line marks the 99% confidence interval. The x indicates a crossing of the confidence interval. Because there are 400 bins, about 4 crossings are expected on average.

D. Simultaneous recordings

To assess the degree of cross-fiber correlation in responses to the DPT, we simultaneously recorded responses to the DPT from pairs of fibers, and computed the cross-correlation histogram. We examined lags up to ±40 ms at a resolution of 200 μs (thus any correlation due to the period of the stimulus, 200 μs, is averaged out). An example cross-correlation histogram is plotted in Fig. 9. The histogram appears flat, suggesting no significant correlation between the two fibers.

To determine whether the peaks that appear in the cross-correlation histograms are within limits of statistical variability, we designed a statistical test. For each pair, a 99% confidence interval was established for each lag by randomly shuffling the spike intervals in each record and computing the cross-correlation histogram 100 times. Because the

Data Record	Duration of Recording	Number of Spikes	Number of Crossings
lml13-87-2a lml13-87-2b	10 s	3069 1799	5
lml13-88-3a lml13-88-3b	10 s	2721 2953	7
lml13-89-2a lml13-89-2b	60 s	5013 1118	3
zms2-3-2a zms2-3-2b	200 s	40,158 85,724	4.2*

* average number of crossings for 10 consecutive 20-second windows of data.

Table 3. Number of times that the cross-correlation histogram crosses the 99% confidence interval. To avoid effects of slow adaptation in discharge rate that occurred for both fibers in record zms-2-3-2, the test was applied separately to 10 consecutive 20-second windows of data. This division was predicated by applying the test to simulated spike trains from two independent, but slowly adapting, Poisson processes. Both the average rate, and the speed of adaptation of the simulated spike trains were matched to those in zms2-3-2. Although these spike trains are independent by design, the cross-correlation test failed for windows that were longer than 20-seconds. Because the simulated non-adapting records passed the test, we assume that adaptation is responsible for this failure.

shuffling was performed independently for each fiber in the pair, the produced spike times are guaranteed to be uncorrelated. Under the null hypothesis, we expect the cross-correlation histogram to cross the confidence intervals on average four times. More than 8 crossings would falsify the uncorrelated response hypothesis ($p < 0.01$).

We failed to detect significant correlation in all four pairs from which we recorded (Table 3). This result, which resembles that found by Johnson and Kiang (1976) for spontaneous activity, is consistent with activity evoked by the DPT being statistically independent from fiber to fiber.

E. Vestibular versus auditory fibers

Some of the recordings described in the previous sections may have come from vestibular fibers lacking spontaneous activity. To address this issue, in later experiments, we deliberately recorded from both vestibular and auditory fibers to determine how the two differ in their responses to the DPT.

Fibers confirmed as vestibular often responded spontaneously. Discharge patterns for these units were similar to those encountered in a healthy ear, and included both regular and irregularly discharging responses. In addition, some fibers in the vestibular nerve did

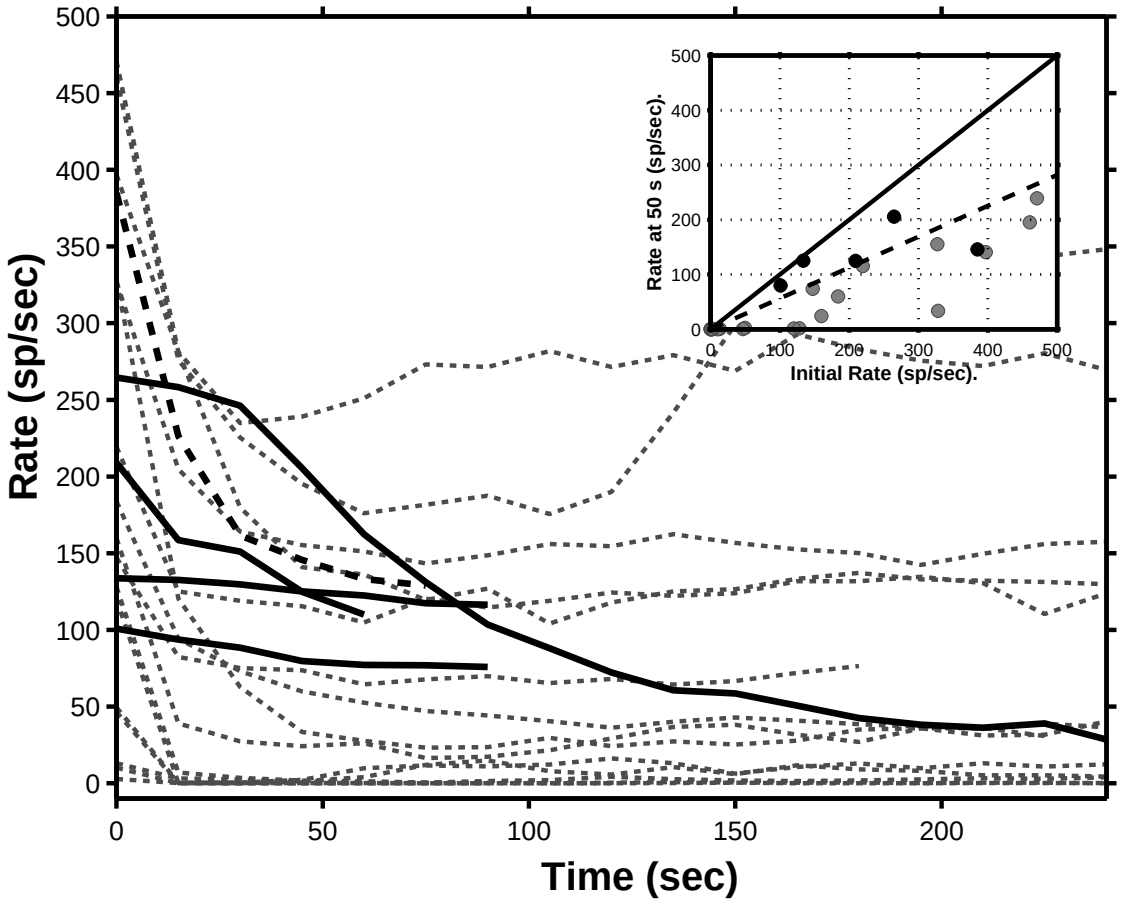


Figure 10. Average discharge rate as a function of time from onset of the DPT for auditory (gray, dotted) and vestibular (black) fibers. The inset plots the average discharge rate from 0 to 15 seconds after DPT onset versus the average discharge rate from 50 to 75 seconds for both auditory and vestibular fibers (gray and black dots respectively). Four out of five vestibular fibers have discharge rates after 50 seconds that are larger than the initial discharge rate times 0.56 (dashed line). All the auditory fibers are below the dashed line. The solid black line of the inset shows where the rate after 50 seconds equals the initial discharge rate. Units whose responses fall on this line would show no adaptation.

not respond spontaneously and were detected only because they responded to the electric pulse train.

Vestibular fibers with spontaneous activity had high thresholds to the electric pulse stimulus. For 70% of the vestibular fibers that we contacted, threshold was higher than the level of the DPT chosen for that animal. However, some vestibular fibers were easily excited by electric stimulation. One fiber's threshold was 7.5 dB below the DPT level.

Fig. 10 shows that the time course in adaptation of responses to the DPT in vestibular fibers with zero spontaneous rate is frequently different from that of auditory fibers. Of the five vestibular fibers (the total number in our database), four adapted more slowly in

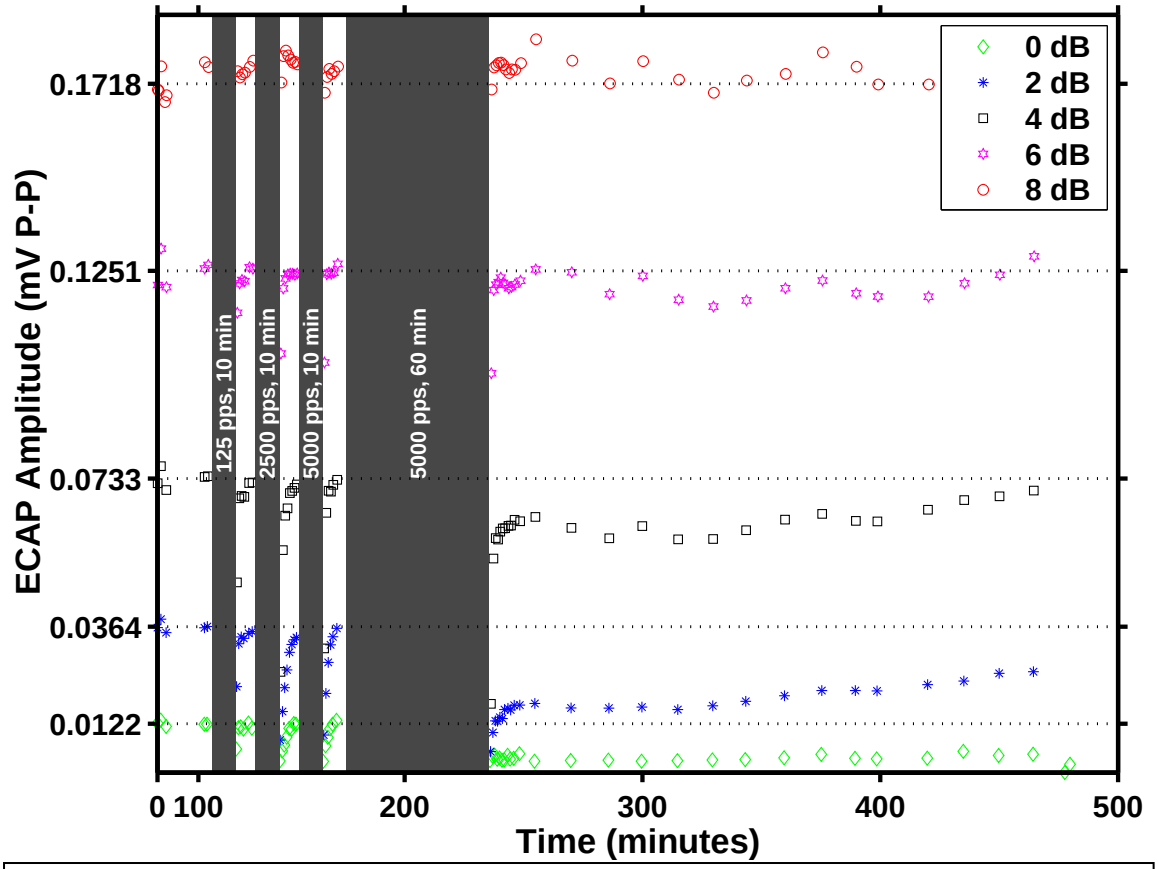


Figure 11. ECAP amplitude plotted as a function of time for one experiment (zms2). Symbols indicate different probe levels. All current levels are in dB re 1 mA 0-p. The gray areas mark times during which a conditioning stimulus was presented. No ECAP measurements were made during the conditioning stimulus. The pulse rate and duration of stimulation are indicated inside the gray areas.

the first 50 s than did any acoustic neuron (solid black lines). For one unit, the adaptation was unremarkable when compared to that of auditory-nerve fibers (dashed black line).

The inset plots the rate after 50 s of DPT stimulation versus the initial discharge rate for vestibular (black) and auditory (gray, dotted) units. For five of the six vestibular fibers, discharge rates at 50 s were greater than those in auditory-nerve fibers when the rates are normalized by the initial discharge rates. Records whose discharge rates 50 s after stimulus onset are greater than 0.56 times the discharge rate immediately after the stimulus onset (dashed line in the inset) were almost exclusively from vestibular units. This criterion was applied to units from earlier experiments to determine which of these units might be vestibular. 10% of the records were labeled as vestibular by this criterion. None of the conclusions of the earlier analyses would be changed if these records were excluded from the analysis.

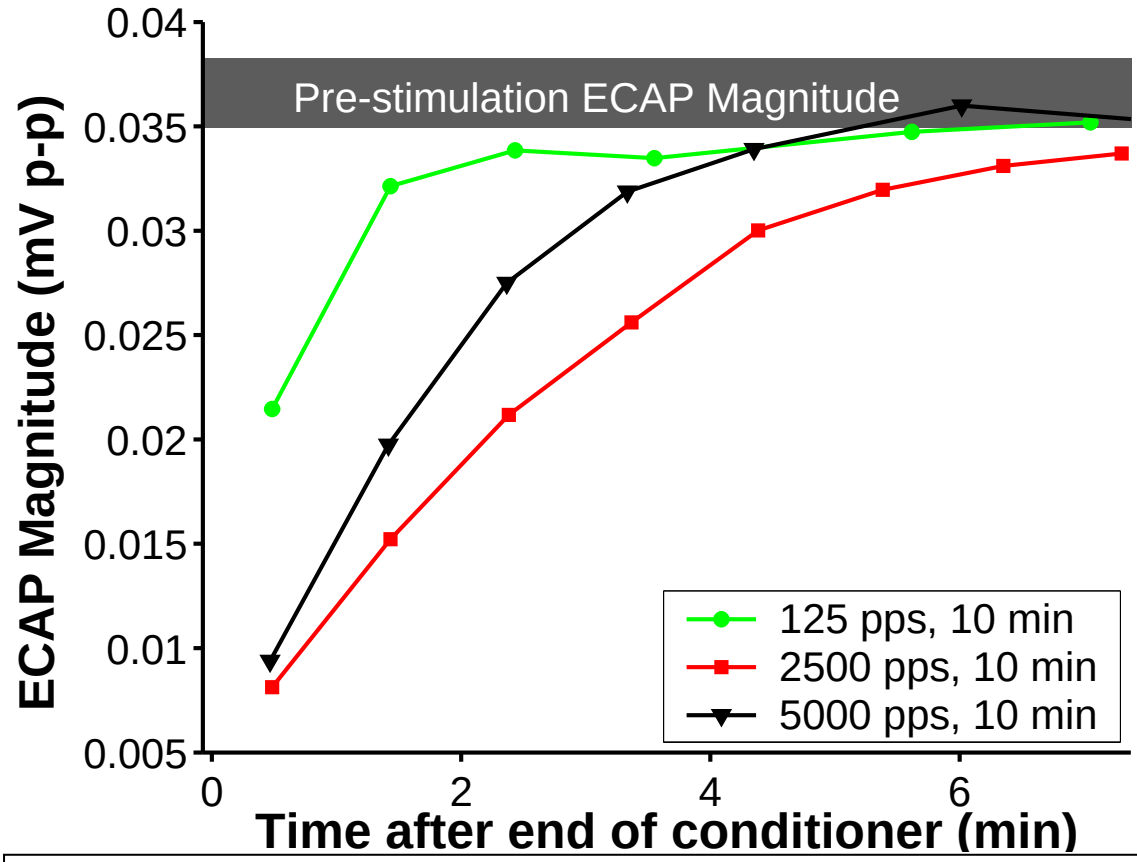


Figure 12. Recovery of response to the probe at 2 dB re 1 mA 0-p after 10 minute conditioners. The x-axis indicates the time from the end of the conditioning stimulus. Different curves indicate conditioning stimuli of different frequency. Suppression of the response immediately following 125 pps stimulation was smaller than after 2.5 and 5 kpps stimulation. In addition, recovery to 125 pps stimulation was faster, and was nearly complete in 2 minutes. Responses to 2.5 and 5 kpps recovered more slowly.

F. Recovery of evoked responses from sustained DPT stimulation

To reliably evoke a sustained discharge rate in an ANF, the DPT level must exceed its single-pulse threshold by at least 2 dB. Consequently, only high-level DPTs evoke sustained responses in a large number of fibers. The influence of such a DPT on subsequent evoked responses was investigated methodically in two cats at the end of the single-unit recordings. The results were similar in the two cats; consequently, we focus on results from experiment zms2. This cat had little residual hearing (click-evoked CAP threshold of 2 dB re 1 V) in the non-implanted ear at the time that the conditioning experiment was undertaken.

Fig. 11 plots the peak-to-peak amplitude of the ECAP as a function of time for five different probe levels (0, 2, 4, 6 and 8 dB re 1 mA 0-p). In between ECAP measurements,

we stimulated the cochlea with a conditioning stimulus. In all cases, the conditioning stimulus was a biphasic pulse train, presented at a level of 6 dB re 1 mA 0-p. The rate and the duration of the conditioning pulse train were varied.

As a control, ECAPs were first measured over a course of two hours without conditioning stimulation, and found to be stable (within ±0.25 dB). Next, a 125 pps conditioning pulse train was turned on for ten minutes. Evoked responses were measured immediately and repeatedly after the end of the conditioning stimulus. For all probe levels, evoked responses exhibited rapid recovery from this low-frequency stimulation, with the recovery being nearly complete 2 minutes after the conditioning stimulus was turned off.

Next, 10 minute conditioning stimuli of 2.5 kpps and 5 kpps were applied. These higher pulse rates are appropriate for DPT stimulation. Evoked responses recorded immediately after these stimuli were strongly suppressed, particularly for low-level probes. The response to a 0 dB re 1 mA probe was completely suppressed, while the response to an 8 dB probe was not affected by a 6 dB conditioner of either frequency. Despite this suppression, responses recovered to near pre-stimulation level after 5 to 10 minutes.

After 1 hour stimulation with a 5 kpps conditioner, responses to low-level probes were strongly suppressed. The suppression was long-lasting, with responses recovering only partially 4 hours after the end of stimulation. However, remarkably, the response to the 8 dB probe was not affected by the 1 hour DPT.

Fig. 12 shows the recovery time course after 10 minutes of 125 kpps, 2.5 kpps, and 5 kpps stimulation for the 2 dB re 1 mA probe. While the response conditioned by the 125 pps pulse train nearly recovered after 100 seconds, the responses to higher rate conditioners took at least 250 to 300 s for comparable recovery. In this animal, the recovery after 2.5 kpps was somewhat slower than recovery after a 5 kpps pulse train. However, in animal lml21, recovery was faster after the 2.5 kpps pulse train. In all cases, nearly full recovery of the evoked response was achieved after 7 minutes.

These results suggest that, high-frequency electric stimulation lasting an hour or longer may cause long term or even permanent shifts in neural excitability. On the other hand, two or three applications of high-frequency pulse trains of 10 minutes or less

separated by recovery periods of 10 minutes or more cause only temporary shifts in neural thresholds.

IV. Discussion

The goals of introducing a DPT are two-fold. The primary goal is to improve the representation of the modulations of an electric pulse train in auditory-nerve fiber responses. We postpone the discussion of modulation coding until the next paper. The second goal is to elicit a pseudo-spontaneous response in the auditory nerve. Such pseudo-spontaneous activity may be of clinical significance, because lack of spontaneous activity has been suggested as a cause of tinnitus (Kiang et al., 1970a). Restoring spontaneous activity on the auditory nerve may reduce or eliminate tinnitus in some tinnitus patients.

A. Discharge rates and adaptation

To produce a sustained discharge rate in a fiber over the course of 10 minutes, the DPT needs to be at least 2 dB above the threshold for that fiber. On average, the sustained discharge rate is larger for fibers with low thresholds re DPT level. However, for fibers with thresholds at least 2 dB below the DPT level, there is great variability in discharge rate even for units with similar thresholds.

The result that threshold poorly predicts the discharge rate after the DPT has been on for 2 minutes is consistent with earlier studies of responses to short, high-frequency electric stimuli. These studies have found that units differ strongly in the degree of adaptation to supra-threshold stimuli (Dynes and Delgutte, 1992; Killian, 1994; Litvak et al., 2001). This is true even if the stimulation level is set at a fixed value relative to fibers' thresholds (Litvak et al., 2001).

The fiber-to-fiber variability in both adaptation and sustained discharge rate may be related to a common mechanism, namely a difference in the ability of different fibers to maintain homeostasis in the face of increasing metabolic demands. One factor that may account for variability in metabolic resources across auditory-nerve fibers is fiber diameter. Larger fibers, which may have more metabolic resources per membrane area, may adapt slower to a taxing stimulus such as the DPT. Consistent with this hypothesis, when differences in threshold are controlled for, vestibular fibers, which are larger, adapt

more slowly than do auditory fibers. Peripherally, ANF diameters can differ by a factor of 2 (Liberman and Oliver, 1984). Interestingly, in a healthy ear, fibers with low spontaneous discharge rates have smaller diameters than the fibers with high spontaneous rate. Because smaller fiber diameter may be correlated both with faster adaptation and higher thresholds to electrical stimulation (Rubinstein, 1991), it is possible that many fibers with low spontaneous discharge rates respond transiently to the DPT.

B. Interval histogram patterns

Using a biophysical model of the neural membrane, Rubinstein *et al.* (1999b) predicted that responses to a 5 kpps pulse train DPT will exhibit exponential interval histograms, resembling those of spontaneously active fibers in an intact ear. Only about 20% of the interval histograms computed from DPT responses were exponential (Type IA). Most of the records with exponential interval histograms had discharge rates below 50 spikes/s, and showed no serial dependence from one interval to the next (as revealed by the conditional mean test).

Interval histograms were generally non-exponential for fibers whose discharge rates exceeded 100 spikes/s. We distinguished between two kinds of non-exponential interval histograms. One type (IB) was non-exponential in only the first 20 ms. About 50% of responses associated with this type of histogram failed the conditional mean test, indicating a serial dependence of adjacent intervals. Type II histograms were non-exponential after 20 ms. These late deviations in the interval histograms were highly correlated with serial dependence. Bursting was noted in some of these units. The inter-burst intervals were frequently regular (e.g. rightmost unit in Fig. 7, showing inter-burst interval of 60 ms).

Interval histograms of Type IA and IB were previously noted by Litvak et al. (2001) in responses of ANFs to short (<500 ms) electric pulse trains of 5 kpps. These recordings were generally not long enough to distinguish Type-II responses. Litvak et al. also noted that responses to 1 kpps pulse trains were mostly of Type-IA (exponential).

Mechanisms that generate the complex interval histograms are not well understood. Random sequences of long and short spikes have been observed in responses to DC stimulation in a squid giant axon (Guttman and Barnhill, 1970; Guttman et al., 1980).

These responses can be understood in terms of an interaction between the dynamics of the Hodgkin-Huxley model and noise in membrane channels (Schneidman et al., 1998).

Although the high-frequency pulses used in this paper differ from DC injections used in the study of the squid giant axon, some parallels may exist between the responses evoked by the two stimuli. In particular, rectification and lowpass filtering, both of which are known to occur in neural membranes, may transform the extracellular pulse train into an intracellular stimulus with a significant DC component. A model that correctly captures the dynamics of the ion channels present in auditory-nerve fibers might accurately predict the IB responses of ANFs to high-frequency electric pulse trains.

The complex Type II interval histograms may be related to processes other than those of ion channel dynamics. One example of such a process is the dynamic maintenance of sodium, potassium, and calcium concentrations in the intracellular and extracellular spaces. Slow oscillations in the concentration of these substances inside the neuron may lead to excitability changes on the scale of several tens of milliseconds, and produce burst responses. Consistent with this hypothesis, neurons that exhibit complex histograms have higher discharge rates. Higher discharge rates may indicate greater load on homeostasis, and therefore greater fluctuations in ion concentrations. Intracellular recordings near the stimulating site during DPT stimulation may be required to establish a precise basis for the complex interval histograms.

C. Safe DPT stimulation

Our results confirm that hour-long suprathreshold DPT stimulation can cause threshold shifts in low-threshold auditory-nerve fibers. This finding is consistent with earlier findings that sustained electric high-frequency stimulation may cause long-term threshold shifts in auditory nerve responses (Tykocinski et al., 1995b). The threshold shifts that we observed are distinct from those reported for DC stimulation (Tykocinski et al., 1995a), because a DC blocking capacitor was used to filter out any residual DC component of the stimuli. While the threshold shift is not necessarily deleterious in itself, it may indicate permanent damage to the cell that may eventually lead to cell death. Alternatively, over a period of days, the cell metabolism may change and even reverse part of the original shift.

In this study, a strongly supra-threshold DPT was presented on a single electrode. The level of the DPT was chosen to stimulate nearly the entire cochlea. Consequently, this DPT exceeded the threshold of the most sensitive fibers (presumably those that are near the stimulating electrode) by as much as 8 dB. Alternatively, a DPT could be presented simultaneously on several intra-cochlear electrodes. Each electrode would stimulate only a local population of fibers. This approach may be safer because the DPT level would be lower relative to threshold for most fibers.

We should note that near-threshold, sustained high-frequency stimulation is already used routinely in some cochlear implant devices (Eddington, personal communication). In these devices, the carrier pulse train is at a subthreshold level even when there is no sound. This stimulation, while safe, is probably at a level well below that required to evoke sustained responses from most auditory-nerve fibers.

Our data suggests that ten-minute stimulation followed by a rest period of ten minutes or more might be safe. Therefore, short-term laboratory tests to evaluate the effectiveness of the DPT concept may be appropriate.

Finally, a cochlear implant processor could evoke pseudo-spontaneous activity temporarily on auditory-nerve fibers by utilizing a DPT that fluctuates slowly (on the scale of several minutes) in magnitude over the course of several minutes. More work is necessary to establish the safety and effectiveness of fluctuating DPT stimulation.

D. Do DPT responses mimic spontaneous activity?

One of the goals of introducing a DPT is to evoke activity on the auditory nerve that mimics spontaneous activity in a healthy ear. This activity may improve coding of the electric modulations. Restoration of a form of spontaneous activity may also alleviate some forms of tinnitus.

We found that most fibers respond in a sustained manner to a DPT presented at a level more than 2 dB above threshold. After two minutes of DPT stimulation, the electrically-evoked activity exhibited broadly distributed interval histograms, like spontaneous activity. These broad interval histograms suggest that responses are random, and that they may be uncorrelated from one fiber to the next. We have directly tested this hypothesis by simultaneously recording from pairs of auditory-nerve fibers while

stimulating with the DPT. We failed to detect significant correlations across the responses of each of the four pairs tested.

DPT-evoked responses differed somewhat from spontaneous responses in a healthy ear. First, a greater portion of fibers had sustained discharge rates to the DPT below 2.5 spikes/s compared to the proportion of the fibers that have low spontaneous discharge rates in a healthy ear. In addition, 20% of the fibers responded with higher rates than any spontaneously responding auditory-nerve fibers in a normal ear. Interval histograms of DPT-evoked activity also differ in shape from those of spontaneous activity. While spontaneous responses always have nearly exponential histograms (Kiang et al., 1965), more than 80% of the interval histograms computed from DPT responses are non-exponential.

While the differences in interval histograms between the DPT-evoked and spontaneous activity are interesting from a modeling perspective, their functional significance, if any, is unclear. Studies in normal-hearing cats revealed a very small sub-population of spontaneously responding fibers with non-exponential interval histograms (Teich and Khanna, 1985; Gaumond, 1980) that resemble the Type IB histograms in this study. Interestingly, these fibers also tend to have high spontaneous rates (Miller and Wang, 1993).

Nearly 50% of responses to the DPT have serial inter-spike interval correlations that exceed the correlations reported for spontaneously active fibers by a factor of 10. Because, at least on a time scale of 40 ms, the DPT responses appear to be uncorrelated from fiber to fiber, and because many cells in the central nervous system integrate activity across many ANFs, these serial correlations may be of little functional significance.

In a normal ear, spontaneous activity shows long-term fractal fluctuations in discharge rate (Teich, 1989; Kelly et al., 1996). It is of interest to determine whether similar fluctuations occur in electrically-stimulated auditory-nerve fibers. Laneau and Wouters (2001) analyzed DPT responses of 3 ANFs from our data. They reported fractal fluctuations in all three of these units. The fractal onset time (defined as duration of time window for which the Fano Factor begins to increase with increasing duration) of these responses was near 100 ms, and was consistent with the fractal time reported for

spontaneously responding ANFs in a healthy ear. Thus, it is probably not necessary to introduce fractal correlations into the electric stimulus in order to mimic fractal properties of the auditory nerve responses, as suggested by Lowen and Teich (Lowen and Teich, 2000).

While our experiments suggest that a high-level DPT may introduce pseudo-spontaneous activity on the auditory nerve, psychophysical experiments will be necessary to determine whether the subjects can learn to interpret DPT responses as spontaneous activity (i.e. silence). Recent experiments indicate that this is indeed the case (Rubinstein, personal communication). Experiments are also under way to test the value of the DPT as a tinnitus treatment.

Chapter 4

Responses of auditory-nerve fibers to sustained high-frequency electric stimulation. II. Sinusoidal modulators

I. Introduction

Any sound can be decomposed into the product of a slowly-varying envelope and the more rapidly-varying fine-time structure. Continuous interleaved sampling (CIS) processors used in today's cochlear implants make use of an envelope detector in each frequency channel to derive the waveforms that are presented to the electrodes by modulating carrier pulse trains. These envelope detectors discard the information carried in the fine-time structure.

Smith *et al*. (2002), evaluated the relative importance of envelope and fine-time structure information using acoustic simulations called "auditory chimeras". The acoustic chimeras are sounds, which, in each frequency band, have the fine-time structure of one sound, and the envelope of another. Smith *et al*. (2002) found that, with 4 to 8 frequency bands, speech comprehension is better using the envelope information than the information in the fine-time structure. However, for the same number of analysis channels, subjects performed melody identification and sound localization tasks better with the fine-time structure than with the envelope information. These results suggest that modifying CIS processors to include the fine-time structure information might improve pitch perception of cochlear implant patients. Fine-time structure may also be essential for taking full advantage of binaural cues delivered by bilateral implants.

In a healthy ear, auditory-nerve fiber (ANF) responses encode the fine-time structure information in their temporal discharge patterns for frequencies up to 5 kHz (Johnson, 1980). Although it would be conceptually straightforward to include the fine-time structure in the modulation waveforms used to modulate the CIS pulse trains, it is unclear whether this information would be accurately encoded in the ANF responses. ANFs are

poor at coding high-frequencies with electric stimulation. For example, in response to a 1 kpps electric pulse train, electric compound auditory potentials (ECAPs) to each pulse alternate between strong and weak responses (Wilson et al., 1987). Alternation suggests that, while many neurons respond to the first pulse, these neurons are in a refractory state during the second pulse and only respond again to the third pulse. If most neurons fire together on every other cycle, then the auditory nerve population represents half of the stimulus frequency rather than the actual frequency.

Rubinstein *et al*. (1999b) proposed that the coding of stimulus fine-time structure might be improved by introducing a sustained, high-frequency, "desynchronizing" pulse train (DPT) in addition to the CIS modulated pulse train. In a previous report (Litvak et al., 2002c), we recorded ANF responses to a sustained, 5 kpps DPT. For many fibers, we found that, after one to two minutes of stimulation, the DPT produced activity that, in many respects, resembled spontaneous activity in a healthy ear. We identified several types of responses to the DPT. Some fibers (roughly 40%) only responded transiently to the DPT, while others showed a sustained response during 10 minutes of DPT stimulation. Among the latter fibers, some responses (20%) had interval histograms with an exponential shape, as does spontaneous activity. However, the majority of sustained responses had a non-exponential interval histogram.

In this paper, we directly test the hypothesis that a DPT improves the representation of the fine-time structure in the temporal discharge patterns of ANFs.

There are three ways in which a DPT could enhance responses to electric modulations and make them better resemble acoustic responses of a healthy ear. First, a sustained DPT may evoke activity that is desynchronized across-fibers. For higher-frequency (>500 Hz) stimuli, desynchronization may cause different neurons to respond to different simulus cycles. By simultaneously recording from pairs of fibers, we showed that such desynchronization does in fact occur after a few seconds of DPT stimulation (Litvak et al., 2002c).

Second, small electric stimuli may modulate existing ANF responses to the DPT. To the extent that the responses to the DPT are stochastic, the modulated DPT responses may be stochastic as well. In the majority of ANFs in a healthy ear (Liberman, 1978), sounds are encoded as rapid modulations of the ongoing spontaneous activity. Such

modulation of random activity allows faithful transmission of the stimulus fine-time structure in neural discharges by a process akin to stochastic resonance (Collins et al., 1995; Wiesenfeld and Moss, 1995). If the stochastic nature of the responses could be restored by the DPT, then by a similar mechanism coding of electric stimulus fine-time structure may also be improved.

Finally, computer simulations of auditory-nerve fibers suggest that the DPT may increase the dynamic range of single ANFs to electric stimulation (Rubinstein et al., 1999a). The increased dynamic range is particularly significant when one considers the representation of complex modulators such as vowels in the temporal discharge patterns. These waveforms contain peaks of widely different heights in each modulation period. Fibers with a wide dynamic range would best represent all of the waveform peaks in their temporal responses rather than just the largest peaks.

Previous tests of the ideas underlying the DPT (Rubinstein et al., 1999b; Wilson et al., 1998) have focused on measuring gross electric compound auditory potential (ECAP) responses. ECAP measures provide only indirect evidence of single-unit activity. In addition, these studies investigated short desynchronizers (30 ms), which are a poor model of a sustained DPT. The nearly steady-state responses to the DPT occur only after one to two minutes of DPT stimulation (Litvak et al., 2002c). In this work, we directly tested the hypotheses underlying the DPT by recording from single fibers from the auditory nerve of deafened cats. In our study, we focused on responses that occurred after adaptation to a DPT presented for 10 minutes. We compared temporal discharge patterns of ANF responses to electric sinusoidal modulations of the DPT with responses evoked by pure tones in a healthy ear. We also test the hypothesis that the DPT increases the dynamic range to electric stimulation by recording responses to sinusoidally modulated DPTs over a range of modulation depths.

II. Methods

The animal preparation, electrical stimulation, and recording methods have been described previously (Litvak et al., 2002c). Briefly, cats were anesthetized with dial in urethane (75 mg/kg), then deafened by co-administration of kanamycin (subcutaneous,

300 mg/kg) and ethacrinic acid (intravenous, 25 mg/kg) (Xu et al., 1993). Two intracochlear stimulating electrodes (400 mm Pt/Ir balls) were inserted into the cochlea through the round window. One electrode was inserted approximately 8 mm and was used as the stimulating electrode. The second electrode was inserted just inside the round window and served as the return electrode. The opening in the round window was sealed with connective tissue.

Standard techniques were used to expose the auditory nerve via a dorsal approach (Kiang *et al*., 1965). We recorded from single units in the auditory nerve using glass micropipettes filled with 3M KCl. For unmodulated pulse trains, a digital filter was used to separate neural responses from stimulus artifact (voltage excursions recorded at the micropipette that are not spikes) in real time. Spike and artifact waveforms were stored for later analysis. To remove the artifact from these records, we first filtered the waveform using a boxcar filter of length equal to the carrier period (0.2 ms). This filter removed most of the 5 kpps periodic component from the artifact. For modulation depths below 2.5%, the remaining artifact was smaller than the noise in the recording for nearly all of the records. For modulation depths between 2.5 and 15%, the artifact was of comparable level to the spikes. To remove the artifact from these records we utilized two algorithms, one for modulation frequencies below 500 Hz, and the other for higher modulation frequencies. Both algorithms are described in Appendix A.

A. Stimuli

We conducted a neural population study by investigating responses to a single DPT level for each animal. To maximize the yield in our experiments, the DPT level was chosen to be 8-10 dB above the compound action potential threshold. Methods for setting the DPT level are described in detail in the companion paper (Litvak et al., 2002c).

Small modulations of a sustained DPT (modulation depth ≤ 15%) were used to imitate the effect of presenting modulated pulse trains superimposed upon DPT stimulation. We assumed that neural responses to a high-frequency pulse train with a low modulation depth are similar to those elicited by a stimulus that is a sum of a sustained DPT and a highly modulated pulse train (Fig. 1). This assumption may hold if the membrane time constant is large compared to the intervals between pulses. It is exactly true if the same

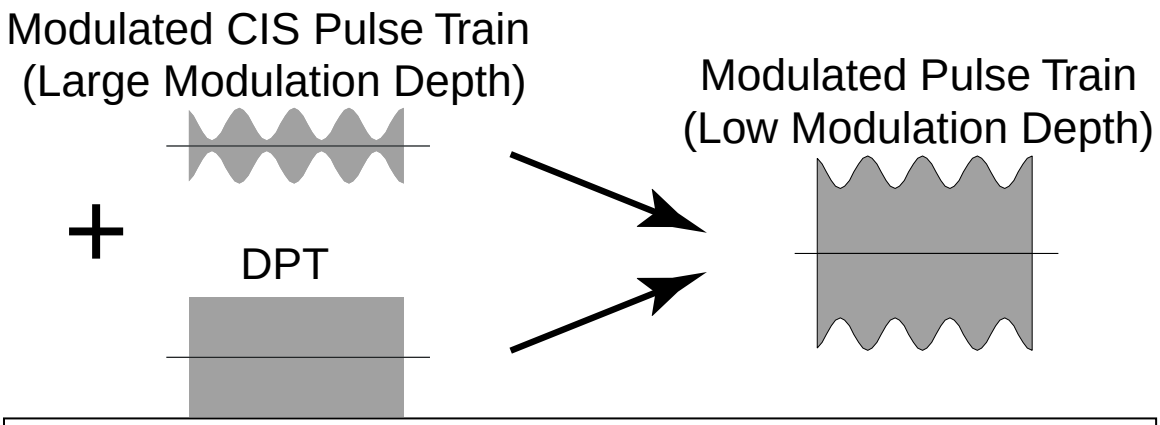


Figure 1. The left panel shows two stimuli generated by a cochlear implant to implement a DPT protocol, as suggested by Rubinstein *et al.* (1999). This stimulus is composed of two pulse trains: one, strongly modulated, is the CIS signal, and the other, unmodulated, is the DPT. The carrier for the two signals need not be identical. We assume that the DPT-enhanced stimulus can be modeled by a carrier of the same frequency as the DPT that is more weakly modulated than the original CIS stimulus (right). This approximation is exact if the CIS and the DPT have identical carriers.

pulse train is used both as the DPT and as the modulation carrier. In this case, modulation depth is linearly proportional to the amplitude of the modulated stimulus for small modulations.

Fig. 3 (top panel) schematizes the envelope of the electric stimuli used in this paper. The carrier was a 5 kpps pulse train composed of biphasic (cathodic/anodic) pulses (25 μs per phase). In order to acquire responses to both the unmodulated DPT, and modulations of the DPT for several modulation depths and modulation frequencies, the stimulus was composed of alternating modulated (400 ms) and unmodulated (600 ms) segments. Modulation depth, and modulation frequency was changed on each successive modulated segment. Modulation depth was varied from 0.5% to 15%, while modulation frequency was either 100, 420 or 830 Hz. The entire sequence of modulated and unmodulated segments was repeated continuously for 10 minutes or until contact with the fiber was lost. To expedite data collection, in the later experiments we shortened the duration of the unmodulated segments to 300 ms when the preceding modulated segment had a modulation depth below 2.5%.

Modulation was applied such that the mean amplitude of the carrier pulses was the same prior to and during modulation. Specifically, during modulation, the envelope was defined as $A \cdot (1 + m \cdot \sin(2\pi f_m t))$, where f_m is the modulation frequency, A is the amplitude of the unmodulated pulse train, and m is modulation depth. To avoid beating with the carrier, modulation frequency was always set to an integer fraction of the carrier

frequency. To facilitate the comparison of responses to modulations of different frequencies, the sinusoids were adjusted in phase relative to the carrier so that the peak of the modulator coincided with a carrier pulse. The root-mean-squared (rms) magnitudes of the sampled sinusoids differed by no more than 0.002 dB as a function of modulation frequency.

B. Analysis

Responses collected during the unmodulated DPT segments were used to classify each fiber using the same scheme as in the companion paper. Some fibers exhibited only a transient response to the DPT, and adapted to near zero discharge rate after a minute of DPT stimulation. We will refer to these fibers as "transient DPT responders." The fibers that responded throughout the unmodulated segments over the entire stimulus duration will be referred to as "sustained responders."

In the companion paper, we analyzed the temporal discharge patterns of sustained responses to the DPT using interval histograms. We found that, while some fibers had interval histograms that were nearly exponential, others had strongly non-exponential interval histograms. We quantified the degree of "exponentiality" of the histogram using the interval histogram exponential shape factor (IH-ExpSF) (Litvak et al., 2001). The IH-ExpSF was computed by first fitting the interval histogram with both a single exponential and piecewise function of three exponentials. We measured the root mean squared error of each fit to the data, and defined the IH-ExpSF as the ratio of the error for the piecewise fit to that for the single exponential fit. The IH-ExpSF for samples from a Poisson process is approximately 1.

C. Stochastic Threshold Model

As a concise way of summarizing the data, we developed a simple functional model of responses of ANFs to modulations of the DPT (Fig. 2). The model takes as input the modulation waveform $m(t)$. For a sinusoidal modulator, $m(t) = m \cdot \sin(2\pi f_m t)$, where f_m is modulation frequency, and m is modulation depth. Next, $m(t)$ is compared to a threshold. A spike is produced by the model whenever $m(t)$ crosses threshold. The threshold is the sum of a deterministic threshold $\theta_n(t)$, and a noise term. To account for the refractory

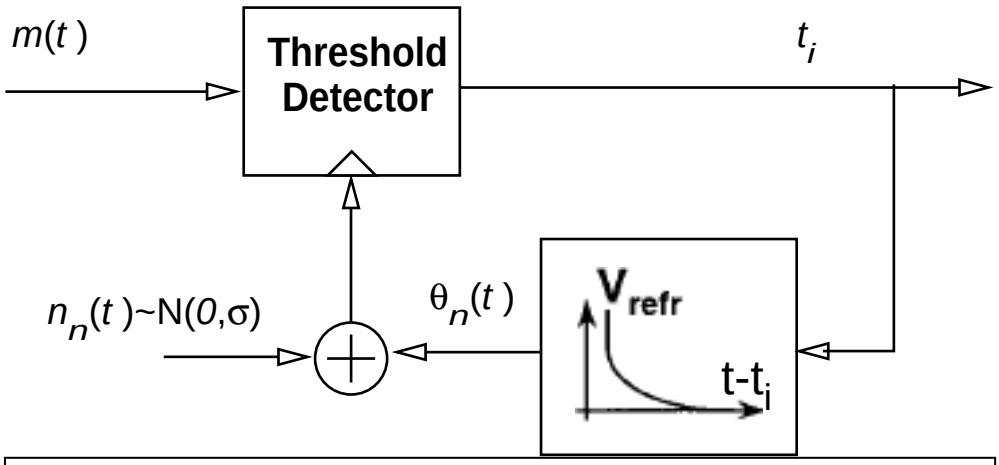


Figure 2. The stochastic threshold model (STM) of ANF responses to modulations of the DPT. Model takes as input the modulation waveform *m*(*t*). The output of the model is spike times {t_i}.

properties of ANFs, the deterministic threshold $\theta_n(t)$ depends on the time since the preceding spike t_i (Bruce et al., 1999),

$$\theta_n(t) = \theta_0 \cdot r(t) = \theta_0 \cdot \begin{cases} \infty & , t - t_i < 0.6ms \\ 1 + 0.97\exp\left[-(t - t_i)/1.32ms\right] & , t - t_i \geq 0.6ms \end{cases}. \quad \textbf{(1)}$$

The threshold recovery function $r(t)$ was chosen to fit the absolute and relative refractory periods reported for electrically stimulated ANFs (Dynes, 1995). For the noise term we used computer-generated white Gaussian noise with zero mean and standard deviation σ. Because the noisy threshold is the critical element of the model, we will refer to this model as the stochastic threshold model (STM).

We computed responses of the STM to sinusoidal modulators of different frequencies and modulation depths. The conditional probability that the modeled neuron will fire at time t given that the last spike occurred at t_i is,

$$p_f(t) = \Gamma\left(\frac{m}{\sigma} \cdot \sin(2\pi f_m t) - \frac{\theta_0}{\sigma} r(t - t_i)\right), \quad \textbf{(2)}$$

where Γ is the cumulative Gaussian function, and $r(t\text{-}t_i)$ is the recovery function described by Eq. (1). When the DPT is unmodulated (m=0), spontaneous discharge rate, which represents the discharge rate during the unmodulated DPT, is entirely determined by the ratio of θ_0/σ. For other values of m, the discharge rate depends on the ratio m/σ as well. The model was simulated using 0.2 ms time step, giving an effective noise

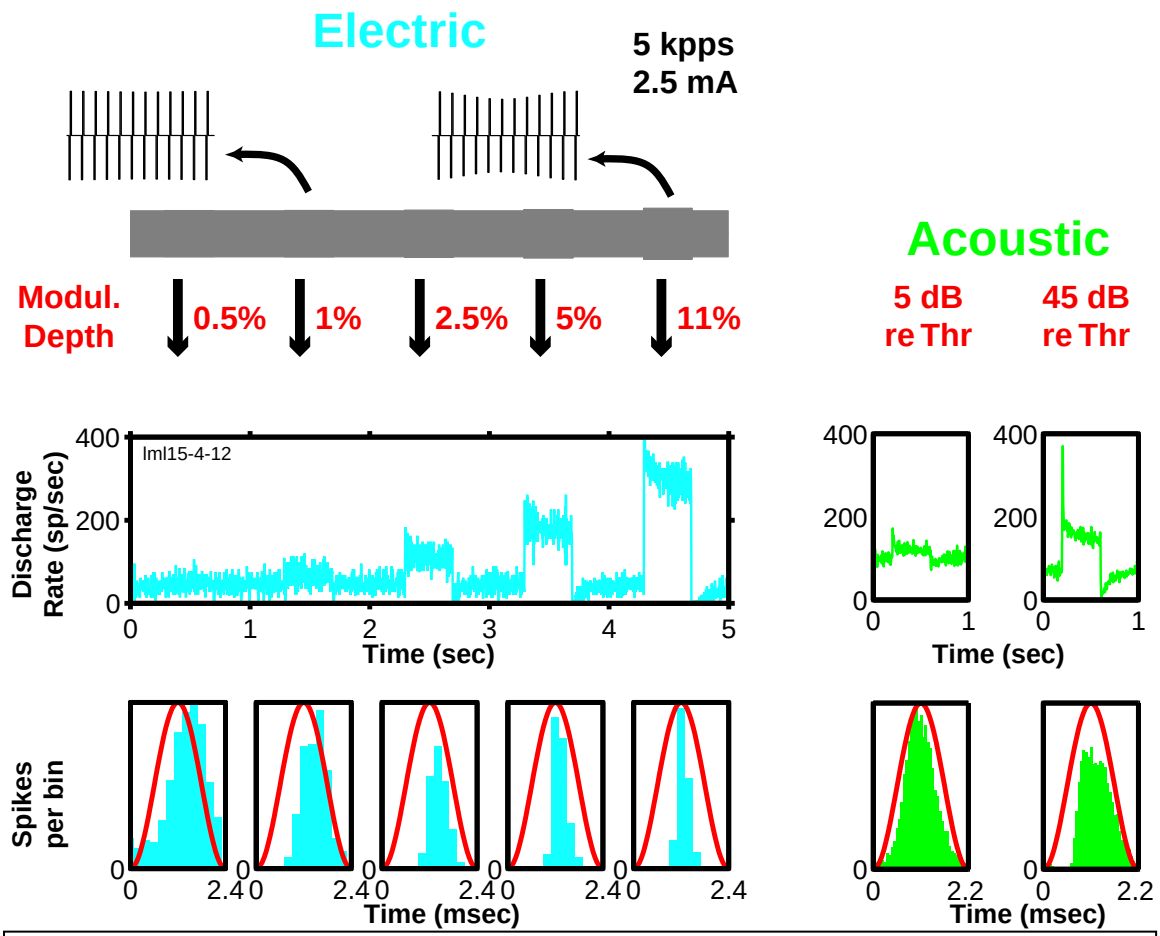


Figure 3. The top panel shows one cycle of the electric stimulus (2.5 mA 0-p). 420 Hz sinusoidal modulator was applied to the stimulus every second. Modulation depth was changed from 0.5% to 11%. The entire sequence of modulations was repeated every 5 seconds. The left panel of the middle row plots the discharge rate of fiber lml15-4, computed using 2.4 ms bins. For comparison, the right two panels of the middle row show the discharge rate computed from the response of an ANF from a healthy ear to tones at near-threshold (left) and supra-threshold (right) levels. The bottom row of panels show the period histograms computed from responses during the electric modulation. The right column shows the period histograms from the ANF response to 440 Hz tones in a healthy ear.

bandwidth of 2500 Hz. We compared STM outputs to ANF responses to sinusoidally modulated DPT.

III. Results

A. Basic response characteristics

We recorded responses from 65 units in 4 cats to sinusoidally modulated electric pulse trains. Each recording included responses to a series of modulated segments with different modulation frequencies and modulation depths. Because ANF responses to the

DPT adapt most rapidly during the first minute of DPT stimulation, we report only the responses that occurred after 50 seconds of DPT stimulation.

Fig. 3 shows the response of an auditory-nerve fiber to an electric pulse train that was sinusoidally modulated at 440 Hz every second for 400 ms. Modulation depth was increased from 0.5% to 11% on each successive modulation segment. The pulse train was left unmodulated between these modulated segments. The entire cycle of 5 modulation depths was repeated every 5 seconds for 10 minutes. During the unmodulated segments, the fiber's average discharge rate was 48 spikes/s, and its interspike interval distribution was nearly exponential (IH-ExpSF=0.99).

The left panel in the middle row shows the discharge rate as a function of time from the onset of the stimulus cycle. For comparison, the right two panels in the middle row show the response pattern of a high spontaneous-rate fiber from a healthy ear to near-threshold (left), and strongly supra-threshold (right) tone burst (440 Hz, repeated every second) (McKinney, 1998).

In response to electric stimulation, discharge rate of this fiber grew monotonically with modulation depth. For modulation depths between 0.5% and 5%, the discharge rate was in the range of the discharge rates (<300 spikes/s) reported for tone responses in a healthy ear (Kiang, 1965; Liberman, 1978). At the largest modulation depths, discharge rate exceeded the rates seen in Liberman's data.

In response to a tone burst, the discharge rate of the ANF rapidly decreases in the first 10 ms of stimulation. This rapid adaptation is followed by slower adaptation with time constant near 100 ms. While the responses to modulations of the DPT show a form of slow adaptation, there is no evidence of rapid adaptation.

The bottom row of Fig. 3 shows period histograms locked to the modulator frequency computed from responses during the modulated segments. The period histogram for the electric pulse train with a modulation depth of 0.5% is already nearly fully modulated, indicating exquisite sensitivity to modulation (left-most panel). The spikes become more precisely locked for the larger modulation depths. Synchrony to the stimulus is already large at modulation depths that evoke no noticeable increase in average rate. Similar behavior is found in responses of high-spontaneous ANFs from an intact cochlea to tones (Johnson, 1980).

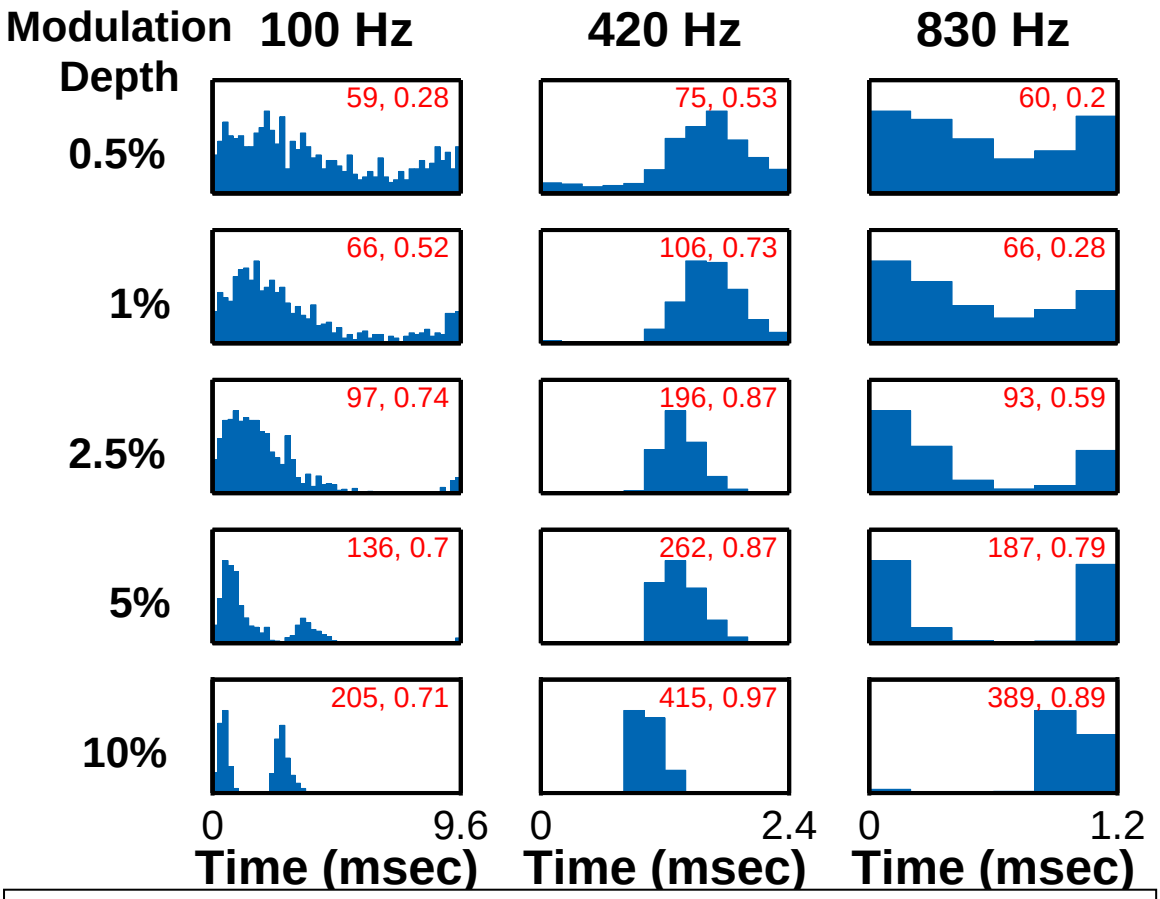


Figure 4. Each panel in this figure shows period histograms computed from responses at one modulation frequency and modulation depth. All responses in this figure are from fiber lml17-19. During the unmodulated segments, this fiber responded with a discharge rate of 54 spikes/s. The interval histogram computed from responses during the unmodulated stimulus segments had a nearly exponential envelope. In each panel, the first number is the discharge rate during the modulated segment. The second number shows the synchronization index (Johnson, 1980).

For a modulation depth of 1%, the period histogram is nearly sinusoidal in shape, suggesting that the responses may be accurately representing the sinusoidal modulator waveform (solid curve). For this modulation depth, the period histogram also resembles that for a response to a soft tone in a healthy ear (right). For modulation depths above 2.5%, however, the period histogram consisted of a very sharp mode restricted to a small fraction of the stimulus cycle, and did not resemble the response to a pure tone at any level.

B. Temporal discharge patterns

Fig. 4 shows period histograms computed from responses to 100, 420 and 830 Hz modulations of the DPT for another ANF. Modulation depth was varied from 0.5% to

10%. During the unmodulated segments, this unit responded with a discharge rate of 54 spikes/s and had a nearly exponential interspike interval distribution (IH-ExpSF=0.97). The first number in each panel indicates the average discharge rate in spikes/s during the corresponding 400 ms modulated segment. The second number is the synchronization index (Johnson, 1980).

For all modulation frequencies, stimuli with modulation depths below 2.5% evoked average discharge rates below 300 spikes/s, in the range reported for ANFs in a normal ear in response to pure tones (Liberman, 1978). In addition, the synchronization index was always below 0.9, similar to acoustic responses to pure tones. The sinusoidal shape of the histograms for the smallest modulation depths suggest that the modulation waveform is accurately represented throughout the range of modulation frequencies between 100 and 830 Hz.

For modulation depths of 0.5% and 1%, phase locking to the modulator and average discharge rate were greater for the 420 Hz modulation frequency that for the other two frequencies. At the modulation depth of 10%, the responses to 420 Hz and 830 Hz modulators had discharge rate that exceeded those seen in a normal ear. For 100 Hz modulation, the period histogram reveals multiple peaks in the response for each cycle in the stimulus. Unlike the “peak splitting” often observed in responses of ANFs to low-frequency tones (Kiang and Moxon, 1972; Johnson, 1980) which occurs on opposite phases of the tone, both peaks of the responses to electric modulation occur during the same half cycle of the modulator.

Interval histograms were examined to determine whether the electrically-evoked interval distributions resemble distributions evoked by tones in a healthy ear. Fig. 5 shows the interval histograms that correspond to the period histograms shown in Fig. 4. For comparison, the right-most column shows an interval histogram from responses of a fiber in a healthy cochlea to a 440 Hz tone. Phase locking to the modulator can be seen in these histograms as the clustering of intervals around integer multiples of the modulator period (dashed lines).

For modulation depths below 5%, interval histograms show some similarity to the acoustic responses. In particular, they show modes at multiples of the modulation period. One exception is the response to the 420 Hz modulator at a modulation depth of 2.5%,

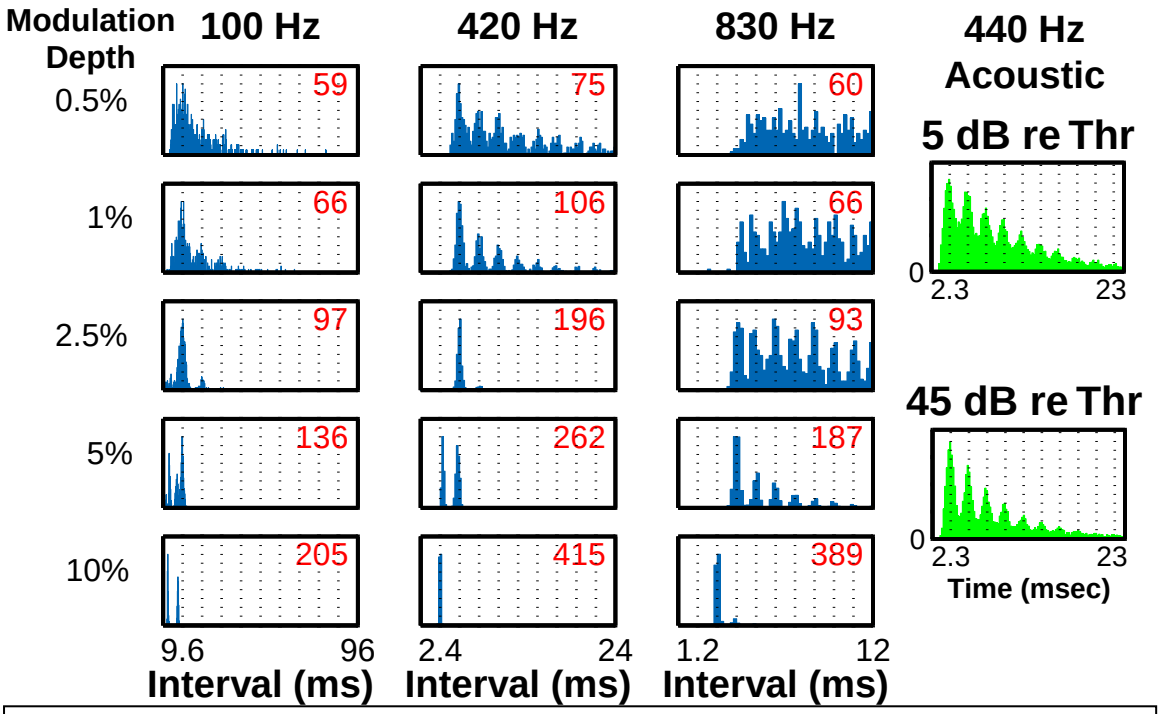


Figure 5. The first three columns from the left show the interval histograms corresponding to the period histograms of Fig. 4. The right column presents interval histograms for tone responses of a high-spontaneous rate ANF from a healthy ear. Dashed lines mark multiples of the modulation period.

which appears to occur primarily on every other stimulus cycle. However, this response pattern was unusual for this modulation frequency and depth.

Detailed examination of the interval histograms for modulation depths below 2.5% reveals additional differences between the DPT responses and the responses to pure tones in a healthy ear. Normal responses to the 440 Hz tone show a pronounced mode at the tone's period (2.2 ms). However, responses to the 420 Hz modulation of the electric pulse train show no mode at the modulation period for modulation depths below 5%, with the first mode occurring at twice the period. These differences are similar to those previously reported for responses to relatively short (200 ms) modulated pulse trains presented without a DPT (Litvak et al., 2001).

For 10% modulation, the interval histograms of responses to electric stimulation differ strongly from the normal-ear responses to tones. In response to the 100 Hz modulator, the fiber fired twice on each modulation cycle, so that the interval histogram was composed of two modes, one at 2.3 ms, and the other one at 7.3 ms. The sum of these two modes equals the modulation period of 9.6 ms. For 420 Hz modulation, the response occurred exactly once per modulation cycle. The interval histogram for this response had a single

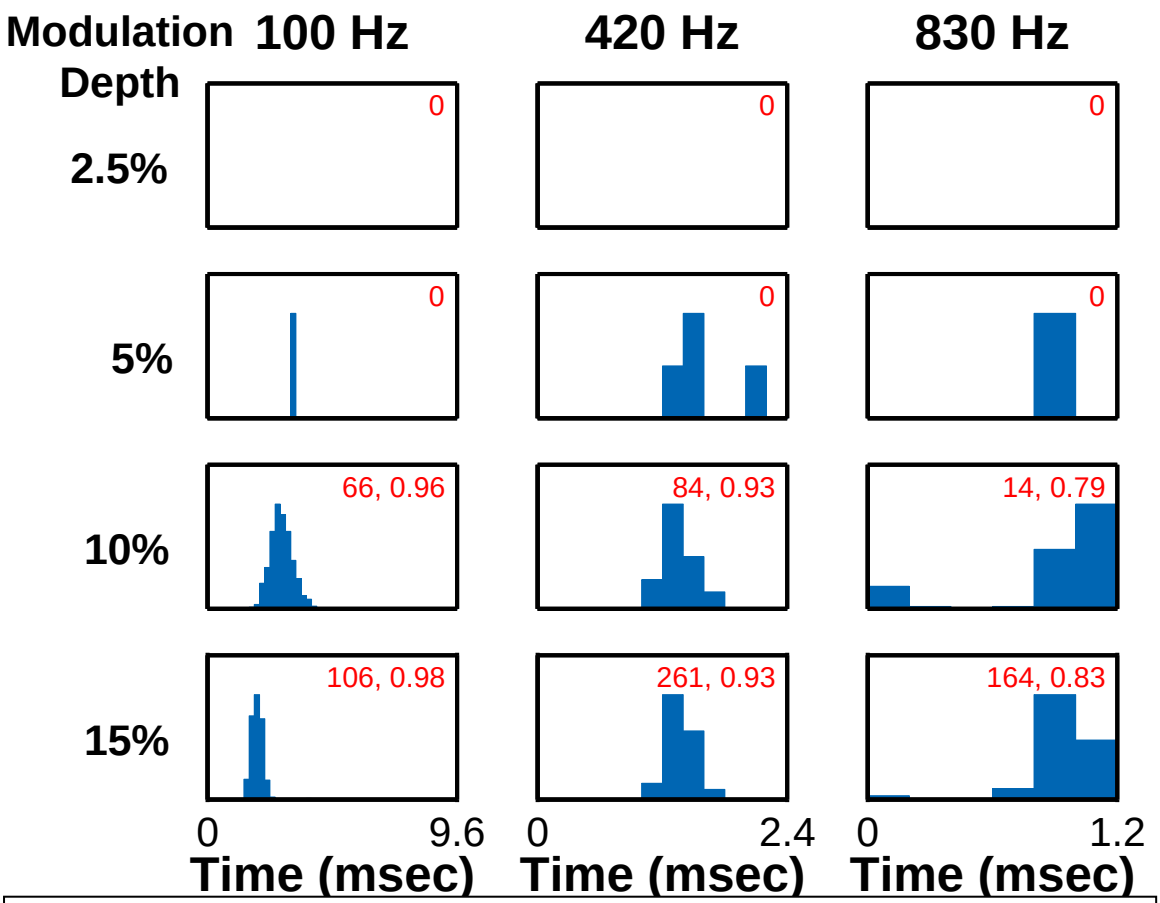


Figure 6. This figure shows period histograms computed from responses of ANF that did not respond during the unmodulated segments of the DPT. The numbers in the top of each panel indicate the discharge rates. The numbers on the bottom show synchronization indexes. This fiber is less sensitive to DPT modulation than fibers in Fig. 3 and Fig. 4, and responded only for modulation depth of 10%. The period histogram poorly resembled the modulation waveform (solid curve).

mode at the modulation period. Such entrainment is not seen in normal-ear ANF responses to tones. For the 830 Hz modulator, the response occurred on nearly every other modulation cycle (dominant mode at twice the modulation period). This distortion is particularly disturbing, because, if it occurs in many fibers, and if those fibers fire in synchrony, then the response of the fiber population would represent a signal at half the actual modulation frequency.

Fig. 6 shows an example of responses to modulations for a transient DPT responder. While this unit responded to modulations of the DPT, the modulation depth necessary to evoke a significant response exceeded 5%, larger than that for any of the sustained responder units. In addition, for the 100 and 420 Hz modulators, the synchronization index was invariably higher than that for pure tone responses from a healthy ear. For all

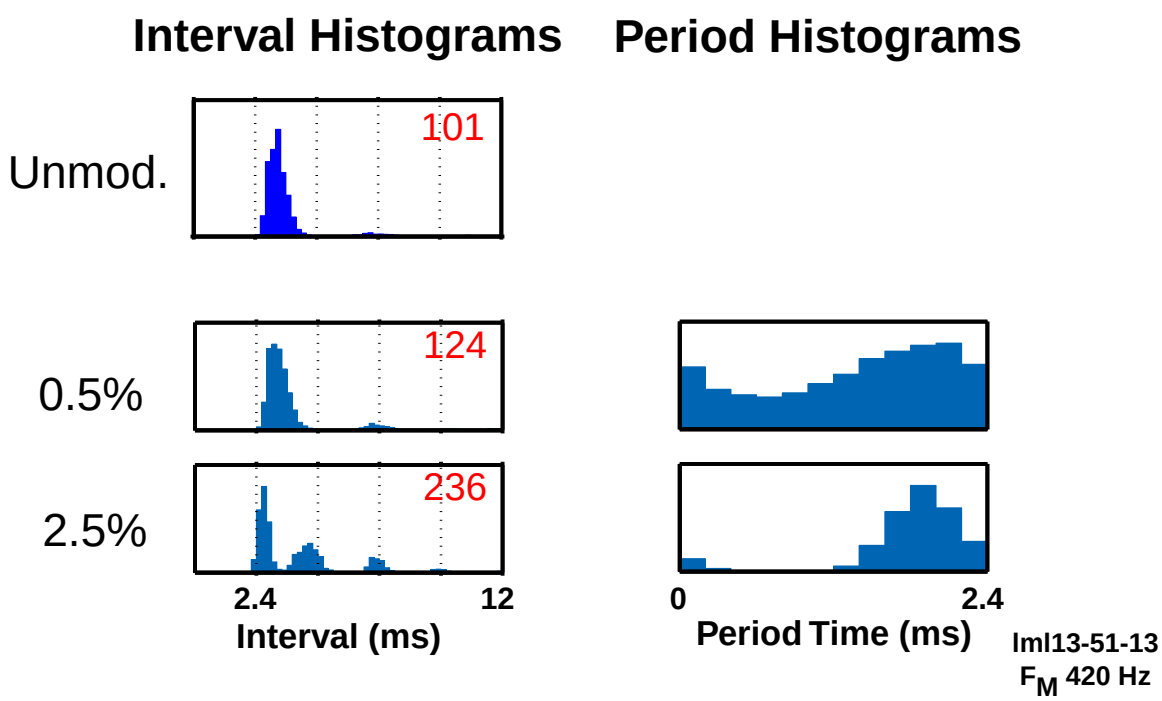


Figure 7. Interval and period histograms of unit lml13-51 responses to modulated (420 Hz) segments of the stimulus pulse train. The interval histogram computed from responses during the unmodulated segments is plotted on the top. This histogram shows a strong mode around 3.5 ms that is followed by a longer tail. The middle row shows the interval histogram (left) and the period histogram (right) of this unit's response to the 420 Hz modulation at modulation depth of 0.5%. Compared to units in Fig. 3 and 4, this ANF showed much less synchronization to the modulator. The bottom row shows the response to the modulation depth of 2.5%. Although the response much more synchronized than for the modulation depth of 0.5%, the modes in the interval histogram are still not at the modulation periods (dashed lines).

modulation depths, the period histogram of responses of these fibers poorly represented the sinusoidal modulation waveform.

C. Discharge patterns for non-exponential fibers

Fibers with non-exponential interval histograms during the unmodulated DPT often exhibited a pronounced mode in their interval histogram (Litvak et al., 2001; Litvak et al., 2002c). We refer to the location of this mode as the preferred interval. During sinusoidal modulations, temporal response patterns of non-exponential fibers depend not only on the modulated stimulus, but also on the intrinsic dynamics that are manifested by the preferred interval. Fig. 7 shows the responses of one fiber that had a non-exponential interval histogram during the unmodulated segments. The interval histogram computed from responses to the unmodulated DPT was non-exponential (IH-ExpSF=0.51), and had a pronounced mode at 3.3 ms (Fig. 7, top panel). The panels on the left of Fig. 7 show the

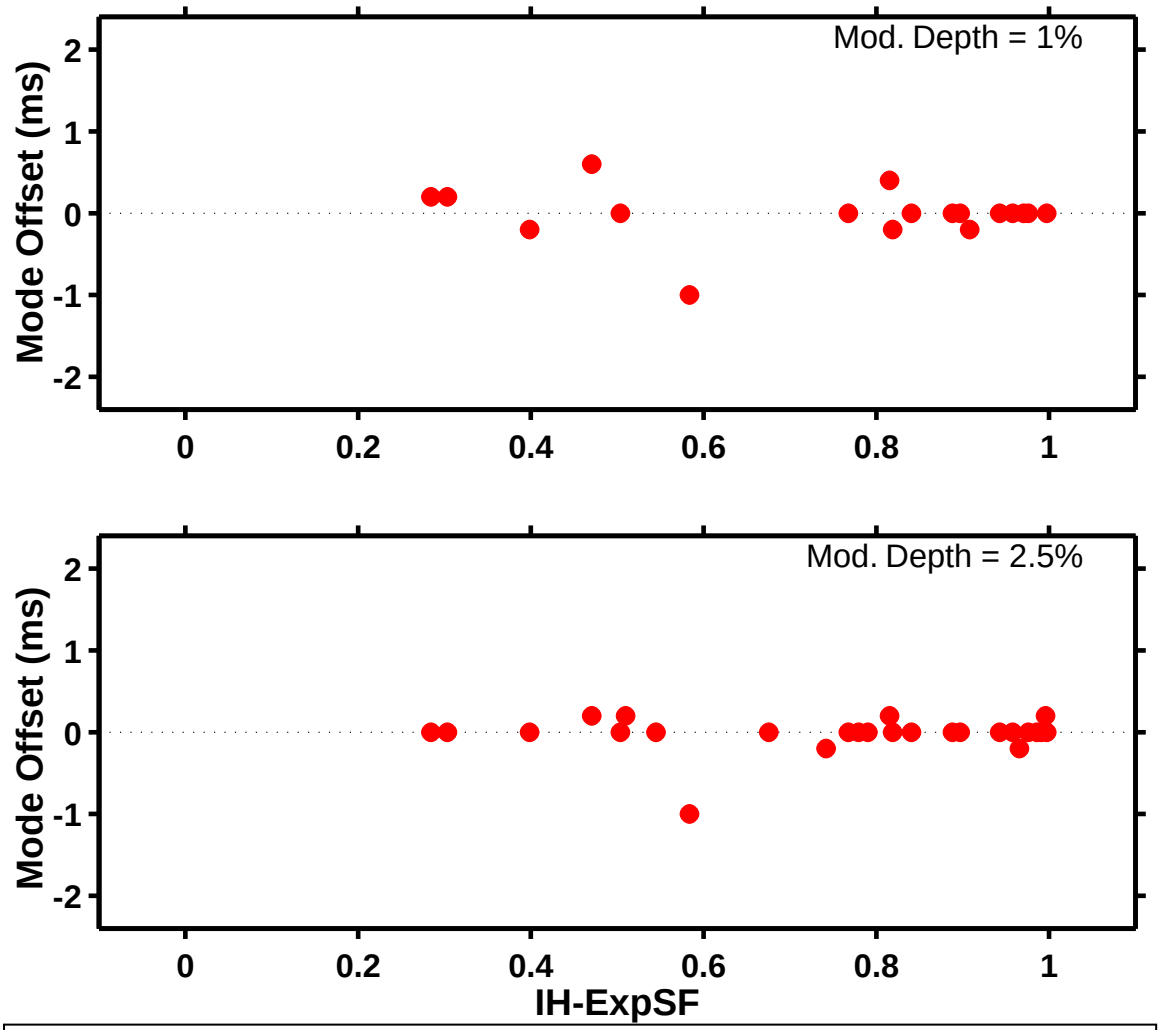


Figure 8. This figure plots the mode offset during the 420 Hz sinusoidal modulation (modulation depth of 1%, top panel, and 2.5%, bottom panel) versus IH-ExpSF computed based on responses to the unmodulated DPT. For the smaller modulation depths, offsets are zero for all units with IH-ExpSF above 0.9, and as much as 1 ms for units with IH-ExpSF below 0.9. Except for one fiber, mode offsets are smaller than 0.2 for 2.5% modulation depth.

interval histograms of the responses to a 420 Hz modulator for modulation depths of 0.5% (top) and 2.5% (bottom). The modes of the interval histograms are systematically offset from multiples of the modulation period. We will refer to these shifts as mode offsets. The direction of the mode offset is towards the preferred interval for the unmodulated DPT (top). Mode offset is most pronounced at 0.5% modulation depth, but is still detectable at 2.5% modulation depth. In addition, the envelope of the interval histogram does not resemble the exponential envelope seen for the normal-ear responses to tones (Fig. 5).

The right panels of Fig. 7 show the period histograms corresponding to the interval histograms on the left. At the lower modulation depth, this unit locked weakly to the modulator (synchronization index of 0.24). In contrast, at the same modulation depth, the response of the exponential unit shown in Fig. 4 was already nearly fully modulated. The non-exponential unit also locked less to the modulation depth of 2.5%. Thus, this unit is less sensitive to modulation than the unit in Fig. 4, which had an exponential inter-spike interval distribution during the unmodulated portions of the stimulus.

To further investigate the hypothesis that mode offsets are specifically related to non-exponential DPT responses, Fig. 8 plots the mode offsets versus IH-ExpSF for a modulation frequency of 420 Hz, and modulation depth of 1% (top) and 2.5% (bottom). Mode offsets were estimated by first computing the interval histogram of the responses, finding the bin containing the most intervals, and then determining the distance between the center of this bin and the nearest multiple of the modulation period. Mode offsets are positive if the mode is longer than the period, and negative otherwise.

For the 1% modulation depth, the near-exponential responders (IH-ExpSF > 0.9) had very small mode offsets. In contrast, some non-exponential responders had mode offsets as large as ±1 ms. At the larger modulation depth of 2.5%, mode offsets were small for both the exponential and the non-exponential responders, with only one exception. Thus, large mode offsets occur primarily at the smaller modulation depths and for non-exponential units.

Fig. 9 plots the mode offsets for 1% modulation at 420 Hz versus the mode offsets for responses to the unmodulated segments of the stimulus. The latter mode offsets were computed by assuming that modulation frequency during the unmodulated DPT was equal to 420 Hz. Only non-exponential DPT responders (IH-ExpSF < 0.9) were included in this plot, because only these fibers have a pronounced mode in their interval histogram. As for the unit in Fig. 7, the mode offsets for 1% modulation depth are strongly related to the location of the mode in the interval histogram during the unmodulated segments. The line in Fig. 9 (slope 0.65) is based on linear regression and accounts for 90% of the variance (if the circled outlier is removed). The slope, which is positive but less than one, suggests that for these fibers, the mode in the interval histogram during weak modulation

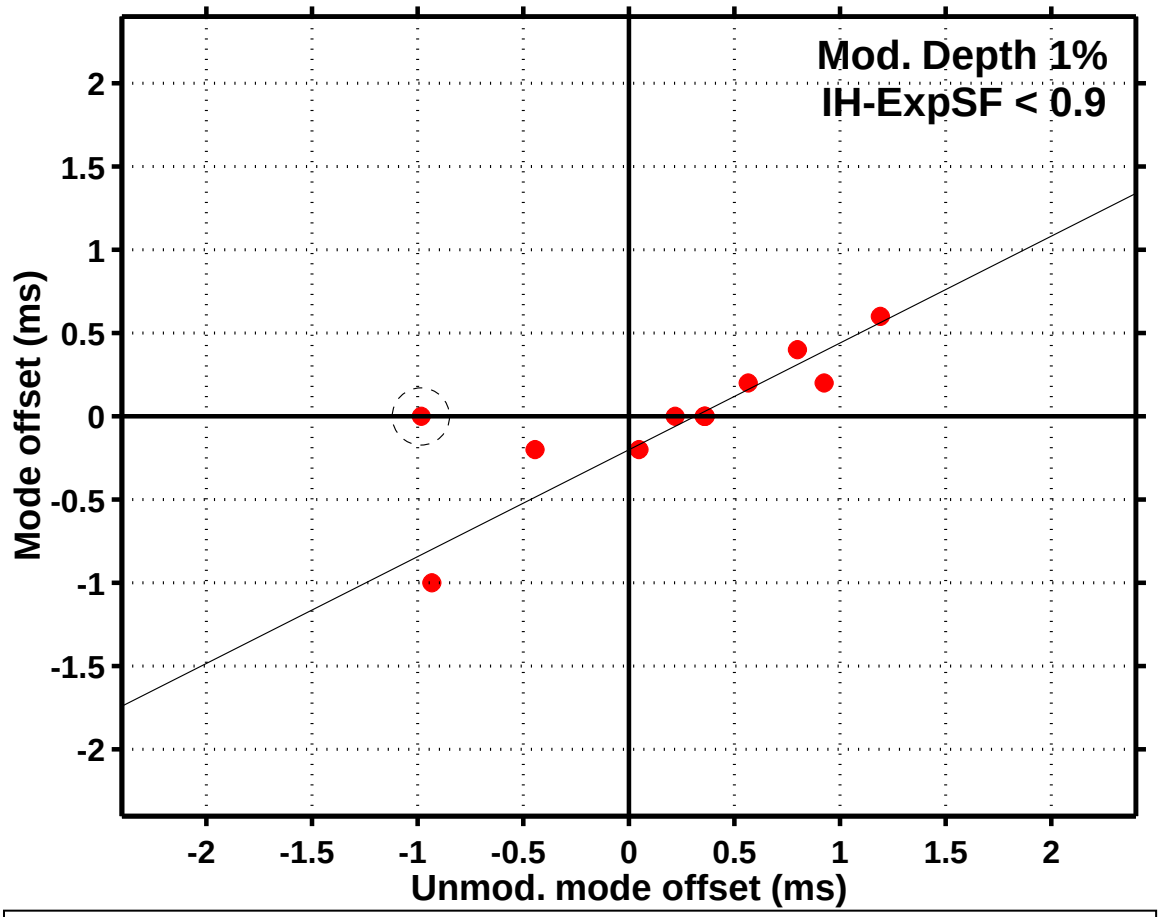


Figure 9. This figure plots the mode offset during the 420 Hz sinusoidal modulation (modulation depth of 1%) as a function of the relative location of the mode of the interval histogram computed from responses during the unmodulated DPT. Only the non-exponential responders (responders with IH-ExpSF below 0.9) were included in the computation, because only these responders contained a clear mode in their responses to the unmodulated DPT. The thin line represents a linear regression fit to the data (excluding the circled outlier). For the non-exponential responders, the modes during and prior to modulation are highly correlated.

(modulation depth < 2%) is affected by both the modulation frequency and the intrinsic dynamics of the fiber.

The response of the fiber shown in Fig. 7 suggests that non-exponential responders may be less synchronized to modulation. However, this observation is not borne out in all cases. Fig. 10 shows the synchronization index computed from responses to a 420 Hz sinusoidal modulator (modulation depth: 0.5%) against the interval histogram exponential shape factor (IH-ExpSF). For units with IH-ExpSF above 0.8, the synchronization index was uniformly high (between 0.3 and 0.5). In contrast, for units with IH-ExpSF below 0.8, the synchronization index varied from 0.1 to 0.5. The correlation between IH-ExpSF and synchronization index was highly significant (correlation coefficient 0.52, P=0.003, permutation test). We showed in the companion paper that non-exponential units respond

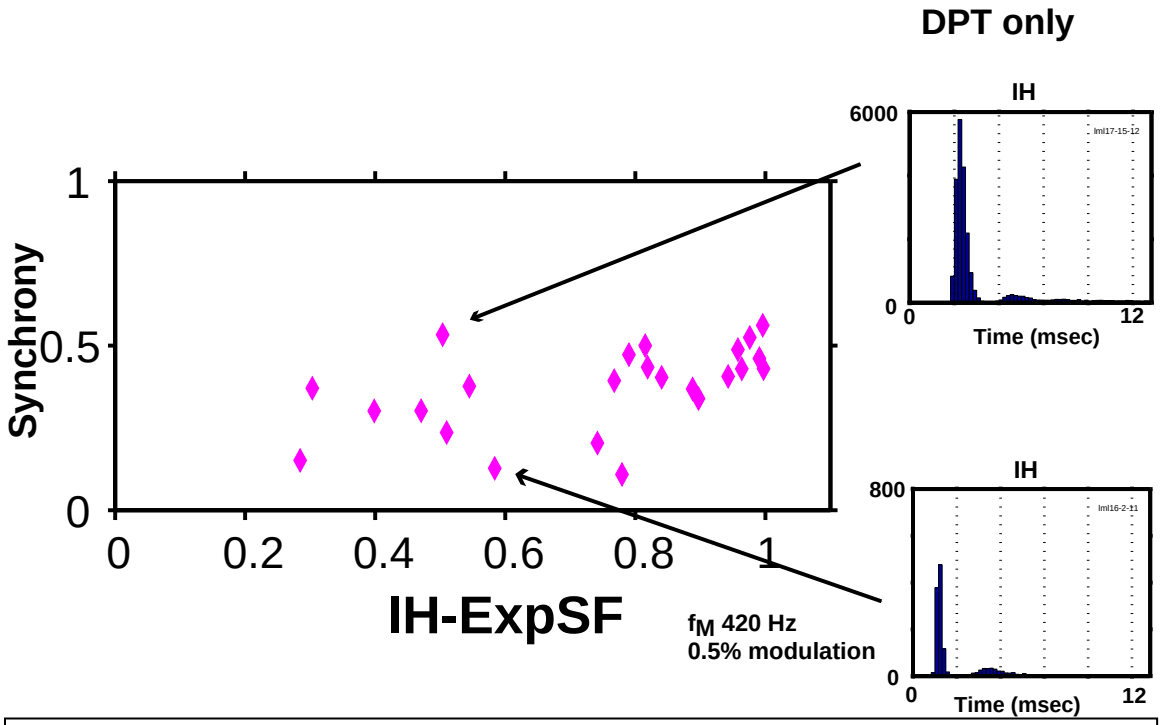


Figure 10. The center panel plots the synchronization index to 420 Hz modulator (modulation depth 0.5%) as a function of the interval histogram exponential shape factor (IH-ExpSF) for 25 fibers from 4 cats. The IH-ExpSF was computed based on responses during the unmodulated segments of the stimulus. The data show a significant correlation between the IH-ExpSF and the synchronization index. The bottom panel plots the synchronization index as a function of discharge rate during the unmodulated segments of the stimulus for the same fibers. The insets on the right are the interval histograms of responses to the unmodulated DPT for two units with low IH-ExpSF. The vertical dashed lines in the insets mark multiples of the modulation period. The top inset shows that the responses of this unit, which has a relatively high synchronization index, occur at intervals near the modulation period during the unmodulated part of the stimulus. In contrast, the unit with low modulation sensitivity has modes in the interval histogram that are far removed from the multiples of the modulation period.

to the DPT with higher discharge rates, so the observed correlation might reflect these rate differences. However, the synchronization index poorly correlated with discharge rate during the unmodulated DPT (correlation coefficient -0.02, P = 0.54, permutation test, not shown).

The variability in the synchronization index that is seen in neurons with IH-ExpSF below 0.8 can be understood by analyzing the relationships between the modes in the interval histogram for the unmodulated DPT and the modulation period. When the mode of the interval histogram is near the modulation period, synchrony during the modulation is high (upper inset). In contrast, when the mode does not coincide with the modulation period, synchrony to the modulator is low (lower inset). This further supports our hypothesis that the mode in the interval histogram during weak modulation (modulation

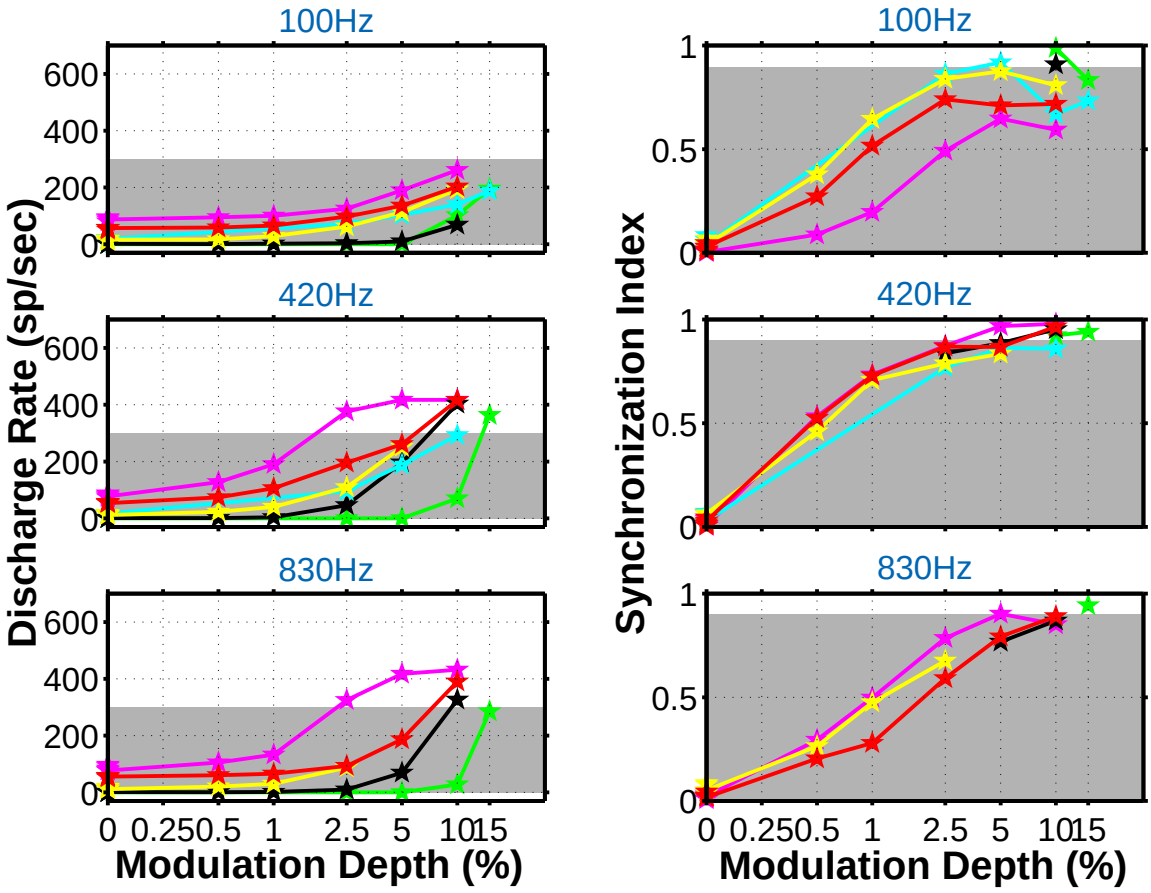


Figure 11. Each panel plots the discharge rate (left) or the synchronization index (right) for 6 fibers as a function of the modulation depth for different modulation frequencies (top 100 Hz, middle 420 Hz, bottom, 830 Hz). Different curves represent different fibers.

depth < 2%) is affected by both the modulation frequency and the intrinsic dynamics of the fiber.

D. Threshold and Dynamic range

Because greater modulation depths evoke higher discharge rates, changes in modulation depth can be used in a sound processor to encode sound pressure level. To determine the range of modulation depths for which the responses to modulated pulse trains in the presence of the DPT partially resembled acoustic responses to pure tones, we methodically investigated the discharge rate and synchrony as a function of modulation depth.

Fig. 11 plots the average discharge rate (left column) and synchronization index (right column) as a function of modulation depth for 6 randomly chosen fibers from our data pool. Each row shows responses to a different modulation frequency. The left-most point

in each curve represents response in the absence of modulation. In a healthy ear, responses of ANFs to pure tones saturate at discharge rates that rarely exceed 300 spikes/s (gray shading) (Liberman, 1978). The maximal discharge rate is even lower for low-spontaneous rate fibers. In the right panels, the gray areas indicate the range of synchronization indexes reported for responses of ANFs to low-frequency tones in a healthy ear (Johnson, 1980), which are below 0.9.

For fibers that have a non-zero discharge rate during the unmodulated portions of the stimulus (sustained responders), discharge rate grows slowly over a 20 dB range of modulation depths. This dynamic range is large compared to the 2-3 dB range reported for responses to electric modulated pulse trains without a DPT (Litvak et al., 2001). With one exception, for modulation depths below 5%, both discharge rates and synchronization indices were always in the acoustic range (gray area). In addition, synchrony for these sustained responders grew rapidly with modulation depth in the range 0.5 to 2.5%, while discharge rate increased minimally in this range. This behavior resembles responses of high-rate spontaneous fibers to tones in an intact ear (Johnson, 1980). At the higher modulation depths, most fibers had discharge rates that exceeded those of acoustically stimulated fibers. For 100 Hz modulator, as the modulation depth was increased, synchrony first increased and then decreased. This decrease could be accounted for by multiple response peaks in the period histogram as shown in Fig. 4 for the largest modulation depth.

Two fibers in Fig. 11 were transient DPT responders. Modulation depths necessary to evoke responses was higher than that for sustained DPT responders. In addition, on the logarithmic scale, discharge rates of transient DPT responders grew more rapidly than those of the sustained DPT responders, suggesting that fibers which respond transiently to the unmodulated DPT have a smaller dynamic range.

Two criteria were used to quantify the modulation detection threshold for individual fibers. Average rate threshold was defined as the modulation depth at which the average discharge rate during the 400 ms modulation exceeded the discharge rate during the following 400 ms unmodulated segment for 75% of the stimulus presentations. Synchronized rate threshold was defined as the modulation depth at which the synchronized discharge rate (synchrony coefficient times the discharge rate) during the

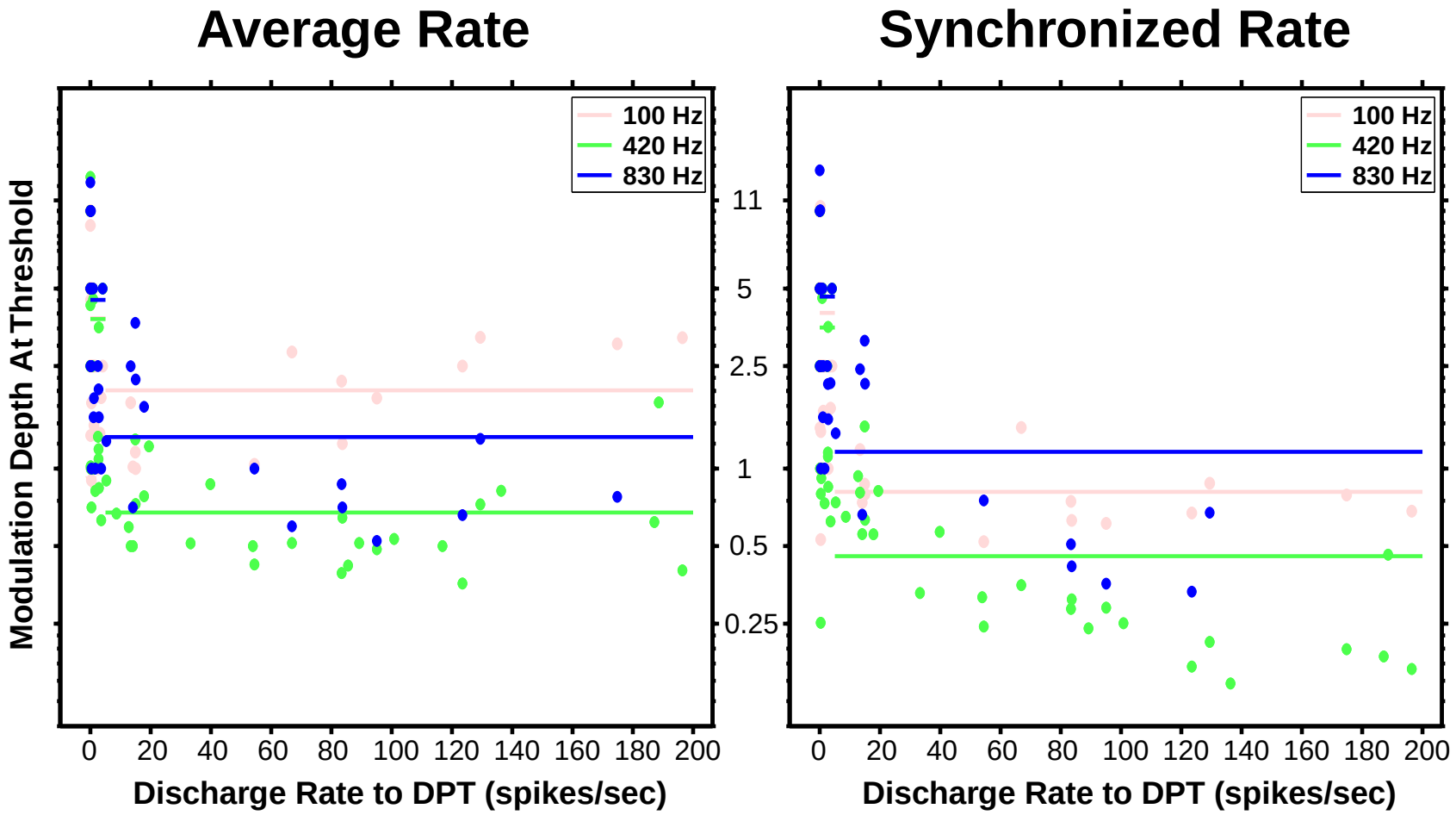


Figure 12. This figure plots the modulation detection threshold based on discharge rate (left) and synchronized rate (right) as a function of discharge rate. Color is used to indicate the modulation frequency. The solid lines indicate the average threshold for units with discharge rates above 5 spikes/sec during the unmodulated portions of the stimulus.

modulation exceeded the synchronized discharge rate during the following unmodulated segment for 75% of the stimulus presentations. This 75% criterion level was chosen because it is similar to that attained by up-down two-alternative choice procedures commonly utilized in psychophysics for threshold estimation (Levitt, 1971). Thresholds estimated by our method can therefore be directly compared to psychophysical thresholds. Because the sampling of modulation depths was rather coarse, and because most modulation depths were above threshold, thresholds were estimated using a special algorithm described in Appendix B.

Fig. 12 plots modulation thresholds based on the average discharge rate (left panel) and synchronized rate (right panel) against the discharge rate to the unmodulated DPT. Both the average rate and the synchronized rate thresholds depend on discharge rate during the unmodulated DPT. Fibers with DPT rates above 5 spikes/s had the lowest thresholds. For these fibers, the synchronized rate thresholds were lower than average-rate thresholds. A similar relationship between synchronized rate and discharge rate thresholds is observed in responses of high spontaneous-rate fibers to pure tones (Johnson, 1980). The difference between the average and the synchronized rates were largest for neurons responding at higher discharge rates during the unmodulated DPT, and could be as large as 7 dB. On the average, threshold is lower for the 420 Hz modulator than for the other two frequencies. The difference between the 100 Hz and 420 Hz mean thresholds is highly significant for both average and synchronized rate threshold ($P<0.005$, permutation test). Similar results hold for the difference between the 830 Hz and 420 Hz. If thresholds for individual fibers are compared instead of the population means, then the differences between the 420 Hz modulation and the other two modulation frequencies is even more pronounced. For the sustained DPT responders (fibers with discharge rate above 5 spikes/s during the unmodulated portions of the stimulus), 12 out of 13 fibers had the lowest synchronized rate thresholds for the 420 Hz modulator.

Transient DPT responders (with discharge rates during the unmodulated DPT below 5 spikes/s) had variable thresholds that were generally higher than thresholds for sustained DPT "responders." In addition, these fibers did not show preferred sensitivity to the 420 Hz modulator (14 out of 27 fibers were more sensitive to the 100 Hz modulator). The 830 Hz discharge rate thresholds were higher than the 420 Hz thresholds for 23 out of 26

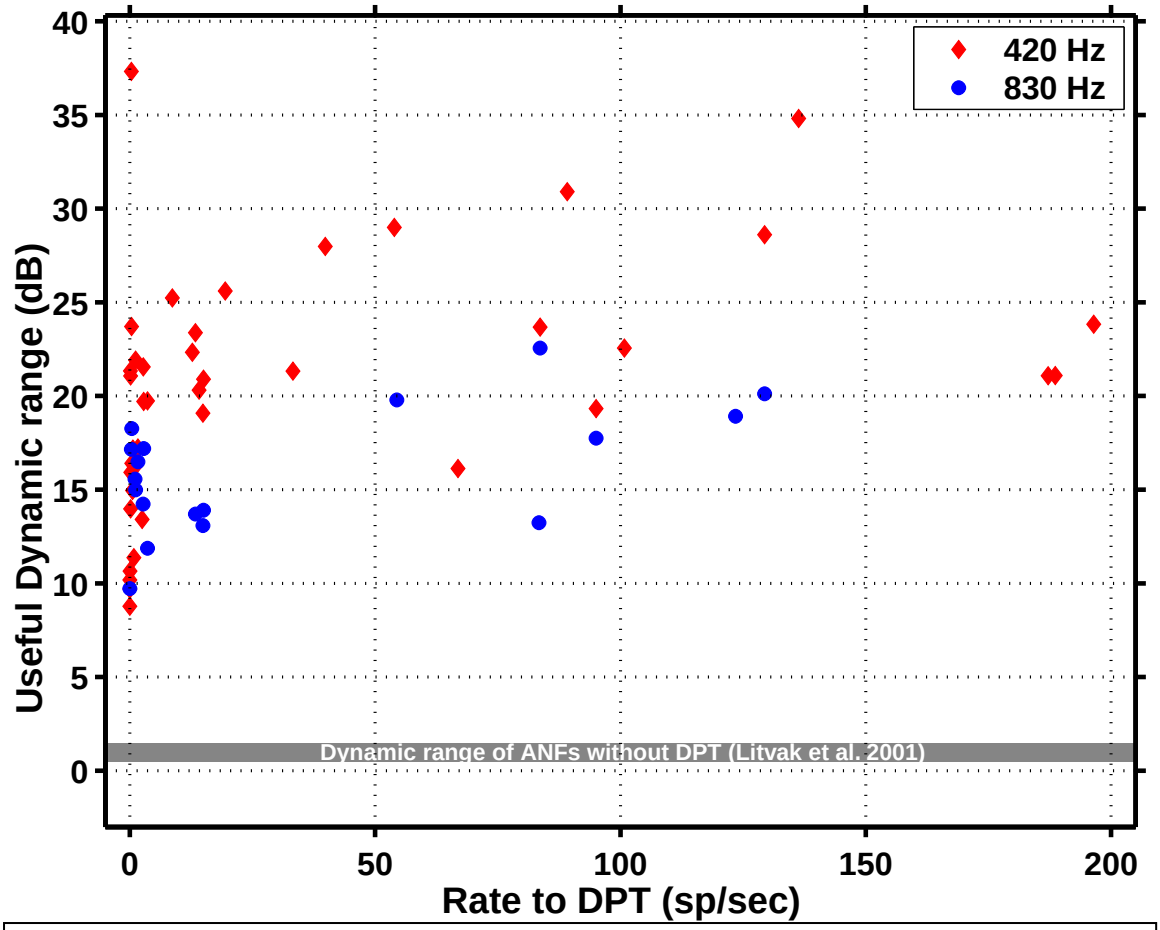


Figure 13. This figure plots the dynamic range in which the electric responses resemble acoustic ones versus the discharge rate during the unmodulated segments. The range of resemblance was defined as the ratio of the modulation depth that evoked a discharge rate of 300 spikes/s divided by the synchrony threshold. Different symbols represent different modulation frequencies. The analysis was not performed for modulation frequency of 100 Hz. For comparison, the gray area shows the approximate dynamic range of responders to electric modulated pulse trains without the DPT (Litvak *et al.*, 2001).

transient DPT responders. Thus, threshold sensitivity appears to be band-pass for the sustained DPT responders, and low-pass for the transient DPT responders.

We defined the "useful dynamic range" as the range of modulation depths for which the response discharge rates and synchrony resemble those seen in a healthy ear. The lower limit of the useful range is the modulation detection threshold based on the synchronized discharge rate. The upper limit is the modulation depth that evoked a discharge rate of 300 spikes/s. As shown in Fig. 11, for these discharge rates, most DPT sustained responders still had synchrony below 0.9, and were therefore within the range

of synchrony evoked by acoustic tones in a healthy ear (Johnson, 1980). The modulation depth that produced a 300 spikes/s discharge rate was estimated by fitting the rate versus modulation depth curve with a sigmoid function. The saturation of the sigmoid was assumed to occur at 420 spikes/s. This limit is appropriate for the 420 Hz modulator, in response to which most fibers saturated at one spike per stimulus cycle. Although at the largest modulation depths, responses to the 830 Hz modulator occasionally had discharge rates slightly higher 420 spikes/s, we chose to set the discharge rates for these points to 420 spikes/s prior to the fit. Our estimates of useful dynamic range do not strongly depend on this assumption.

Fig. 13 plots the "useful dynamic range" as a function of discharge rate during the unmodulated DPT. This analysis was only performed for the responses to the 420 and 830 Hz modulation frequencies, because responses to the 100 Hz modulation frequency that reached 300 spikes/s did not resemble the normal-ear responses, occurring more than once during a single phase of the modulator.

This figure confirms the trend suggested by Fig. 11, that the sustained DPT responders have larger dynamic ranges than transient DPT responders. The average useful dynamic range of sustained DPT responders is nearly 24 dB for the 420 Hz modulator (diamonds), and about 17 dB for the 830 Hz modulator (circles). The difference between the two means is highly significant ($P<0.01$, permutation test (Efron and Tibshirani, 1993)). Because multiple firings in a single phase in responses to 100 Hz modulators occur at relatively low modulation depths (Figs. 4,11), the useful dynamic range may be somewhat smaller for responses to 100 Hz modulators.

Transient DPT responders had useful dynamic ranges as low as 9 dB. This value is still considerably larger than the dynamic range reported for modulated electric pulse trains without a DPT, which is near 2 dB (gray bar) (Litvak et al., 2001). The DPT increases the dynamic range even for units that do not respond to the DPT. On the other hand, it is lower than the typical 20-40 dB dynamic range obtained with acoustic stimulation (Sachs and Abbas, 1974).

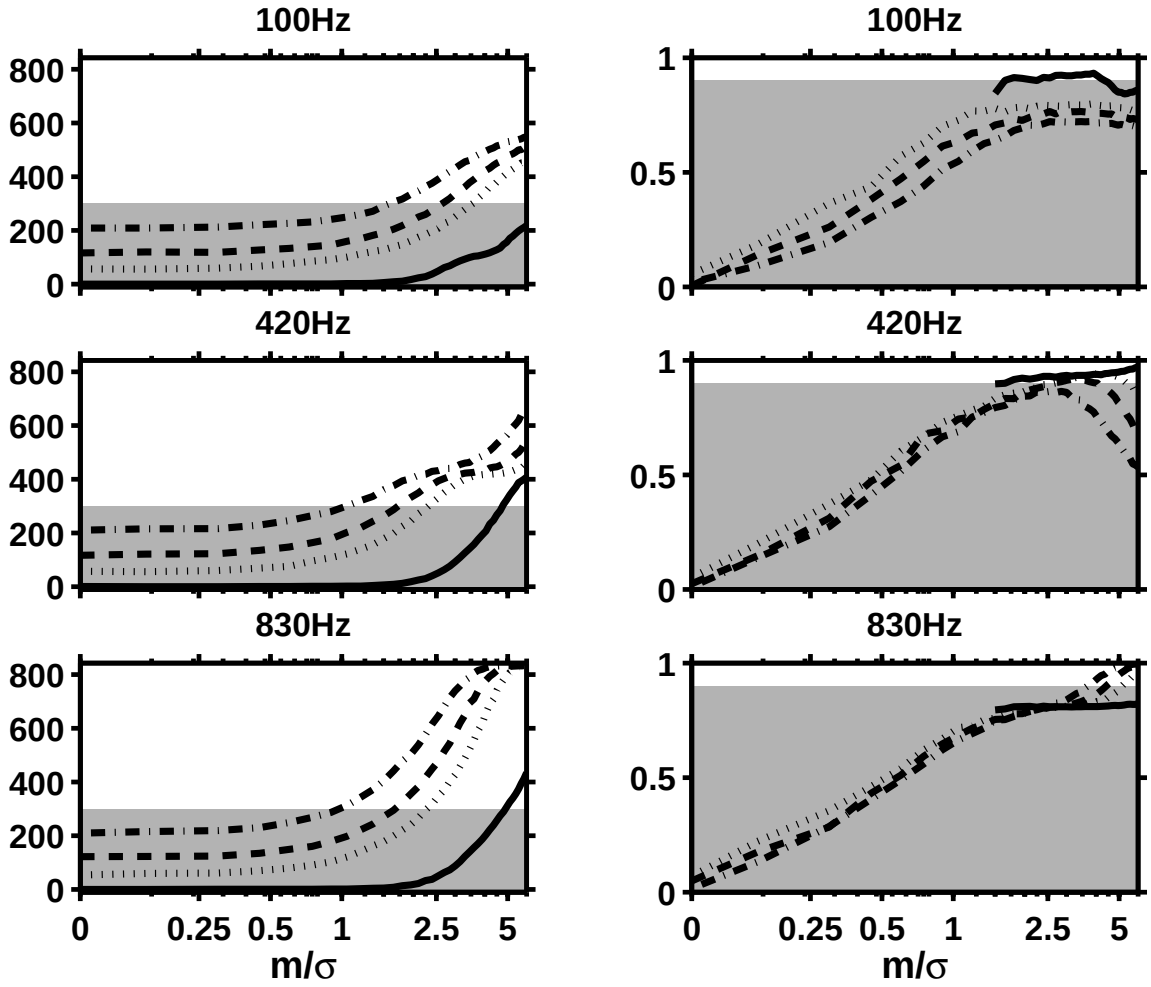


Figure 14. Discharge rate (left column) and synchronization index (right column) of the stochastic threshold model as a function of m/σ. Each panel corresponds to a response to different modulation frequency. The model was run for four choices of θ_0/σ: 1.4 (dot-dash), 1.8 (dash), 2.4 (dot), and 4 (solid).

E. Responses of the stochastic threshold model

Several key features of ANF responses to sinusoidally modulated DPTs are predicted by an extremely simple stochastic threshold model (STM). STM is composed of a noisy threshold that is compared with the modulation waveform $m(t)$ (Fig. 2). The response of the model is determined by two parameters: m/σ and θ/σ. Parameter m/σ characterizes the effective stimulus level. Parameter θ/σ determines the discharge rate in the absence of a modulator. We will refer to this rate as spontaneous rate. We will demonstrate that the relationship between the STM's spontaneous rate and modulation responses is similar to the relationship between discharge rate of ANFs during the unmodulated DPT and their responses to sinusoidal modulators.

Fig. 14 plots the predicted discharge rate (left column) and synchrony (right column) as a function of m/σ for fibers with different spontaneous discharge rates (different values of θ_0/σ). The format of the plot is similar to that of Fig. 11. The model predicts several features of the data. (1) Both discharge rate and synchrony grow over a wide range of modulation depths. (2) Both transient DPT responders and model fibers with low spontaneous rates have higher modulation thresholds and lower dynamic ranges than sustained DPT responders. (3) Both sustained DPT responders and model fibers with spontaneous discharge rates above 5 spikes/s have a range of modulation depths for which synchrony changes rapidly, while the discharge rate changes little. In contrast, both the transient DPT responders and model fibers with spontaneous rates below 5 spikes/s have high synchrony for all supra-threshold modulations. (5) For both the data and the model, synchrony does not grow monotonically with modulation depth for 100 Hz modulation. The non-monotonicity is related to multiple responses in each stimulus cycle (not shown).

There are also some differences between the model and the data. First, the maximum discharge rates exhibited by the model are larger those seen in the data. In addition, while neural synchrony grows monotonically with modulation depth for a modulation frequency of 420 Hz, synchrony can be somewhat non-monotonic for the model responses. Both behaviors suggest that the recovery function $r(t\text{-}t_i)$ inaccurately represents the true recovery function during DPT stimulation.

Fig. 15 plots the model averaged and synchronized rate thresholds as a function of spontaneous discharge rate. As in Fig. 14, threshold is expressed as a ratio m/σ. These plots should be compared to thresholds for neural data (Fig. 12). Note however, that while in the neural data, the largest discharge rate during the unmodulated DPT is below 200 spikes/s, we have extended the model to rates near 600 spikes/s.

The left panel plots the average rate threshold. For neurons with extremely low spontaneous rates, the rate threshold changes rapidly with spontaneous rate, with higher thresholds corresponding to lower spontaneous rates. For discharge rates above 5 spikes/s, the rate threshold decreases only slowly with discharge rate, and, when the discharge rate to the DPT exceeds the modulation frequency, the predicted average rate

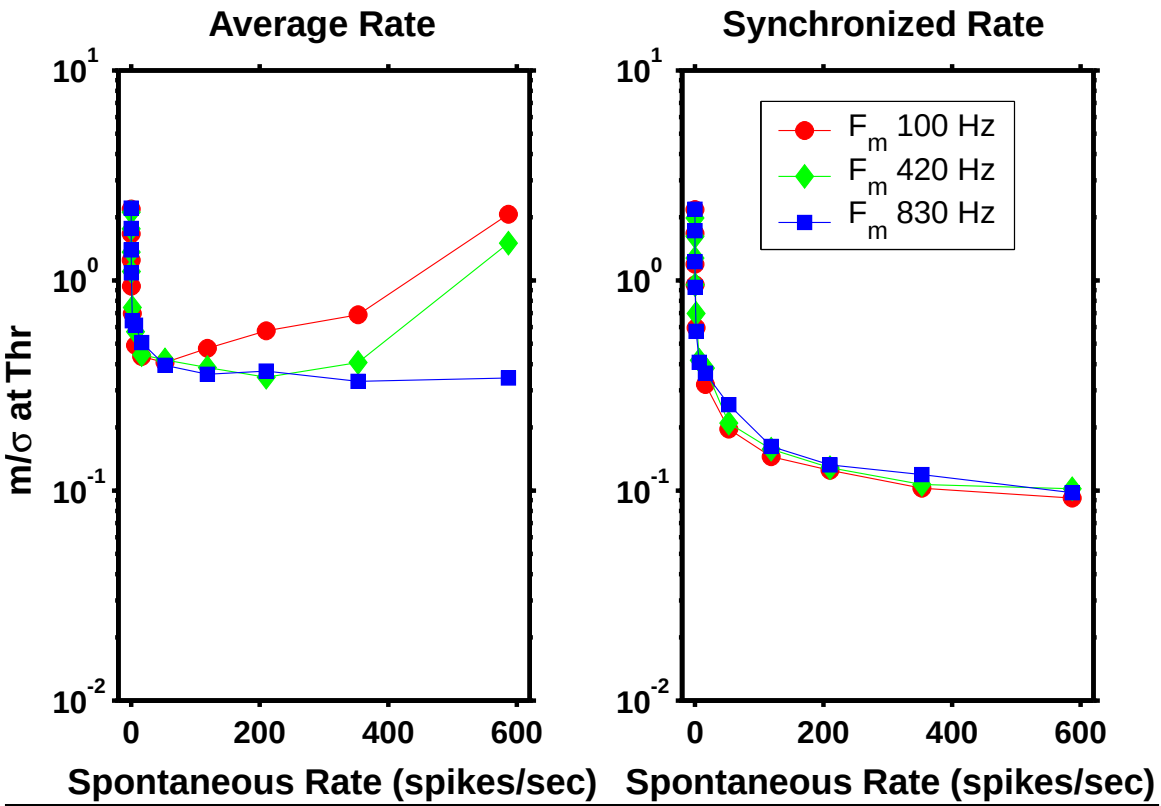


Figure 15. This figure plots the model's threshold to modulation (left: average rate threshold, right: synchronized rate threshold) as a function of spontaneous rate. Spontaneous rate was changed by changing θ_0/σ. Different curves indicate thresholds to different modulation frequencies (100 Hz circles, 420 Hz diamonds, 800 Hz squares).

threshold starts to increase. Similar behavior can be discerned in the data for the 100 Hz average rate thresholds in Fig. 12.

The predicted synchronized rate threshold (right panel) decreases monotonically with spontaneous rate for all modulation frequencies. Over the range of spontaneous rates from 5 to 200 spikes/s, the synchronized rate threshold decreases more rapidly than the average rate threshold. The same trend is exhibited by the data as well, particularly for the 420 Hz modulator. However, only the data show a dependence of threshold on modulation frequency (with lowest thresholds achieved for modulation frequency of 420 Hz). One way to incorporate dependence of threshold on modulation frequency is to assume that the modulation waveform is first filtered by a linear modulation filter. We will explore consequences of the modulation filter in a companion paper (Litvak et al., 2002b).

Fig. 16 plots the predicted useful dynamic range as a function of spontaneous discharge rate. As in Fig. 13, useful dynamic range is defined as a ratio of the modulation

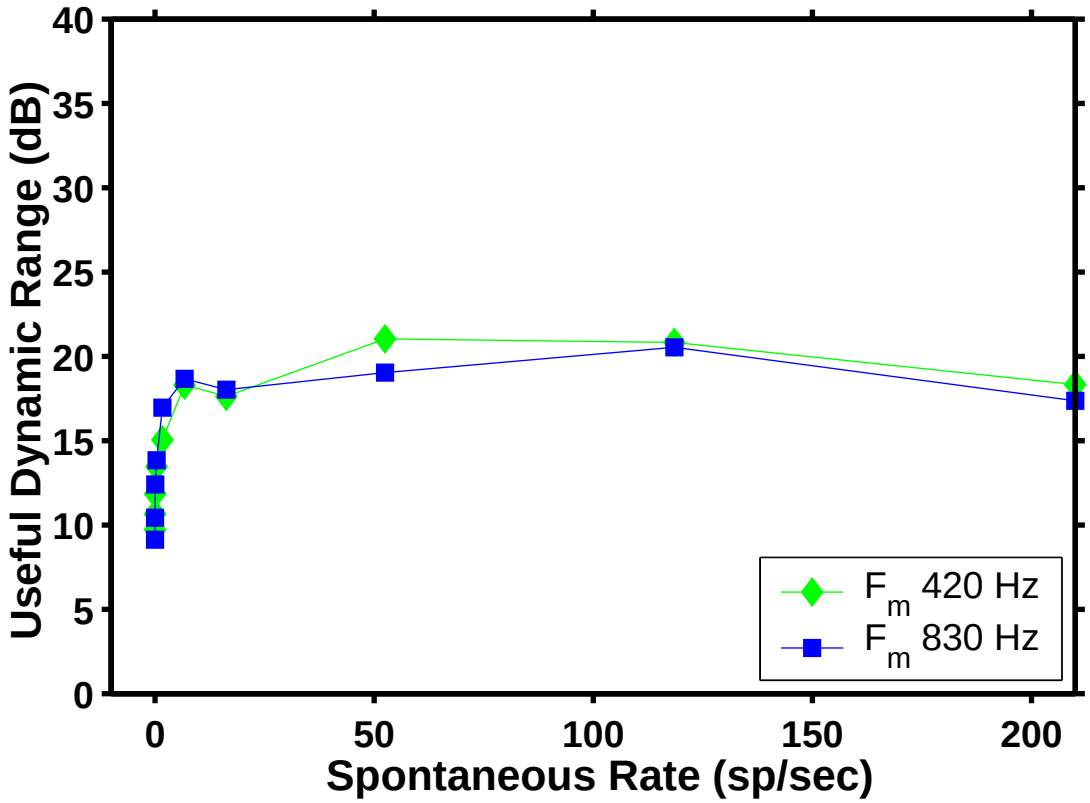


Figure 16. This figure plots stochastic threshold model's useful dynamic range for 420 Hz (diamonds) and 830 Hz (squares). The definition of useful dynamic range is the same as in Fig. 13. As before, useful dynamic range was only computed for 420 Hz and 830 Hz modulation frequencies.

depth that evokes a discharge rate of 300 spikes/s to the synchronized rate threshold. Because useful dynamic range is defined as a ratio of modulation depths, there are no free parameters in this plot. For neurons with spontaneous rates below 5 spikes/s, the model predicts that the dynamic range is strongly dependent on spontaneous rate, increasing with spontaneous rate. A similar trend is observed in the data in Fig. 13. In contrast, for higher spontaneous rates, the dynamic range of the model is nearly independent of spontaneous rate. The predicted useful dynamic range for these neurons is 19 dB for both frequencies. The observed average dynamic range is 24 dB for 420 Hz modulation frequency and 17 dB for 830 Hz modulation frequency. The model does not predict the difference in dynamic range between responses to the two modulation frequencies.

Overall, the STM predicts the relationship between discharge rate of ANFs during the unmodulated DPT and their responses to sinusoidal modulators. However, the STM cannot account for dependence of ANF sensitivity to modulation on modulation frequency.

IV. Discussion

A. Representation of frequency of the sinusoid

All neurons in our study responded to sinusoidal modulations of the DPT. Neurons that had sustained responses to the DPT were most sensitive to modulations. These neurons showed strong phase locking to the modulator for modulation depths as low as 0.5%. In these neurons, the average discharge rate grew monotonically with modulation depth over a range of nearly 20 dB. In addition, the temporal discharge patterns to modulation depths below 2.5% broadly resembled the acoustic responses of ANFs to pure tones. These responses were stochastic, with interspike intervals occurring at random at one, two, three or more stimulus cycles. The period histograms of these responses were similar to period histograms of responses to pure tones in a healthy ear. The discharge rates of these neurons were below 300 sp/s, consistent with rates evoked by tones in a healthy ear.

Although the interspike interval distributions show modes at multiples of the stimulus period for both electric and acoustic responses, there are some differences between the two. While in a healthy ear the interspike distributions of ANF responses to a 420 Hz tone show a large mode at the tone period (2.5 ms), this mode was absent in some ANF responses to 420 Hz modulations. In addition, some neurons whose responses during the unmodulated DPT had non-exponential interspike interval distributions had modes in their interval histogram that were systematically shifted away from multiples of the pitch period, particularly for low modulation depths. These responses are larger than the systematic offsets seen in responses to pure tones in a healthy ear (McKinney, 1998).

Although interspike distributions of some ANF responses to small modulations differed from ANF responses of a healthy ear to tones, these distortions may not severely interfere with accurate representation of the stimulus frequency. Some neurons in the auditory brainstem are likely to integrate information across many fibers. Because there is no evidence for cross-fiber correlation in responses to an unmodulated DPT (Litvak et al., 2002c), responses to small modulations of the DPT may also be uncorrelated from one fiber to the next. By responses to modulations being uncorrelated we mean that a response on a particular cycle of the modulator on one fiber should not make it more likely that a response would occur on the same cycle on another fiber. By integrating

across several auditory-nerve fiber inputs, some central auditory neurons could contain a mode at the stimulus period in their interval histograms even if such a mode is lacking in the ANF inputs. Similarly, because deviations from the multiples of the stimulus period differ from one unit to the next and can be positive or negative (Fig. 8), these deviations may be averaged out at the output of the integrating central neurons. Nevertheless, both distortions could degrade somewhat the accuracy of frequency representation. Direct recordings from central neurons (e.g. cochlear nucleus) while stimulating with a modulated DPT would test whether integration suffices to properly encode the modulation frequency.

For a 10% modulation depth, neurons that had a sustained response to the DPT exhibited responses to sinusoidal modulations that differed from ANF responses to pure tones. These modulation responses had higher discharge rates than those seen during acoustic stimulation. In addition, for modulation frequencies above 420 Hz, some neurons fired on every other stimulus cycle. These neurons were more precisely phase locked than the acoustic responses.

Neurons that responded transiently to the unmodulated DPT nevertheless responded to sinusoidal modulations. However, larger modulation depths were needed to produce a response in these neurons. At these modulation depths, the responses of these neurons were more synchronized than responses to sounds in a normal ear, and poorly represented the stimulus waveform. In addition, transient DPT responders had a smaller useful range than sustained DPT responders. In this respect, transient DPT responders differ from the low spontaneous-rate ANFs in a healthy ear, which have a larger dynamic range than do high spontaneous-rate fibers (Liberman, 1978). The transient responders comprised roughly 50% of the total number of neurons that we recorded from. Even though a strategy that uses only small modulation depths may stimulate only half of the neurons that are excited electrically, it may still produce substantial benefit by mimicking some aspects of normal ANF responses in the neurons that it does stimulate.

B. Comparison with models of the electrically stimulated fibers

We presented the STM as a model of ANF responses to modulations of the DPT. Despite its simplicity, the STM accounts for the dependence of both threshold and dynamic range on the average discharge rate during the unmodulated DPT. In addition,

STM quantitatively predicts (within 2-5 dB) the average useful dynamic range of ANFs. However, the dependence of threshold sensitivity on modulation frequency is not predicted by the STM. A parsinomous way to account for that dependence would be to assume that the modulation is first filtered by the neuron using a linear band pass filter. We show in a companion paper, that a linear filter is also necessary to account for ANF responses to complex modulators.

It is at first sight surprising that the dependence between the discharge rate during the unmodulated DPT and useful dynamic range are quantitatively predicted by a very simple stochastic threshold model. We believe that the agreement may not be accidental. Many complex neural models can be approximated by a "driving" function, which depends only on the stimulus, and another (possibly noisy) function describing threshold, which depends only on previous activity (e.g. Hill (1936)). In general, the driving function may depend on the stimulus in a complex way. However, because the stimuli in our study are composed of a large signal (the DPT) which is perturbed by small modulators, driving function may be approximated by linearization. Thus, to the extent that the threshold part of the model captures the complex dynamics of recovery of ANFs from the spike, for small modulation depths the stochastic threshold model can be considered an approximation of a more general neural model.

In some respects, responses of ANFs to a modulated DPT also resembled the responses of a biophysical model to electric sinusoids in presence of a 5 kpps conditioner (Rubinstein et al., 1998). At low levels of the electric sinusoid, period histograms of model responses resembled the sinusoidal waveform. We obtained a similar result for responses to small modulations of the DPT. The useful dynamic range of the biophysical model responses was near 20 dB, and changed little with DPT level, so long as it evoked discharge rates from 10 to 250 spikes/s. The measured responses to the modulated DPT showed a similar trend, with no systematic relationship between the sustained discharge rate during the unmodulated DPT and the useful dynamic range for discharge rates that were larger than 5 spikes/s (Fig. 13). Because Rubinstein *et al*. did not report interval histograms of model responses during the sinusoid, it is unclear whether their model captures the mode shifts, and the missing first mode reported for some modulation responses. In conclusion, although responses of the biophysical model to a modulated

DPT remain to be studied, the data so far suggest that the biophysical model of Rubinstein *et al*. (1999b) captures the main features of responses of ANFs to sinusoidal electric stimulation.

C. Comparison with psychophysical modulation detection thresholds

It is of interest to compare modulation detection thresholds of single neurons estimated in this study with the psychophysical thresholds reported by Shannon (1992) for human cochlear implant listeners. In that study, subjects were stimulated continuously with a 0.5-2 kHz sinusoid. Beats were produced by presenting another sinusoid with a frequency slightly off the sinusoid frequency. Because our threshold criterion closely parallels that used in psychophysics, direct comparison of threshold values is appropriate.

Shannon reported that subjects were most sensitive to 100 Hz beats. At these frequencies, a modulation depth of 1% could be detected. These modulation depths are comparable to the most sensitive rate modulation detection thresholds that we observed for 100 Hz modulators, and are similar to the average synchronized rate detection thresholds. Unlike the neural data, however, sensitivity of human subjects to beats dropped rapidly with increasing beat frequency. The best psychophysical modulation detection thresholds for the 420 Hz modulator were near 3%. In contrast, the neural detection thresholds were lower for the 420 Hz than for the 100 Hz modulator, and were as low as 0.5%.

Shannon (1992) suggested that the drop-off in the subjects' ability to detect modulations with increasing beat frequency may be related to a central limitation in detecting high-frequency modulations. Because the peripheral neurons encode 420 Hz modulations as well as, or better than, the 100 Hz modulations, our data supports this view. However, the comparisons between the physiological thresholds reported in this study for a 420 Hz modulator and psychophysical results of Shannon (1992) should be interpreted with caution. It is possible that the difference between the neural and the psychophysical data can be accounted for by the differences in the stimuli used in the two studies (sinusoidal modulations of a 5 kpps pulse train versus beats of a 0.5-2 kHz sinusoid). Psychophysical study of stimuli that are more similar to the ones used in this study should be conducted to resolve the question of whether small, high-frequency modulations of a DPT can be detected by human listeners.

D. Implications for cochlear implant processors

Overall, our results suggest that a DPT can enhance the representation of the temporal fine-time structure in auditory nerve responses for frequencies up to at least 800 Hz. The enhancement is maximal for small (<5%) modulations of the DPT. In addition, responses to such small modulators in many respects resemble responses to sound in a healthy ear in terms of discharge rate and synchrony. These observations suggest that a strategy that properly incorporates a DPT, and that preserves the temporal fine-time structure information may improve performance of cochlear implant users. In particular, proper coding of the sound fine-time structure may lead to improved perception of musical pitch, and to more effective utilization of binaural cues in bilateral implants (Smith et al., 2002).

One modern strategy that preserves the fine-time structure information in the electric stimuli is the continuous analog (CA) strategy. Because a DPT may allow for better encoding of the fine-time structure, performance of CA strategy may be substantially enhanced by addition of the DPT. In the current implementation, the CIS strategy throws out the fine-time structure at the envelope detector stage. The function of the envelope detector is to produce a strictly positive modulator waveform. Because negative waveforms could be represented with the DPT as decreases in DPT level, it may be possible to entirely eliminate the envelope detector stage in a DPT-enhanced CIS strategies.

1. How to present a DPT

We have demonstrated that sustained DPT responders are best at representing the modulations. In the companion paper, we showed that the DPT needs to be at least 4 dB above CAP threshold to evoke a sustained response in a sizeable fraction of neurons, but that DPT levels that are more than 8 dB above CAP threshold evoke long-lasting changes in neural excitability, and may therefore be harmful to the nerve. To stimulate a large number of fibers, while simultaneously meeting the safety constraints, it may be optimal to present a DPT at a low (4 dB above CAP threshold) level on several electrodes. However, more work needs to be done in working out the safety of such stimulation.

2. Effect of the DPT on threshold and dynamic range

Our results predict that a DPT would decrease thresholds to electric stimulation. In the case of CIS, if the CIS carrier and the DPT are in phase, then our results suggest that

there would be sufficient information on the auditory nerve to detect modulations as small as 0.5% of the DPT level for 420 Hz modulation, and near 1 to 2% for 100 Hz and 830 Hz modulations. While we predict that a DPT would increase the dynamic range of individual neurons, it is unclear whether the psychophysical dynamic range would also be increased. This is because the psychophysical range is determined by the distribution of ANF thresholds as well as by the dynamic range of individual fibers. Because a DPT is expected to modify the threshold distribution, its effect on the psychophysical dynamic range is unclear.

V. General conclusion

We showed that, using a DPT, it may be possible to design strategies for cochlear implants that would properly code the fine-time structure of the sound waveform in the temporal discharge patterns of the electrically-stimulated auditory-nerve fibers. If the central nervous system can interpret this information, then these strategies may substantially improve the quality of auditory experience enjoyed by cochlear implant users.

VI. Appendix A: Methods for artifact cancellation

To remove the artifact from modulation responses off-line, we first filtered the waveform using a boxcar filter of length equal to the carrier period (0.2 ms). This filter removed most of the 5 kpps periodic component from the artifact. For modulation depths below 2.5%, the remaining artifact was smaller than the noise in the recording for nearly all of the records. For modulation depths between 2.5 and 15%, the artifact was of comparable level to the spikes. To remove the remaining artifact from these responses, we developed two algorithms, one for modulation frequencies below 500 Hz, and the other for higher frequency modulations. Both algorithms assumed that (1) the remaining artifact is nearly periodic with a period equal to that of the modulator, (2) the neural responses are composed of spikes, and (3) the spike waveform is known within an overall gain, which we allowed to change slowly throughout the recording. Spike shape was determined by recording the responses to an unmodulated DPT. Specifically, the recorded response $y_k(t)$ during the k^{th} modulator period is assumed to be

$$y_k(t) = a(t) + x_k(t) + n(t), \qquad \text{(A1)}$$

where $a(t)$ is the periodic artifact, $x_k(t)$ is composed of spikes, and $n(t)$ is measurement noise.

A. 830 Hz modulation frequency

One algorithm was particularly effective in isolating responses to 830 Hz modulators. Using a recursive procedure, the algorithm determines $\hat{a}(t)$ that minimizes the energy E in the error signal $e_k(t) = \hat{x}_k^s(t) - \hat{x}_k(t)$, where $\hat{x}_k(t) = y_k(t) - \hat{a}(t)$ is the estimated neural response, and

$$\hat{x}_k^s(t) = \sum_j g_j \cdot s(t - \tau_j) \qquad \text{(A2)}$$

is the model neural response composed only of spikes. j is the spike number, $s(t)$ is the known spike waveform, while g_j and τ_j are spike gains and spike times respectively. In each step, g_j and τ_j were estimated by computing a running correlation between $\hat{x}_k(t)$ and $s(t)$, and determining the peaks in the correlation that are larger than a criterion threshold. We used a method of gradient descent (Menzel, 1960) to determine the $\hat{a}(t)$ that minimizes E. As an initial guess for $\hat{a}(t)$, we used the mean (over k) of the recorded waveform $y_k(t)$. The algorithm was stopped whenever the maximum of the error $e(t)$ averaged over all modulator periods was $1/50^{th}$ of the spike height.

B. Lower modulation frequency

Although the method described above was highly effective in removing artifact for 830 Hz modulators, it was often less so in removing artifact from responses to lower modulation frequencies. The reason for this is that the initial guess used (the mean of the neural response) is closer to the true artifact for the higher frequency responses, where the spikes occur at most during every other modulator period, than for lower modulation frequencies, where the spikes can occur in every modulator period. Another method was developed to specifically remove artifact for modulation frequencies of 100 and 420 Hz.

The top panel of Fig. 17 shows the superimposed response $y_k(t)$ of fiber lml17-38 to 420 Hz modulators for 25 successive modulator periods. Over a portion of the modulator period (bar), the variability is very low, indicating that there are no spikes during this time period. Therefore, the artifact over that period is simply the average (over k) of the

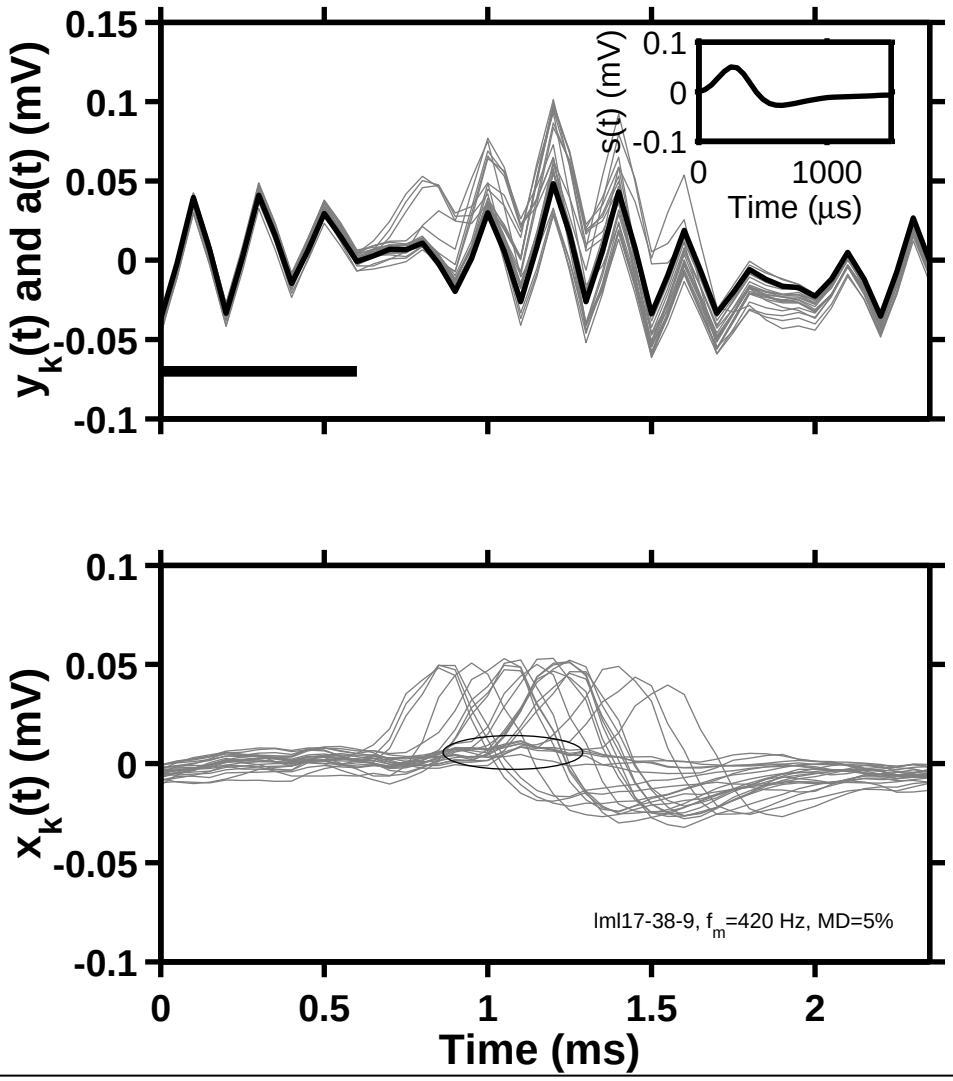


Figure 17. Algorithm for removing artifact from the contaminated recordings allows for accurate detection of spikes. *Top panel*: preprocessed response of fiber lml17-38 to a sinusoidally modulated DPT (modulation frequency: 420 Hz, modulation depth: 5%). 25 successive cycles of the response $y_k(t)$ are shown (gray, thin lines). Spikes can be seen in this response as variations in response between 0.2 and 2 ms. The thick black line indicates the artifact estimated using the first algorithm described in Appendix A. *Bottom panel*: estimated neural response. Circle indicates where the residual artifact is largest.

neural response $y_k(t)$. In contrast, the rest of the modulator period, the response is highly variable, indicating that it is composed of both artifact and spikes. The algorithm next proceeds recursively, starting at the end of the portion of the modulator period for which the artifact is known, to extend the estimate of the artifact from *t-1* to *t*, until the artifact is estimated over the entire modulator period. If the artifact is known from times 0 to *t*-1, then the neural response over that period can be estimated by solving Eq. A1. This neural response is composed of spikes of the type shown in the inset of the top panel of Fig. 17.

Because the responses are preprocessed by differentiation, the spikes are generally biphasic, with a positive initial phase.

The algorithm next determines which periods contain the initial phase of a spike occurring earlier than *t* in the period by cross-correlating the estimate of the neural responses prior to time *t* with the spike waveform. If it is determined that a period contains a spike prior to *t*, then the generic spike waveform is used to estimate the contribution of that spike to the response *y*(*k*,*t*) at *t*. Finally, the artifact at *t* is estimated as the *minimum* (over *k*) of $y_k(t) - \hat{x}_k(t \mid t-1)$ where $\hat{x}_k(t \mid t-1)$ is the extrapolated neural response based on responses up to *t*. The rationale for this estimate is that, for each modulator period *k*, $y_k(t) - \hat{x}_k(t \mid t-1)$ is either equal to (1) the artifact alone, or (2) the artifact plus a spike that starts at *t*. Because the initial phase of the spike is positive, then Case (2) periods will have larger values of $y_k(t) - \hat{x}_k(t \mid t-1)$ than Case (1) periods. Assuming that both types of periods are always present in the response, the minimum response will estimate the artifact. The top panel of Fig. 17 shows an example of the estimate of *a*(*t*) (thick line) derived by the algorithm from the recordings (thin, gray lines). The lower panel shows the neural responses derived using Eq. A1. Although most of the artifact has been removed, the neural responses appear to have some residual artifact near 0.6 ms (circle). Because the residual artifact is at most 10% of the spike height, it does not interfere with spike detection.

VII. Appendix B: Method for estimating modulation threshold

Single fiber modulation threshold was defined in this study as the level at which either the discharge rate or the synchronized rate during the modulation exceeded the discharge rate during a following unmodulated segment on 75% of stimulus trials.

The most straightforward way to determine threshold would be to compute the percentage of trials in which the rate during the modulation exceeded the rate during the unmodulated DPT for each modulation level. We denote such trials as hit trials. To reliably estimate threshold, data need to span the 75% criterion. However, this condition was rarely met, particularly when measuring synchronized rate thresholds for responses

to a modulation frequency of 420 Hz. A common pattern in some 30-40% of the data for that measurement was a percentage of hits between 75% and 95% for the smallest measured modulation depth, and 100% for larger modulation depths (3 out of 4 fibers in Fig. 18, left panel). To extrapolate these data to threshold reliably, they would need to be fit to an accurate model. Because the dependence of hit percentages on modulation depth can differ in both threshold and growth rate, an adequate model would have to include at least these two parameters. However, when there is only one modulation depth with a percentage of hit trials below 100%, there is insufficient data to estimate both parameters independently.

We developed an alternative method for determining threshold which does not directly depend on measurement of hit percentages. The goal of the new method is to utilize data points for which hit percentage is 100%. This method assumes that, over a range of modulation depths, rate of growth of discharge rate near threshold is similar to the rate of growth of discharge rate well above threshold.

To this end, we developed a measure, which we call quartile rate difference, which is 0 when hit percentage is 75%, and depends monotonically on modulation depth. For each modulation depth m, and each trial i, let $r_m(m,i)$ denote the rate during the modulation (rate can be either the synchronized or the average discharge rate). Let $r_u(m,i)$ denote the rate during the corresponding unmodulated segment. Let $r_d(m,i)= r_m(m,i)-r_u(m,i)$. Let the quartile rate difference $r_d^q(m)$ be the 25^{th} percentile of $r_d(m,i)$ over i. At modulation threshold, 75% of the rate differences are positive (the rate during the modulator exceeds

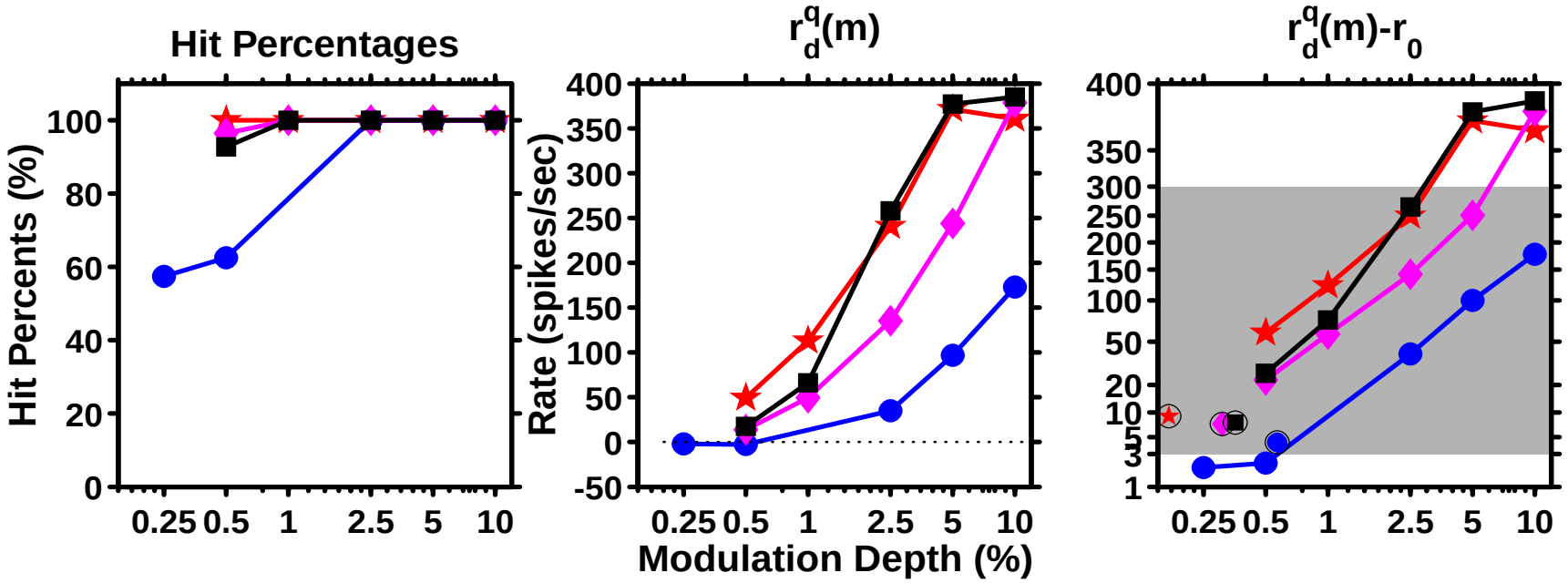


Figure 18. The left panel shows percentage of hit trials versus modulation depth for 4 sustained DPT responders. For 3 of these fibers, hit percentage was below 100% for only one modulation depth. The middle panel plots rd(m) versus modulation depth for the same 4 fibers. The right panel replots the same data using a scale which described by Eq. A4. The gray area indicates the region where the data are approximately linear. The circled points indicate thresholds.

the rate during the corresponding unmodulated segment), and 25% are negative, so $r_d^q(m)$ must be 0. Thus, to determine threshold, we need to determine m_{thr} such that $r_d^q(m_{thr}) = 0$.[3]

The middle panel of Fig. 18 plots the quartile rate difference as a function of modulation depth for the same units as in the left panel. The threshold criterion of 0 is shown as a dotted line. Theoretically, as *m* approaches 0, $r_d^q(m)$ approaches a negative constant r_0. r_0 equals the 25% percentile of $r_u(m,i)$-$r_u(m,j)$ for all pairs (i,j), and was estimated directly for each fiber from responses to the unmodulated DPT. For 55% of the synchronized rate records, $r_d^q(m)$ crossed zero, and threshold was estimated by linearly fitting $r_d^q(m)$ between modulation depths that span zero. However, for many records from the more sensitive sustained DPT responders (such as 3 of the 4 shown in the panel), $r_d^q(m)$ was above 0 even for the smallest *m* recorded. For these data, $r_d^q(m)$ could often be modeled by:

$$r_d^q(m) = r_0 + g_\Gamma \cdot \Gamma(\alpha \cdot \ln(m) + T_0) \qquad \textbf{(A3)}$$

$\Gamma(x)$ is a cumulative Gaussian function, g_Γ is a gain, and α and T_0 are free parameters. This model implies that that,

$$\Gamma^{-1}\left(\frac{r_d^q(m) - r_0}{g_\Gamma}\right) \qquad \textbf{(A4)}$$

and ln(*m*) are linearly related. Fig. 18 (right panel) plots $r_d^q(m) - r_0$ versus *m* for the same data as in the other panel. The abscissa is logarithmic, and the ordinate uses a scale defined by Eq. A4. In practice, the gain g_Γ was set to the modulation frequency for 100 Hz and 420 Hz modulations, and to half the modulation frequency for 830 Hz modulation. In the shaded region (below 300 spikes/s and above 3 spikes/s) the data grow nearly linearly, as predicted by the model. Outside the shaded region, the data deviate from linearity, perhaps because of difficulties in accurately estimating a Gaussian

[3] This problem is ill-defined for transient responders, because $r_d^q(m)$ is zero for small *m*. To resolve this ambiguity, a small random number (at most corresponding to half a spike) was arbitrarily added to each $r_m(m,i)$ and $r_u(m,i)$ that equaled 0. Because thresholds of transient responders are typically elevated, and are typically larger than the smallest modulation depth recorded, this addition has only a minor effect on modulation threshold.

function near its saturation. For most fibers, at least 3-4 data points are inside the shaded region. This contrasts with the hit percentage (left panel), where only one point was often available for estimating threshold.

To estimate threshold, the free parameters α and T_0 were first estimated using nonlinear least squares. This procedure was applied to 45% of the fibers. For 34% of the total number of records, the fit was sufficiently close, and the threshold was estimated by inverting Eq. (3). Because of insufficiently close fit, or lack of data, threshold could not be determined for 11% of the fibers. These fibers are not represented in Fig. 12. Most of these fibers (9% of total) were transient responders, whose thresholds were higher than the highest modulation depth tested. For the other 2% of fibers, $r_d^q(m)$ was above criterion for all modulation depths; presumably, therefore, these fibers have thresholds that are similar to those of the more sensitive fibers.

Chapter 5

Responses of auditory-nerve fibers to sustained high-frequency electric stimulation. III. Vowel modulators

I. Introduction

In a healthy ear, auditory-nerve fiber (ANF) responses encode the fine-time structure in each frequency band in their temporal discharge patterns for frequencies up to 5 kHz (Johnson, 1980). For a complex sound, the auditory nerve is able to code several components simultaneously in the temporal discharge patterns. During a sustained vowel, for example, both the information in the envelope and the information in the fine-time structure (e.g. formant frequencies) are represented in the temporal patterns of ANF responses (Young and Sachs, 1979; Miller and Sachs, 1984; Delgutte, 1984). Future cochlear implant devices may attempt to reproduce these temporal discharge patterns in the deaf cochlea. Doing so may provide more information to the implantees about the incoming sounds than is available through existing implants.

Although it would be conceptually straightforward to include the fine-time structure information into the electric stimuli generated by the cochlear implant processor, it is unclear whether the fine-time structure would be properly represented in temporal discharge patterns of ANFs. Responses to electric vowels occur primarily at the largest peak of the vowel waveform over a wide range of vowel amplitudes (van den Honert and Stypulkowski, 1987b). Thus, only the pitch envelope, and not the fine-time structure, is coded reliably in ANF temporal discharge patterns. This is true even if vowels are compressed to the psychophysical dynamic range (Knauth et al., 1994).

Rubinstein *et al*. (1999b) proposed that the coding of the fine-time structure of the electric stimuli might be improved by introducing a sustained, high-frequency, "desynchronizing" pulse train (DPT). In a previous report, we recorded ANF responses to sinusoidal modulation of sustained, 5 kpps DPT (Litvak et al., 2002a). We found that,

over a range of modulation depths, sinusoidal modulations produced responses on the auditory nerve that mimicked the responses to tones in a healthy ear in several respects. In particular, the responses were stochastic, occurring every one, two or more cycles, and were less synchronized to the modulator frequency than responses to modulated pulse trains without a DPT (Litvak et al., 2001). In addition, response rates and synchronization indices grew over a wide (15-25 dB) range of modulation depths. The large dynamic range to sinusoidal stimulations suggests that complex modulations that have peaks of differing height in their waveform may be properly represented in the temporal discharge patterns of auditory-nerve fibers. To test this hypothesis, in this paper we recorded responses of ANFs to complex filtered-vowel modulations of the DPT. We focused on vowel modulators because of the perceptual importance of vowels, and because their representation in a normally-functioning ear has been extensively characterized (Young and Sachs, 1979; Miller and Sachs, 1984; Delgutte, 1984).

In the previous paper, we proposed a simple stochastic threshold model that describes responses to modulations of the DPT. For sinusoidal modulators, the model was successful in predicting dependence of threshold and dynamic range on discharge rate during the unmodulated DPT. We suggested an elaboration of the stochastic threshold model incorporating a linear modulation filter that would be able to predict responses to arbitrary complex modulators. The modulation filter is introduced to account for differences in sensitivity across modulation frequency. In pursuit of that goal, we estimated the modulation filter directly from the vowel data, and compared the ANF responses to vowel modulators with predictions of the modified stochastic threshold model.

II. Methods

The animal preparation, electrical stimulation, and recording methods have been described previously (Litvak et al., 2002c). Briefly, cats were anesthetized with dial in urethane (75 mg/kg), then deafened by co-administration of kanamycin (subcutaneous, 300 mg/kg) and ethacrinic acid (intravenous, 25 mg/kg) (Xu et al., 1993). Two intracochlear stimulating electrodes (400 mm Pt/Ir balls) were inserted into the cochlea through the round window. One electrode was inserted approximately 8 mm and was

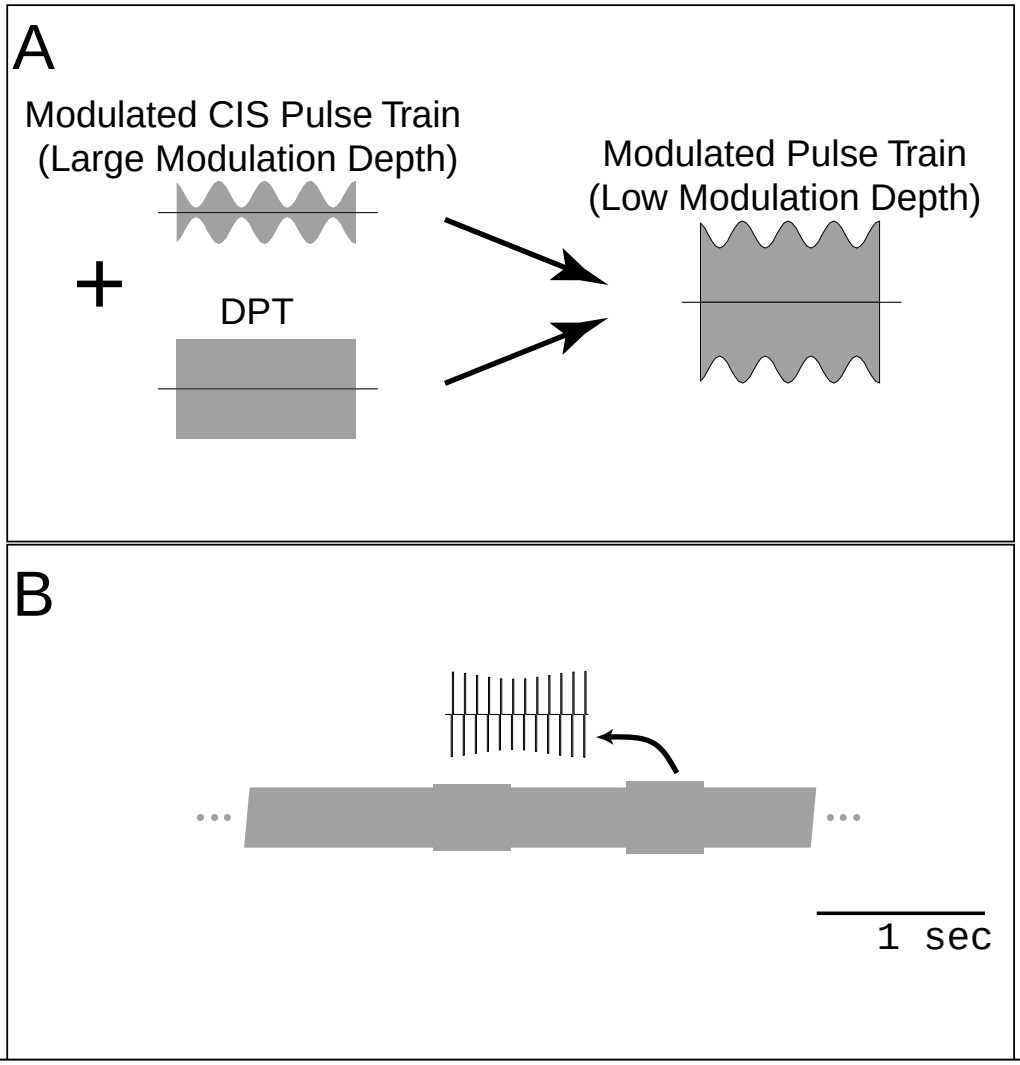


Figure 1. (A) The left panel shows two stimuli generated by a cochlear implant to implement a DPT protocol suggested by Rubinstein *et al.* (1999). This stimulus is composed of two pulse trains: one, strongly modulated, is the CIS signal, and the other, unmodulated, is the DPT. We assume that the DPT-enhanced stimulus can be modeled by a carrier of the same frequency as the DPT that is more weakly modulated than the original CIS stimulus (right). This approximation is exact if the CIS and the DPT have identical carriers. (B) Schematic representation of the electric stimuli used in this study. Shown is one cycle of the stimulus. The stimulus consisted of a sustained DPT, which was modulated every second. The modulation waveform and/or modulation depth were changed on each stimulus presentation. The entire cycle of modulation waveforms and modulation depths was repeated for up to 10 minutes.

used as the stimulating electrode. The second electrode was inserted just inside the round window and served as the return electrode. The opening in the round window was sealed with connective tissue.

Standard techniques were used to expose the auditory nerve via a dorsal approach (Kiang *et al.*, 1965). We recorded from single units in the auditory nerve using glass micropipettes filled with 3M KCl. For unmodulated pulse trains, a digital filter was used

to separate neural responses from stimulus artifact (voltage excursions recorded at the micropipette that are not spikes) in real time. During stimulus modulations, spike and artifact waveforms were stored for later analysis. To remove the artifact from these records, we used the method described in the companion paper (Litvak et al., 2002a).

A. Stimuli

We conducted a neural population study by investigating responses to a single DPT level for each animal. To maximize the yield in our experiments, the DPT level was chosen to be 8-10 dB above the evoked-response threshold. Methods for setting the DPT level are described in detail in a companion paper (Litvak et al., 2002c).

As in the earlier paper on sinusoidal modulations (Litvak et al., 2002a), small modulations of a sustained DPT (modulation depth ≤ 15%) were used to imitate the effect of presenting modulated pulse trains superimposed upon DPT stimulation (Fig. 1A). We assumed that neural responses to a high-rate pulse train with a low modulation depth are similar to those elicited by a stimulus that is a sum of a sustained DPT and a highly modulated pulse train. This assumption may hold if the membrane time constant is large compared to the intervals between pulses. It is exactly true if the same pulse train is used both as the DPT and as the modulation carrier. In this case, the modulation depth is linearly proportional to the amplitude of the modulated stimulus for small modulations.

The carrier was a 5 kpps pulse train composed of biphasic (cathodic/anodic) pulses (25 μs per phase). In order to acquire responses to both the unmodulated DPT, and to modulations of the DPT for several modulation depths and modulation waveforms, the stimulus was composed of alternating modulated (400 ms) and unmodulated (600 ms) segments (Fig. 1B). Modulation depth and/or modulation waveform was changed on each successive modulated segment. Modulation waveforms were derived from vowels and are discussed below. Modulation depth ranged from 0.5% to 15%. Modulation was applied such that the mean amplitude of the carrier pulses was kept approximately constant prior to and during modulation. The entire sequence of modulated and unmodulated segments was repeated continuously either for 10 minutes, or until contact with the fiber was lost. To expedite data collection, in the later experiments we shortened the duration of the unmodulated segments to 300 ms when the preceding modulated

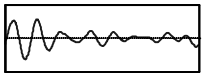
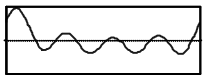
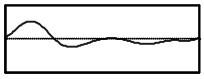
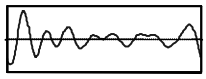
Sound	Band	Formants	Waveform
/ɔ/	2	F1=750 Hz	
/ɜ/	1	F1=450 Hz	
/u/	1	F1=250 Hz	
/ɯ/	2	F2=850 Hz	

Table 1. Stimuli used in this study. The fundamental frequency was 100 Hz. Because the formant frequencies were in all cases in between the harmonics of the fundamental, the peaks in the spectra of the vowels in the vicinity of the formants are less pronounced. In the time domain, this is evident by the fairly complex pattern observed for the first and the fourth vowel (right-most column).

segment had a modulation depth below 2.5%, and therefore produced low discharge rates.

The envelope for vowel-modulation stimuli was derived from synthetic vowels created by Assmann and Summerfield (1990). These vowels were synthesized by a Klatt series synthesizer (Klatt, 1980). The fundamental period of all vowels was 100 Hz. A single period of each vowel was extracted, and digitally resampled at 40 kHz. From each vowel, two modulation waveforms were produced by filtering the vowel period into two frequency bands (0 to 500 Hz and 500 to 1000 Hz). Although some vowels in Assman and Summerfield (1990) had second formants as high as 2.2 kHz, we did not analyze frequencies higher than 1000 Hz because those frequencies cannot be properly represented with a 5 kpps carrier pulse rate. Filtering was accomplished by computing the discrete Fourier transform of the vowel period, zeroing the harmonics outside of the pass band, and computing the inverse Fourier transform. The filtered vowel periods were concatenated to produce the 400 ms modulators. Finally, only band-vowel combinations that contained at least one vowel formant were investigated. Table 1 summarizes the vowel-band combinations used in this study.

Two different definitions of modulation depth were used to specify the amplitude of the modulated vowel stimuli. In early experiments, the modulation depth was chosen to reflect the peak modulation amplitude relative to the DPT. Specifically, if *e(t)* represents the modulation waveform, the envelope was defined as $A\cdot\left(1+\frac{m\cdot e(t)}{\max\left(\left|e(t)\right|\right)}\right)$. With this definition, the rms magnitude of the envelope can differ by as much as 1.5 dB across different filtered vowels. In later experiments, we normalized rms across vowel envelopes. Specifically, for the first vowel in the sequence, which was always the low-passed \ɔ\, the modulation depth was chosen to reflect the peak amplitude relative to that of the unmodulated DPT as before. The other vowels were then normalized such that, for a given modulation depth, the rms of each vowel modulation waveform matches that of the first vowel. Specifically, the modulation waveform was defined as $A\cdot\left(1+\frac{m\cdot e(t)\cdot K}{rms\left(e(t)\right)}\right)$, where *K* is the rms amplitude of the low-passed \ɔ\ divided by the peak amplitude for that vowel.

Analysis of responses to vowel modulators revealed no consistent dependence of discharge rate on the definition of modulation depth. For the purpose of this paper, therefore, we pool the data from the two definitions of modulation depth together when describing population responses.

B. Analysis

Responses collected during the unmodulated DPT segments were used to classify each fiber as in the companion paper. Some fibers exhibited only a transient response to the DPT, and adapted to a discharge rate near zero after a minute of DPT stimulation. We will refer to these fibers as "transient DPT responders." The fibers that responded throughout the unmodulated segments after the first minute of stimulation will be referred to as "sustained responders."

In a normally-functioning ear, the frequency components near the characteristic frequency are coded in the temporal discharge patterns of ANFs (Young and Sachs, 1979; Delgutte and Kiang, 1984). To determine whether the fine-time structure is also coded in responses to the modulated DPT, period histograms of the responses were computed

locked to the 10 ms period of the vowel. To eliminate the effect of locking of responses to the 5 kpps carrier, all histograms were computed using 0.2 ms bins. The amplitude and the phase of response to each harmonic of the vowel were estimated by computing the discrete Fourier transform of the period histogram. If the period histogram is scaled such that the magnitude at each bin represents the instantaneous discharge rate in spikes/s, then each Fourier component also has units of rate in spikes/s, and corresponds to the average discharge rate multiplied by the synchronization index to that component (Young and Sachs, 1979).

To formally characterize the transformation between the spectrum of the modulator and that of the period histogram, we estimated the neural modulation transfer function (MTF). The MTF expresses, as a function of frequency, the complex ratio (magnitude and phase) of the modulation in the neural response to the modulation in the stimulus. Traditionally, the MTF is estimated by recording responses to sinusoidal modulators. At frequency f_m, the MTF magnitude is twice the synchronization index of the response to the sinusoidal modulation of frequency f_m divided by the modulation depth of the stimulus, and the phase is the mean phase of the response minus the phase of the stimulus. Alternatively, if the system behaves linearly, the MTF can be estimated directly from the Fourier transform of the period histogram of responses to a complex stimulus (Moller, 1973; Moller and Rees, 1986) because the synchronization index for frequency f_m equals the ratio of the Fourier component at f_m to the Fourier component at DC (e.g. Young and Sachs (1979)). Estimates of the MTF from responses to complex modulation were computed, for each frequency by dividing twice the ratio of each component of the Fourier transform to the DC component by the same component of the Fourier transform of the modulator. In the case of sinusoidal modulators, this definition is identical to the traditional one. We used bootstrapping techniques to estimate the confidence interval for the MTF computed from response to each fiber (Efron and Tibshirani, 1993). To test the linearity assumption, we compared MTFs computed from responses of the same fiber to different vowel modulators. For a linear system, the MTF should be independent of the stimulus used to estimate it. We also compared the average MTF estimate derived in this paper from vowel responses to the estimate at three frequencies based on responses to sinusoidal modulations recorded in the previous paper (Litvak et al., 2002a).

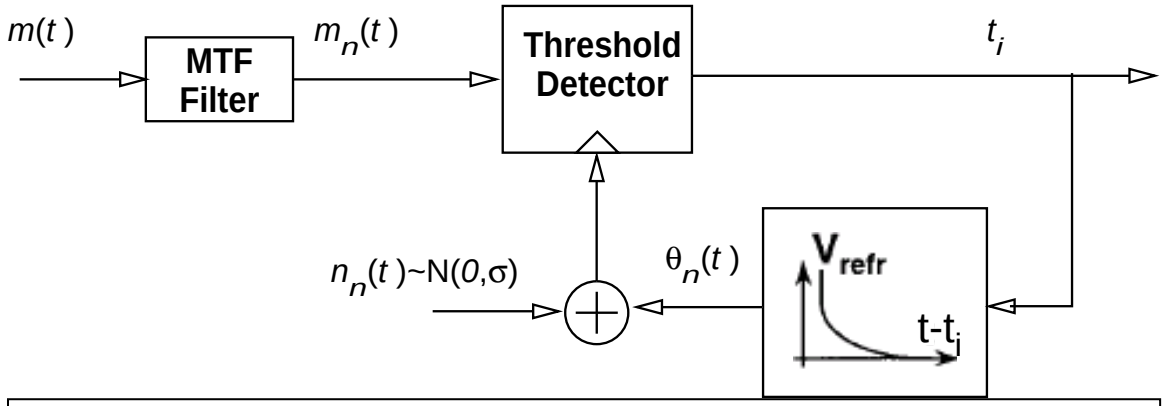


Figure 2. Functional model used to predict ANF responses to arbitrary modulation waveforms *m*(*t*) applied to the DPT. Model can predict response to. The model outputs the spike times $\{t_i\}$.

C. Stochastic Threshold Model

In the previous report, we introduced a simple functional stochastic threshold model (STM) of ANF responses to modulations of the DPT (Litvak et al., 2002a). The model takes as input the modulator waveform *m*(*t*), which is compared to a threshold. A spike is produced by the model whenever *m*(*t*) crosses threshold. The threshold is the sum of the deterministic threshold, which includes both an absolute and a refractory period, and the noise term. We demonstrated that this model was successful in predicting the dependence between threshold and dynamic range of responses to sinusoidal stimuli and discharge rate during the unmodulated DPT. However, the model failed to predict the dependence of modulation detection threshold of single ANFs on modulation frequency. To account for this dependence, we introduce a linear neural MTF into the model (Fig. 2). To distinguish the MTF of the STM model from the empirical MTF computed from responses, we will refer to the former as the *underlying* MTF. We will show in this paper that the underlying MTF can be estimated from ANF responses to complex modulators.

We computed responses of the STM to vowel modulators of different frequencies, and compared its output to ANF responses to a modulated DPT. The model was simulated using 0.2 ms time steps.

III. Results

We recorded responses of 36 units in 4 cats to modulated vowels. Each record was composed of responses to 4 filtered vowels at 4 or 5 modulation depths ranging from 1% to 15%.

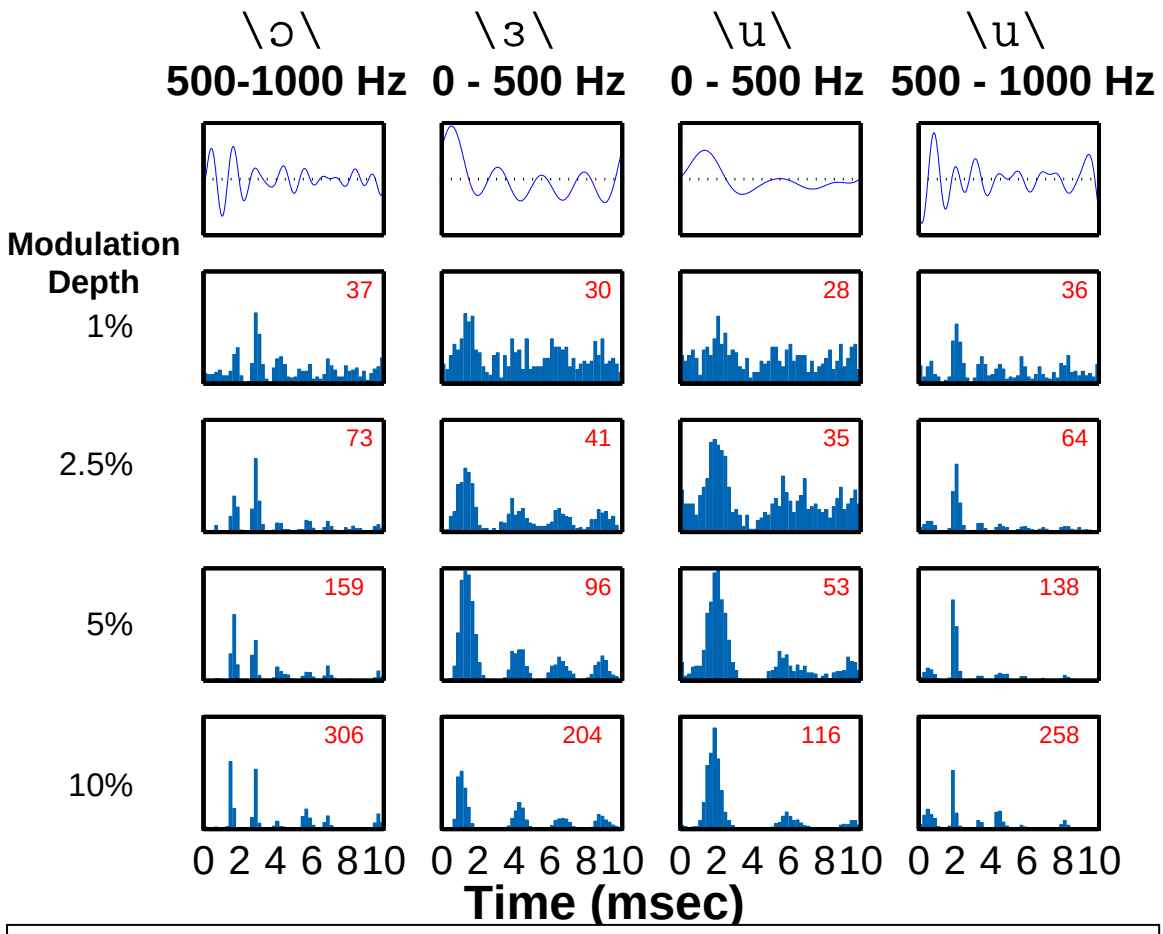


Figure 3. Plots of period histograms of responses to vowel modulators for ANF lml21-53 (discharge rate during unmodulated DPT: 28 spikes/s). *Top Row*: the modulation waveforms, with dashed lines showing each waveform's mean. *Other Rows:* period histograms of responses of the ANF to each modulator for modulation depths 1%, 2.5%, 5% and 10%. Period histograms were computed using 0.2 ms bins. The numbers in each panel indicate the average discharge rate during the corresponding modulated segment in spikes/s.

A. Temporal discharge patterns

Fig. 3 shows the period histograms computed from the responses of one fiber to the filtered-vowel modulators at different modulation depths. During the unmodulated stimulus segments, this fiber's discharge rate was 28 spikes/s. For all modulation depths, the period histograms broadly resembled the modulation waveforms. In particular, modes in the period histograms match peaks of the modulation waveform. The intervals between peaks in the modulation waveforms are related to the vowel formant frequencies. This result suggests that with a DPT, the responses to modulated vowels code the vowel formant frequencies.

Although the peaks of the modulation waveform are represented for all modulation depths, the best resemblance between the modulation waveform and the period

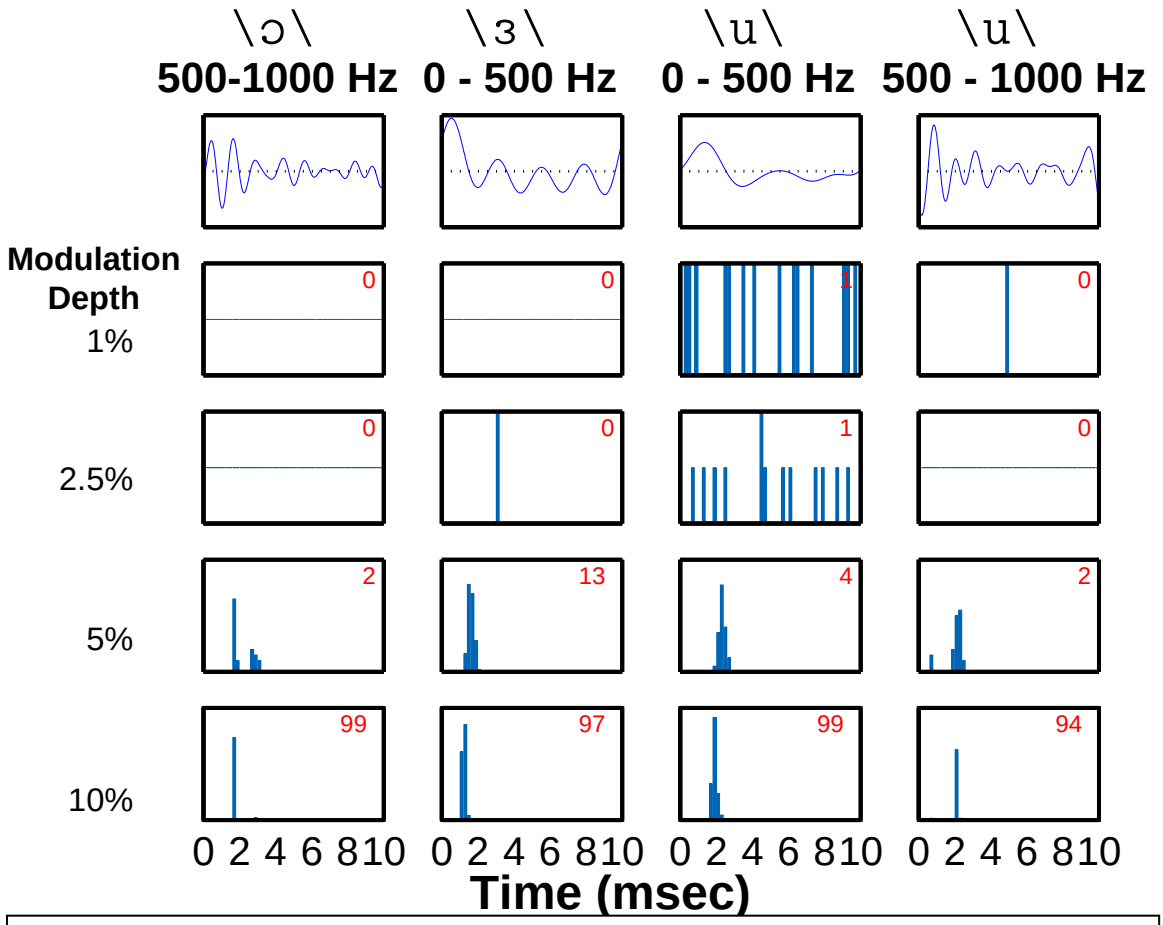


Figure 4. Plots of period histograms of ANF responses to vowel modulators for fiber lml20-8 (transient DPT responder). The format of the figure is the same as for Fig. 3. This fiber was less sensitive to modulation than the fiber in Fig. 3. While the ANF in Fig. 3 represented the entire modulation waveform in its responses, this fiber primarily responded to the peak of the modulation.

histograms occurs for small modulation depths. For example, the distance between the two largest modes in the period histogram computed from the response to the modulator derived by filtering the \ɔ\ modulator is 1.2 ms for modulation depths between 1% and 5%, but is 1.4 ms for the modulation depth of 10%. The time between the two largest peaks in the stimulus waveform is 1.2 ms. Similar distortions can be shown in responses to the vowel \u\. In addition, some of the smaller peaks in the modulation waveform are only represented at the smaller modulation depths.

In contrast to the sustained DPT responders, the transient responders poorly represented complex modulations. Fig. 4 shows the response of a fiber that responded transiently to the unmodulated DPT. For modulation depths below 5%, the fiber did not

respond significantly to the vowel modulation. This finding is consistent with the lower sensitivity of transient DPT responders to sinusoidal modulations. At modulation depths of 10% and 15%, the fiber responded strongly to the modulation. The response was highly synchronized, and occurred almost exclusively to the largest peak in the modulation waveform. One exception is the response to the second largest peak in the vowel \ɔ\ (filtered from 500 to 1000 Hz) at 5% modulation. These responses resemble responses to electric vowels without the DPT (Knauth et al., 1994).

To quantify the resemblance between the modulation waveform and the period histogram of the ANF responses, we computed the correlation coefficient between the two. To account for possible neural delay, the modulation waveform was shifted successively in steps of 0.2 ms (up to 1 ms), until the largest correlation between the modulation waveform and the response was achieved.

Fig. 5 plots the maximum correlation against the average discharge rate during the unmodulated DPT for modulation depths of 2.5% (top) and 10% (bottom) for responses of 20 fibers in 3 animals. Correlation was only plotted if the bootstrapped 90% confidence intervals for the correlation coefficient were below 0.1, and if the period histogram contained at least 10 spikes. At 2.5% modulation, all responses from sustained responders (rate during unmodulated DPT above 1 spikes/s) except one had correlations above 0.65, with an average correlation near 0.75, implying excellent representation of the vowel fine-time structure. In contrast, most transient DPT responders had lower correlations.

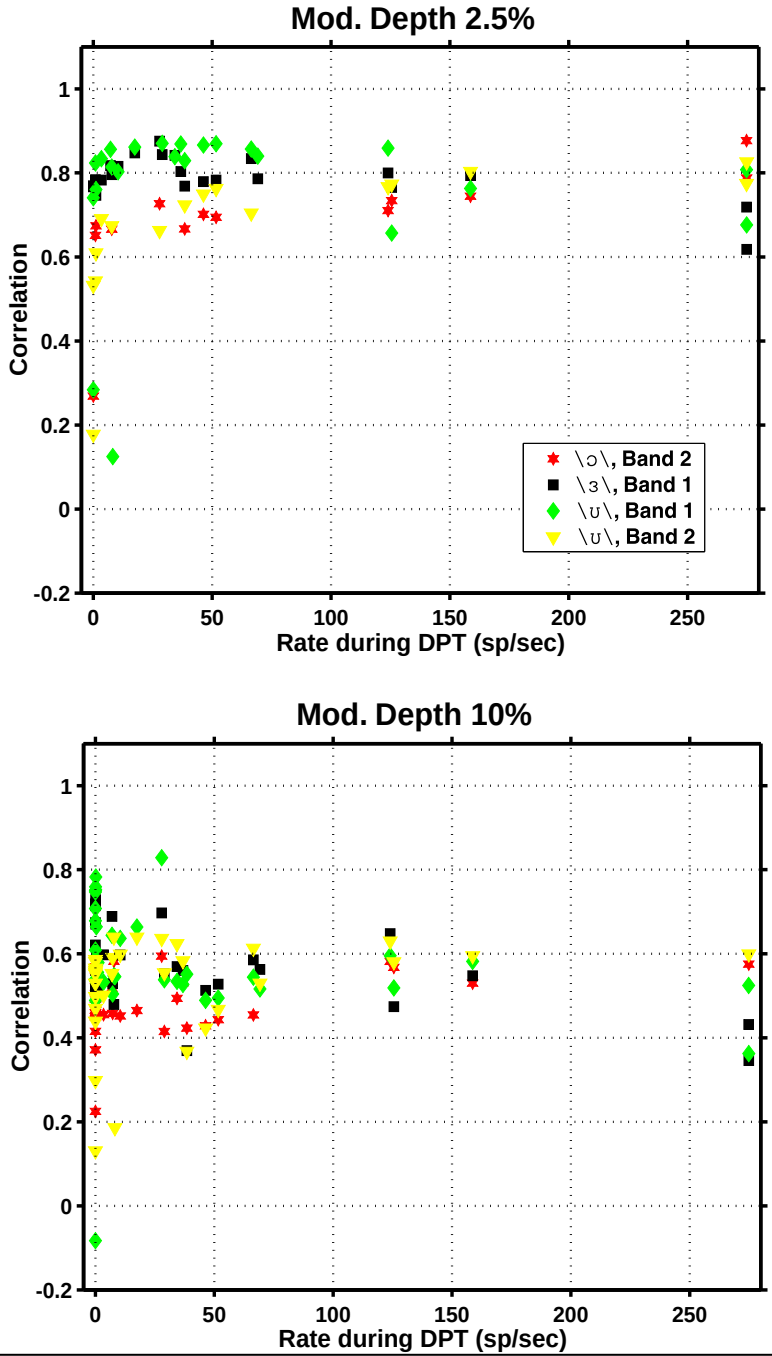


Figure 5. Maximum correlation coefficient between the period histogram and the modulation waveform as a function of average discharge rate during the unmodulated DPT. Correlation was computed separately for each vowel modulator (symbols). The top panel shows the correlation for a modulation depth of 2.5%. Sustained DPT responders (discharge rate during the unmodulated DPT above 1 spikes/s) have high (on average near 0.8) correlations. The correlation is reduced to an average of 0.5 for the modulation depth of 10% (bottom panel).

At 10% modulation depth, the average correlation of the sustained DPT responders is lower (near 0.5). Some transient DPT responders also have much lower correlations. Thus, representation of the vowel waveform in the ANF temporal discharge patterns is

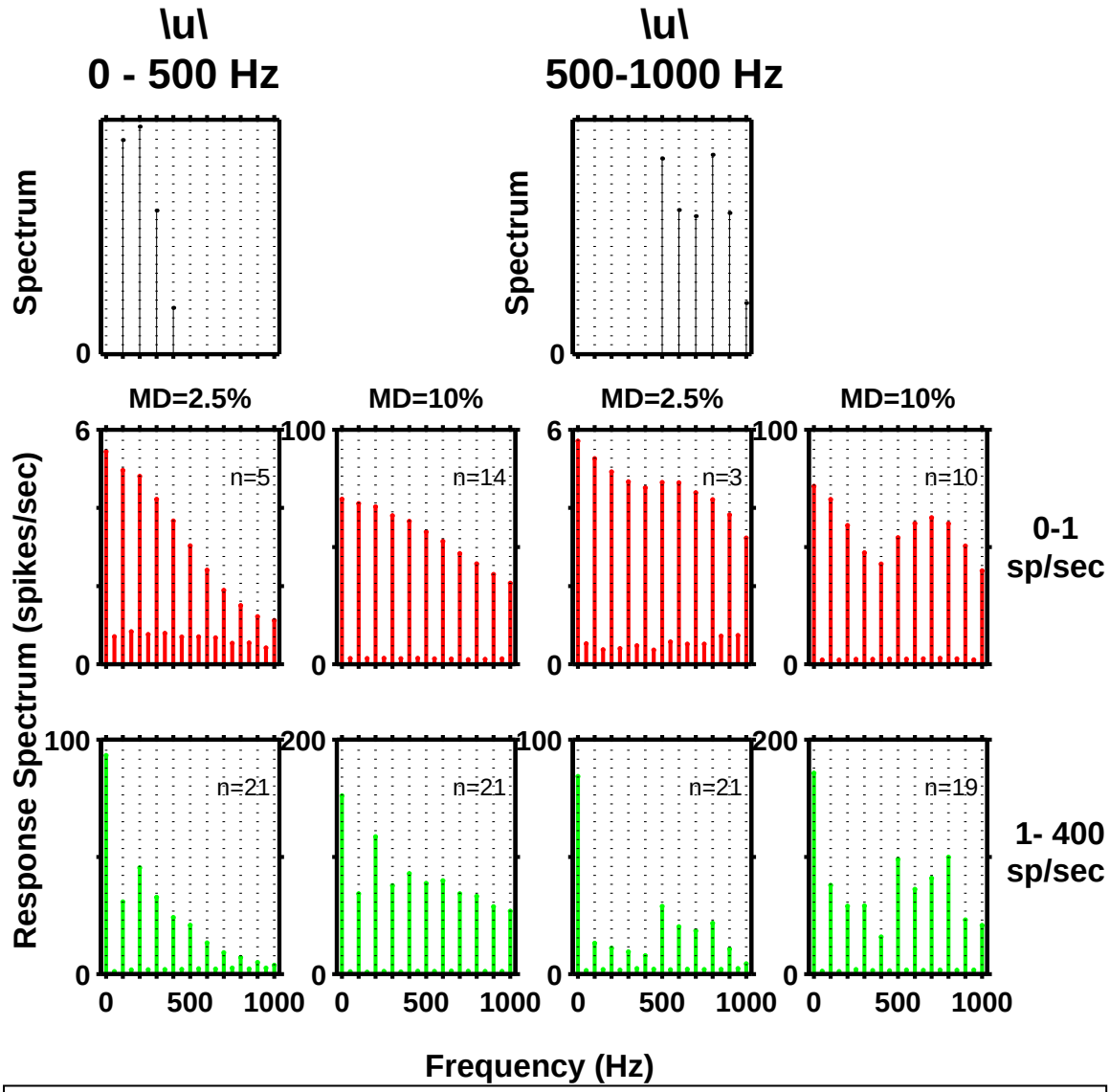


Figure 6. While the fundamental is represented in responses of sustained as well as transient DPT responders, only the former reliably encode formant frequency in their temporal discharge patterns. Each panel in the top row shows the spectrum of the modulator waveforms. Peaks in these spectra are at harmonics of the 100 Hz fundamental that are closest to the formant frequencies. The bottom two rows show the average spectrum of period histograms of responses to the modulators (modulation depths of 2.5% (left) and 10% (right)). The units on the ordinate are spikes/s. The average spectrum is computed separately for transient DPT responders (discharge rate below 1 spike/s) (middle row), and sustained responders (bottom row).

better at the 2.5% modulation depth than at the 10% modulation depth. Representation is also better in the sustained responders than in transient responders.

B. Representation of formant frequencies

The high correlations of the responses to vowel-like modulators do not necessarily mean that the spectral peaks of the modulation waveform are well represented. In the case of vowels, the peaks in the spectrum correspond to formant frequencies, which are

important cues for vowel identification (Peterson and Barney, 1952). To determine whether the spectral peaks of the modulation waveform are accurately represented in the ANF temporal patterns, we computed the discrete Fourier transform of the period histogram (Young and Sachs, 1979). Transforms were computed from histograms containing two vowel cycles, to measure the noise floor between the harmonics.

Magnitude spectra of individual fiber responses depended strongly on average discharge rate during the DPT. While there were notable differences in the shape of the spectrum between transient and sustained DPT responders, within each group the spectra were more similar. Fig. 6 plots the average magnitude spectra for the vowel \u\ (low-passed in the left column and high-passed on the right column) and 2 modulation depths. Separate averages were computed for transient (<1 spike/s top row) and sustained (≥1 spikes/s, bottom row) DPT responders.

For both modulation depths, the average spectrum for transient DPT responders poorly represented the spectrum of the low-pass filtered \u\ (top row, left). However, the presence of the peaks at the harmonics of the fundamental indicates that the response is locked to the fundamental frequency. The largest spectral component (besides DC) is the 100 Hz fundamental. Although the average magnitude spectrum for transient DPT responders shows a local maximum near the second formant of the high-passed \u\, the largest non-DC component of the average spectrum is still the fundamental. Thus, transient DPT responders represent primarily the fundamental in their temporal responses.

In contrast, the sustained DPT responders had peaks in the response spectrum that qualitatively matched those in the vowel spectrum. However, at both modulation depths, magnitude spectra showed significant components at frequencies for which there was no energy in the stimulus modulator spectrum (top panel), indicating that the transformation between the period histogram and the response is in general nonlinear. Part of the non-linearity may be related to rectification, which is apparent in period histograms shown in Fig. 3.

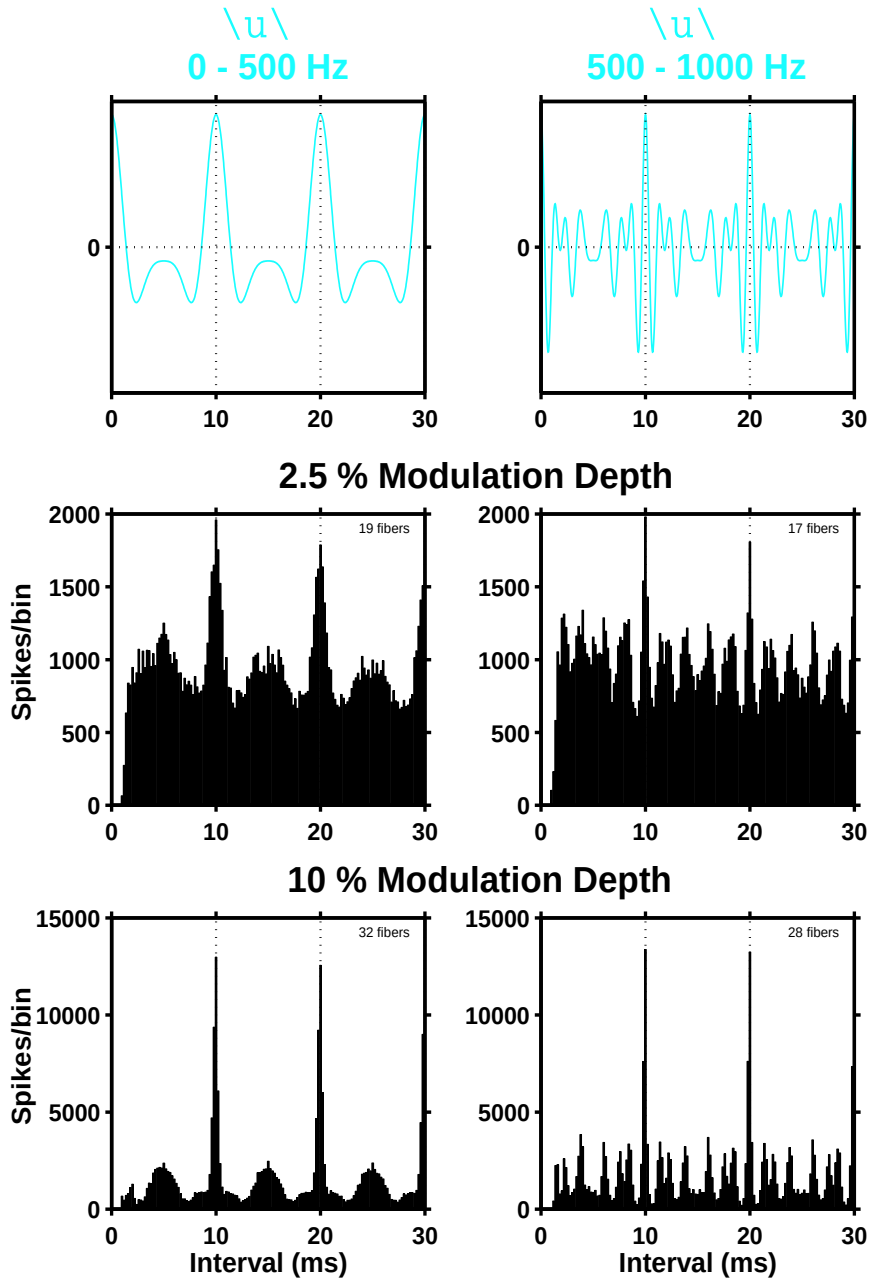


Figure 7. The pooled autocorrelation histogram robustly encodes the fundamental, and resembles the autocorrelation of the stimulus (except for the shortest intervals). Each panel in the top row plots the autocorrelation for two modulation waveforms used in the present study. Dashed lines are plotted at multiples of the vowel period. Panels in the next two rows plot pooled autocorrelation histograms of the electric responses for modulation depths of 2.5% (middle row) and 10% (bottom row). All pooled autocorrelation histograms have a large peak at the vowel fundamental. Smaller peaks in the pooled autocorrelation histograms correspond to smaller peaks in the autocorrelation of the stimulus.

C. Representation of pitch

Because the period histogram analysis assumes that the period is known, it cannot be used to directly test for coding of the vowel fundamental in ANF temporal discharge

patterns. Arbitrary response periodicities can be revealed using all-order interval (a.k.a. autocorrelation) histograms (Moller, 1970; Cariani and Delgutte, 1996b).

We pooled the responses of all fibers by computing the autocorrelation histogram of individual fibers, and then summing their autocorrelation histograms. Fig. 7 plots the pooled autocorrelation of the electric responses at modulation depths of 2.5% (middle row) and 10% (bottom row) for two modulation waveforms. Both waveforms were derived by filtering the vowel \u\. The responses to other vowels show the same general trends and are not shown.

All pooled histograms have a pronounced mode at the vowel period of 10 ms, suggesting that the fundamental is robustly coded in the temporal patterns of ANF responses for both vowels. Nevertheless, the fine-time structure of the two histograms is very different. The smaller modes in the histogram match the peaks in the autocorrelation of the modulator (top row). Because peaks in the autocorrelation of the modulator are related to formants, the pooled autocorrelation histogram codes the formant frequencies in its fine-time structure. A similar result has been reported for pooled autocorrelation histograms computed from ANF responses to acoustic vowels (Cariani and Delgutte, 1996a; Cariani and Delgutte, 1996b).

D. Linearity of responses

The modulation transfer function (MTF) is defined as the ratio of modulation in the neural response to the modulation of the stimulus. Traditionally, MTFs are measured using sinusoidal modulators (e.g. (Delgutte et al., 1999)). However, for a system that is nearly linear, identical estimate the MTF can in principle be obtained using modulators that simultaneously contain several harmonics by dividing the spectrum of the response period histogram by the spectrum of the stimulus. We tested the linearity assumption by computing MTFs from responses to different vowels. In a linear system, the MTF estimate should be independent of the modulator waveform used to estimate it.

Fig. 8 shows the MTFs estimated in this way from responses of individual fibers for discharge rates to the DPT above (right) and below (left) 130 spikes/s respectively. Estimates based on responses to different vowels are denoted by different symbols. For each vowel response, MTFs can only be estimated for frequencies at which there is

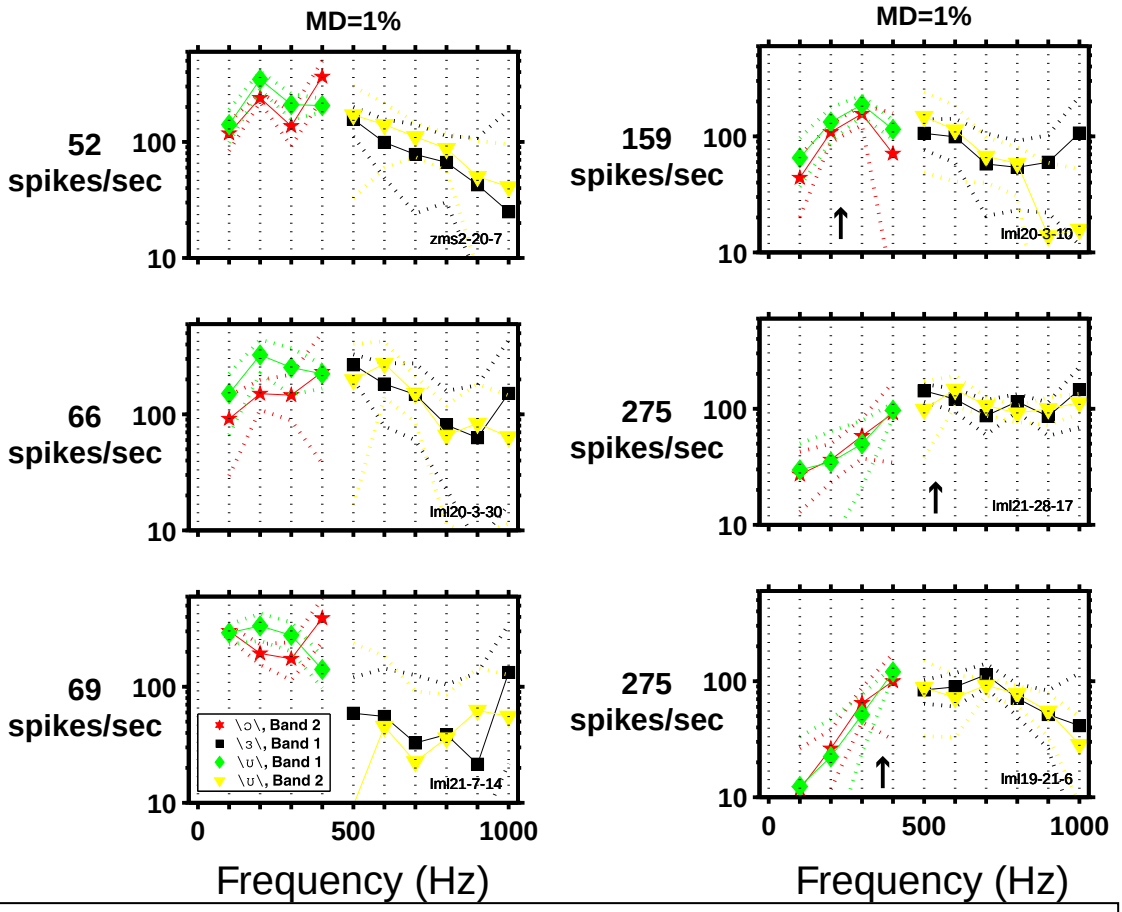


Figure 8. For fibers with average discharge rates during the unmodulated DPT above 130 spikes/s, the MTF is independent of the vowel used to estimate it. Each panel plots the magnitudes of the MTFs estimated from responses of one fiber. Different symbols denote different vowels. Note that the MTF could only be estimated at frequencies for which there was energy in the modulator. Discharge rate of each fiber during the unmodulated DPT is indicated on the left. The modulator was presented at 1% modulation depth. Dashed lines represent the 99% confidence intervals and were computed by bootstrapping ANF responses. For the three fibers on the right, the inherent periodicity frequencies (defined as the inverse of the largest mode in the interval histogram of responses to the unmodulated DPT) are (from top to bottom) 234, 537, and 383 Hz (arrows).

spectral energy in the modulator. The dashed lines of corresponding color represent the 99% confidence intervals in estimating the MTF.

For 1% modulation depth, the modulation gain of all fibers is strikingly large, exceeding 100 for frequency of 400 Hz. MTFs computed from responses of fibers with discharge rates above 130 spikes/s in response to the DPT (Fig. 8, right) are independent of the vowel used to compute the MTF. This can be seen in Fig. 8 from the overlapping confidence intervals computed for the different vowels. The independence of MTF from the vowel used in computing its estimate is consistent with the linearity assumption. However, there are statistically significant differences in MTFs across fibers.

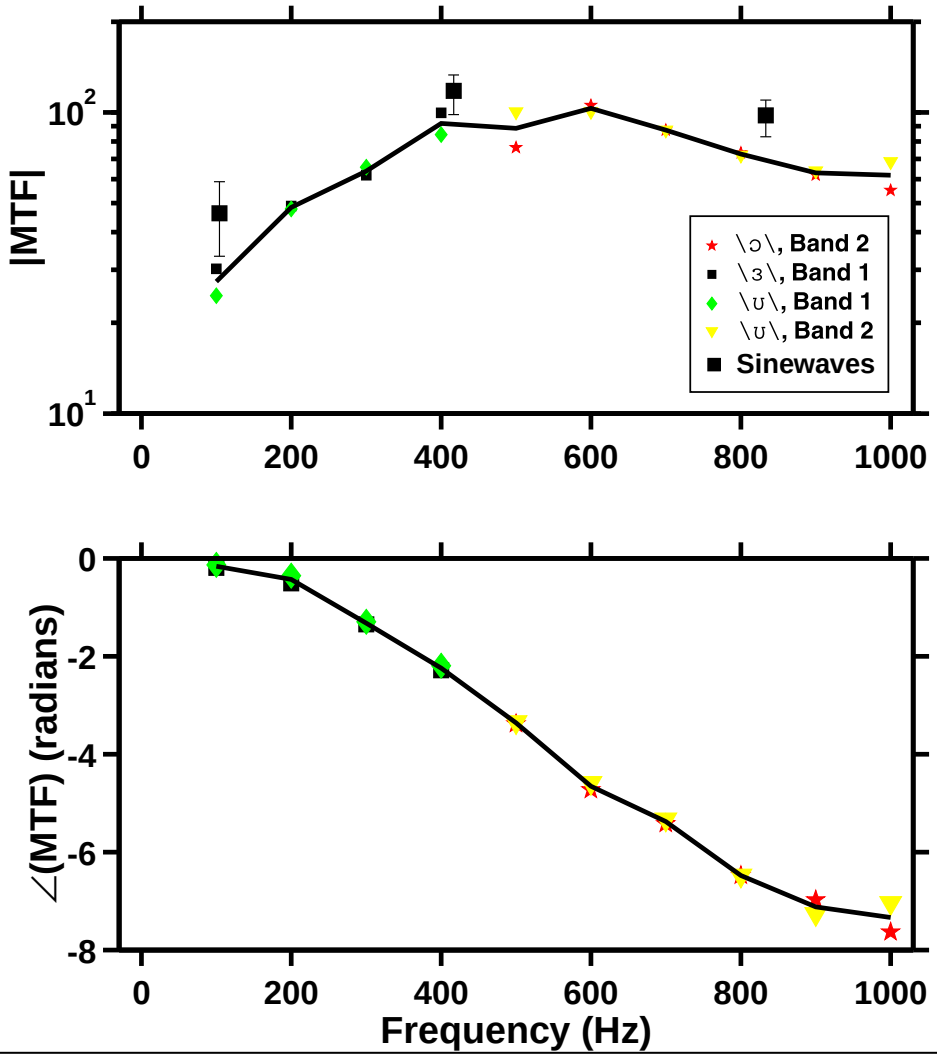


Figure 9. The average MTF computed from responses of fibers with average discharge rates during the unmodulated DPT above 130 spikes/s. The overall spectral shape resembles a bandpass filter and is consistent with estimates of average MTFs based on responses to sinusoidal modulators. *Top*: magnitude of the estimated MTF. The black line indicates the average, which was computed by averaging individual MTF responses for three fibers to the four modulators. The symbols indicate the average MTF based on responses of the three fibers to individual modulators. The black squares indicate average, and the error bars 99% confidence intervals of the MTF estimated from responses to sinusoidal modulators (Litvak et al., 2002b). The MTF was estimated using the formula $MTF(f)=2SI(f)/m$, where f is modulation frequency, SI is the synchronization index, and m is modulation depth. As with the vowels, the MTF was estimated based on responses to modulators with 1% modulation depth. *Bottom*: Estimated MTF phase. The phase is nearly linear between 200 and 800 Hz, but deviates from linear at both the lower and higher frequencies.

For fibers with the lower discharge rates in response to the DPT (Fig. 8, left), MTF estimates are not independent of the vowel used to compute it. Differences between MTF estimates were particularly large for estimates based on low-passed vowels \u\ and \ɜ\. These responses are not consistent with linearity. The most noticeable non-linearity is rectification, which is apparent in period histograms of these responses (e.g. Fig. 3).

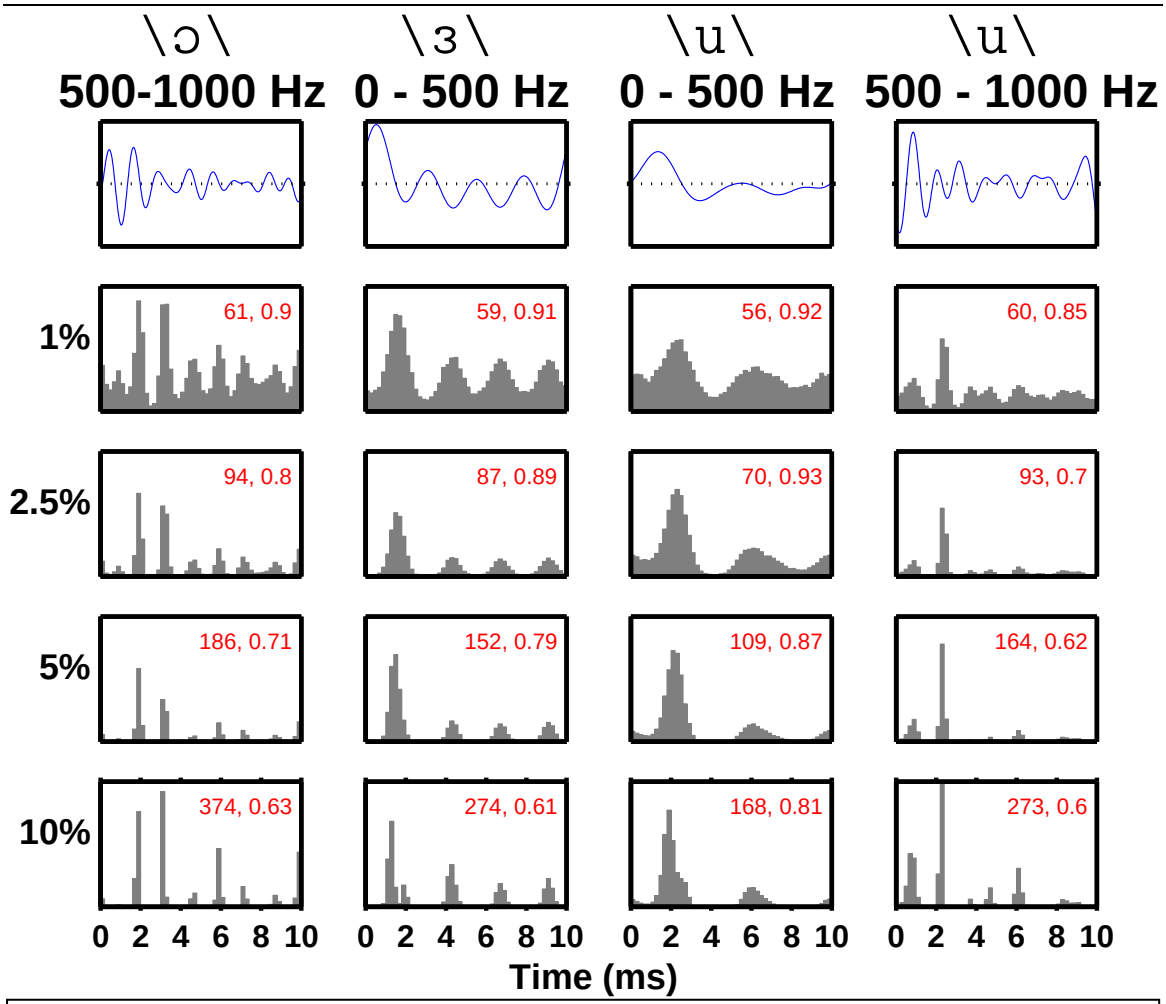


Figure 10. The period histograms of STM responses for a fiber with moderate spontaneous rate were similar to responses to modulated DPTs for at least some sustained DPT responders (e.g. Fig. 3). The STM fiber had a spontaneous rate of 50 spikes/s (θ/σ=2.222). Format of the plot is the same as in Fig. 3. The second number in each panel represents the similarity of the response to the vowel modulator waveform defined as in Fig. 5, while the first number is the average discharge rate during the vowel modulator.

Because any model that produces spikes inherently includes rectification, the dependence of MTF shape on stimulus spectrum does not necessarily suggest that a model with a linear MTF cannot predict the responses of these fibers. The data do suggest, however, that responses to complex modulations cannot be used directly in these fibers to estimate the underlying MTF.

Fig. 9 plots the magnitude and the phase of the average MTF computed from the responses of fibers with rates above 130 spikes/s during the unmodulated DPT (solid line). Although responses to different vowels were used to compute the average, the results would not be very different if only a single vowel (one set of symbols) had been used to estimate each half of the spectrum. The large black squares indicate the MTF

estimated from responses to sinusoidal modulators reported in the previous paper (Litvak et al., 2002a). Consistent with our previous report, the MTF is bandpass, with a center frequency near 400-600 Hz. Although not apparent in Fig. 9, where frequency is on a linear scale, the drop-off in the MTF around the pass band is approximately 6 dB/octave.

E. Predictions of the stochastic threshold model

In a previous report (Litvak et al., 2002a), we proposed a stochastic threshold model (STM) to predict responses of ANFs to modulations of the DPT. We showed that the STM was successful in predicting many aspects of responses to sinusoidal modulations. In this paper, we examined whether the STM can also predict responses to complex modulators by comparing measured and model responses to vowel modulators.

The parameters of the STM are the threshold θ_0, the noise variance σ, and the underlying MTF. The ratio θ_0/σ specifies the spontaneous discharge rate. To adapt the STM to arbitrary waveforms, both the underlying MTF filter and the σ must be specified. We chose the empirical filter shown in Fig. 9 as the underlying MTF. To match the measured MTF of the model fiber (spontaneous rate of 300 spikes/s) at modulation frequency of 400 Hz with the MTF in Fig. 9, we set σ to 1.7.

Fig. 10 shows the period histograms of responses of a model fiber with a spontaneous rate of 50 spikes/s to vowel modulators for modulation depths ranging from 1% to 10%. The predicted period histograms have several features in common with measured histograms for responses of sustained DPT responders (e.g. Fig. 3). First, both encode the modes of the modulator waveform in the peaks of the period histogram. The location and the number of major peaks is similar in the measured and model period histograms. Moreover, the correlation between the vowel modulator waveform and the model period histogram (second number in each histogram) is in the range observed for ANF sustained responses (Fig. 5). For both model fiber and ANF responses, the greatest correlation and the smallest discharge rate are achieved for low-pass \u\. However, at a modulation depth of 10%, similarity to the vowel modulator was somewhat less for the data than for the model.

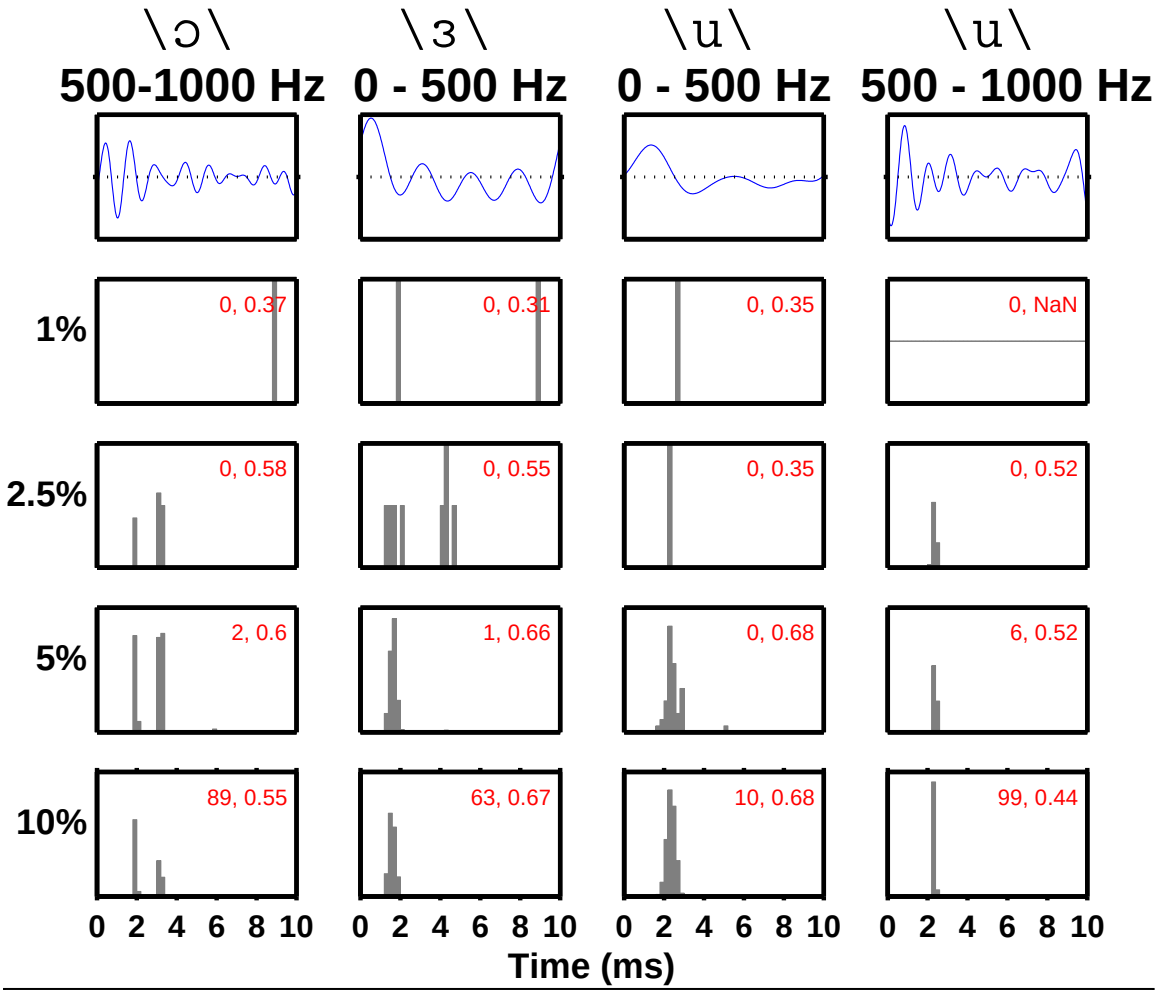


Figure 11. The period histograms of STM responses of a fiber with low spontaneous rate was similar to responses of at least some transient DPT responders to a modulated DPT (e.g. Fig. 4). The spontaneous discharge rate of the STM fiber was 0.0014 spikes/s (θ/σ=5).

Fig. 11 shows the responses of a model fiber with spontaneous rate near 0 spikes/s. The responses of this model fiber to vowel modulations resembled those of transient DPT responders (e.g. Fig. 4). Both the model fiber and transient DPT responders required large modulation depths to evoke a response. For both the model and neural data, responses to most vowels tended to occur only at the peak of the vowel modulator waveforms. These responses, therefore, were most accurate in representing the fundamental and poorly represented the formant frequencies. However, both the model fiber and the transient DPT responder showed responses representing the two largest peaks of the high-passed \ɔ\ modulator.

To quantitatively test whether the STM can predict the shape of the period histogram of ANF responses to vowel modulators, we computed the correlation coefficient between

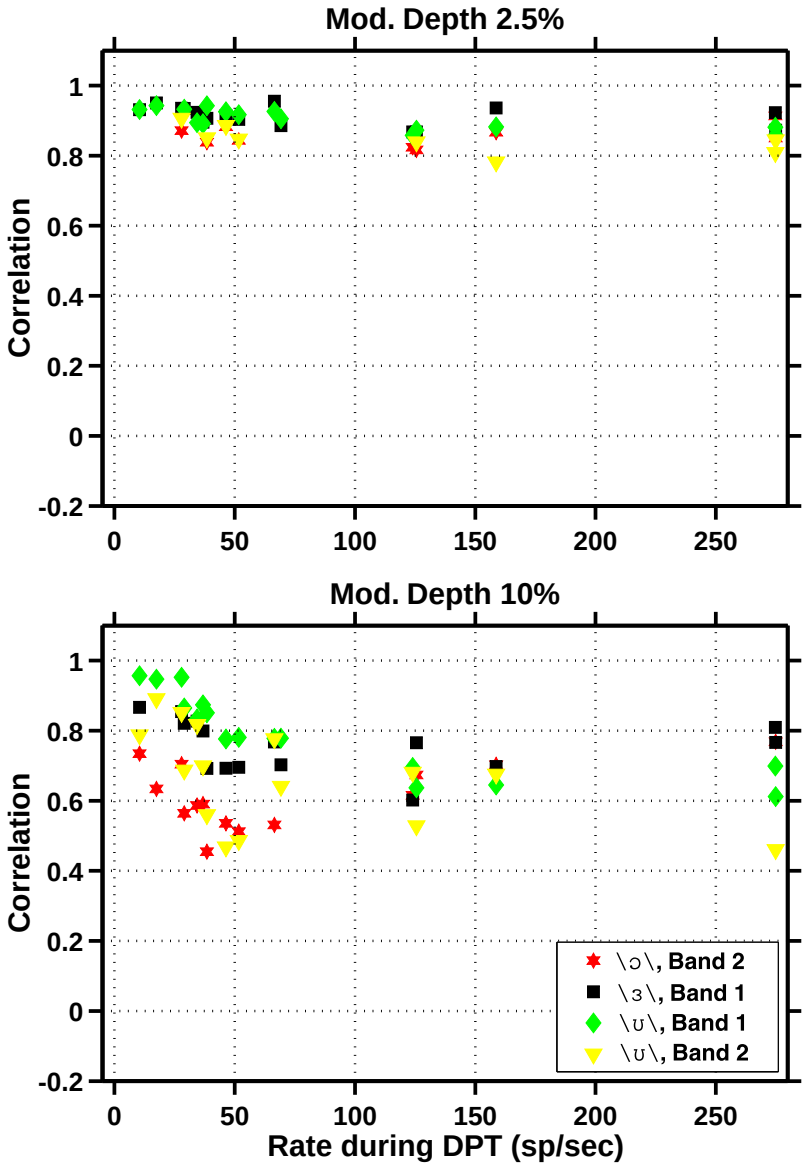


Figure 12. For low modulation depths, STM period histograms are highly correlated with period histograms of responses of ANFs to vowel modulators. *Top*: Correlation of the period histogram of STM responses to vowel modulators (modulation depth of 2.5%) with those of ANF responses as a function of the discharge rate during the unmodulated DPT. Different symbols represent responses to different vowel modulators. To account for differences in neural delay across fibers, we allowed τ to vary between -0.4 ms and 0.4 ms. Due to the difficulty of estimating low discharge rates accurately, the correlation was computed only for fibers with discharge rates during the unmodulated DPT above 10 spikes/s. *Bottom*: Correlation between ANF and STM responses to vowel modulators at a modulation depth of 10%.

the period histogram of the neural response and the period histogram predicted by the STM. The STM was fit to each fiber by adjusting the spontaneous rate (through adjusting θ) to match the average discharge rate of the fiber during the unmodulated DPT. Fig. 12 plots the correlation between the measured and model period histograms for modulation

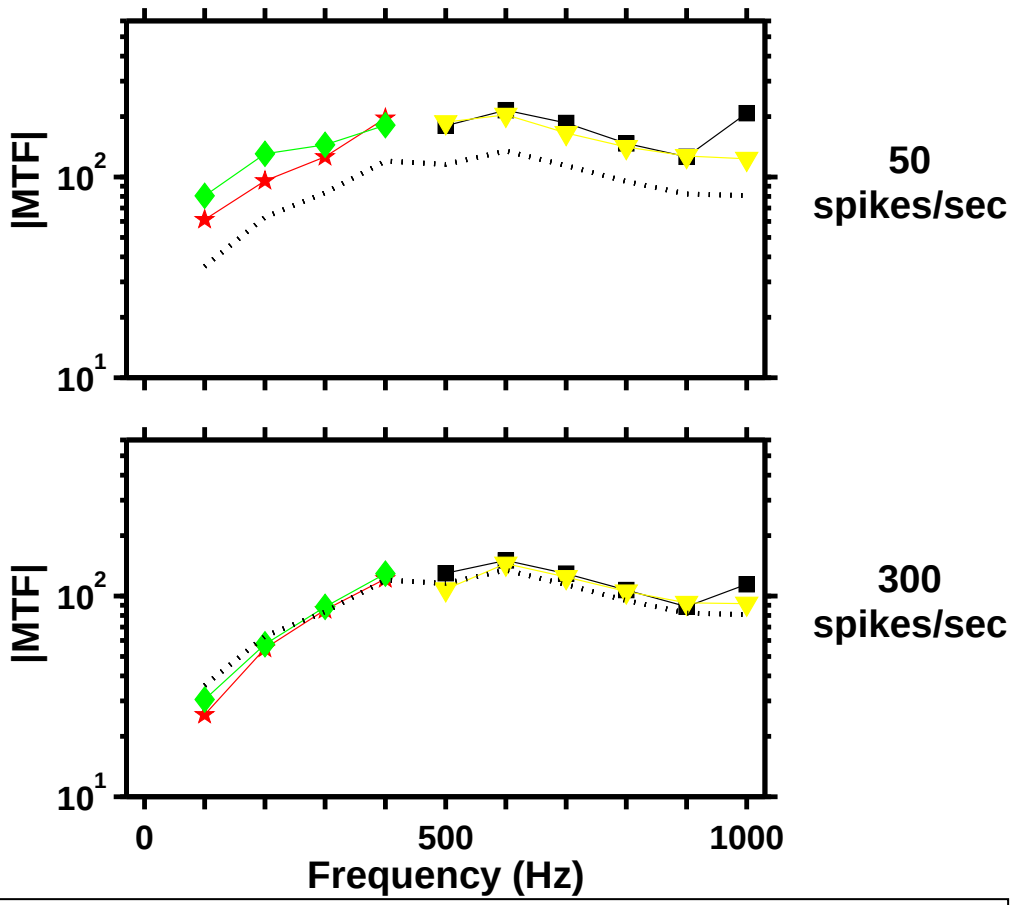


Figure 13. MTFs computed from responses of the STM fibers confirm that responses of fibers with high spontaneous rates to complex modulators (1% modulation depth) can be used to directly estimate the "underlying" MTF of the STM. *Top Row*: Modulator spectrum for each vowel. *Middle Row*: Magnitude of the estimated MTF for STM fiber with spontaneous rate of 50 spikes/s (θ/σ=2.222). The shape of the estimated MTF depends on the modulator used to estimate it. Also, the MTF shape differs from the underlying model MTF (dashed line). *Bottom Row*: Estimated MTF computed from responses of the STM fiber with spontaneous rate of 300 spikes/s (θ/σ=1.1389). The shape of the estimated MTF is nearly independent of the modulator used to derive the estimate, and resembles the underlying model MTF.

depths of 2.5% (top) and 10% (bottom) versus the discharge rate of the unmodulated DPT. At both modulation depths, the correlation between the two is high, significantly better than correlation of the period histogram of ANF responses with the modulator waveform (Fig. 5). For the lowest modulation depth, the correlation is on average 0.89. Although the underlying MTF was estimated from responses to fibers with discharge rate during the unmodulated DPT above 130 spikes/s, the correlation at the lower modulation depth is nearly independent of discharge rate during the unmodulated DPT over the range of discharge rates from 10 to 300 spikes/s. The correlations are somewhat poorer for

higher modulation depth (average of 0.71). Thus, the STM is best at predicting responses at low modulation depths.

Fig. 13 shows the empirical MTFs computed from responses of the model to 1% modulation for two STM fibers that differ only in spontaneous rate. The dashed line in each panel represents the underlying MTF filter of the model. Both responses were computed at a 1% modulation depth. While the empirical MTF computed from the model fiber with spontaneous rate of 50 spikes/s differed in shape from the underlying MTF, the match was very close for the 300 spikes/s model fiber, providing further evidence that responses of fibers with high spontaneous rates can be used to estimate the underlying MTF filter directly from their responses.

IV. Discussion

A. Representation of vowel modulators in ANF responses

In response to filtered vowels, auditory-nerve fibers that had a sustained response to the unmodulated DPT had complex period-histogram shapes whose peaks corresponded to peaks in the modulation waveform. Fourier and autocorrelation analyses revealed that sustained DPT responders represented both the fundamental and the formants in their temporal discharge patterns. Because peaks in the modulation waveform differ greatly in amplitude, only neurons with a large dynamic range would be able to simultaneously encode several of these peaks in their temporal discharge patterns. The ability of sustained DPT responders to code the entire modulation waveform is therefore consistent with the large (15-25 dB) dynamic range of these neurons demonstrated in an earlier paper using sinusoidal modulators (Litvak et al., 2002a).

The closest resemblance between the period histograms and vowel waveforms occurred for modulation depths below 5%. At larger modulation depths, the period histogram showed distortions in the representation of the high frequencies. In addition, for modulation depths above 5%, the average discharge rates exceeded those seen acoustically.

As with sinusoidal modulations, the representation of the vowel is more accurate in sustained DPT responders than in transient DPT responders. A typical response pattern for a transient DPT responder is a single spike per stimulus period occurring at the

modulator peak. For the modulation depths studied, transient DPT responders are poor at representing the formant frequency. They do, however, accurately represent the fundamental. Responses of transient DPT responders are similar to responses to electric vowels without a DPT (Knauth et al., 1994; van den Honert and Stypulkowski, 1987b).

Reproducing the exact spatio-temporal pattern of ANF responses in a healthy ear in response to complex modulation would require an ability to stimulate narrower portions of the cochlea than is possible with modern cochlear implants. However, because the pooled autocorrelation histogram of responses to complex modulations resembles the autocorrelation of the stimulus, reproducing the pooled autocorrelation histogram of acoustic responses in each frequency channel may be achievable through manipulation of the modulation waveform. It remains to be seen whether implanted subjects can make use of that information.

B. Modulation transfer function

We hypothesized that responses to low modulation depths can be characterized by a linear transformation between modulation waveform and the period histogram. We tested the linearity assumption by estimating MTF for individual fibers using responses to different vowels. Under linearity, the estimated MTF should be independent of the vowel used to derive the estimate. We found that for a modulation depth of 1%, linearity held only for fibers with discharge rates during the unmodulated DPT above 130 spikes/s. MTFs estimated from these fibers differed somewhat from one fiber to the next. We have previously reported that fibers with discharge rates during the unmodulated DPTs above 100 spikes/s have inherent periodicities in their responses, and that these periodicities interact with periodicities in the modulation (Litvak et al., 2002c; Litvak et al., 2002a). One possibility is that inherent periodicities, which differ from fiber to fiber, may be responsible for differences in the MTFs observed for fibers with high discharge rates during the unmodulated DPT. Since synchrony to the modulator is highest whenever the period of the modulator coincides with inherent periodicity, (Litvak et al., 2002a), this hypothesis would suggest that modulation gain would be higher than average at frequencies whose period coincides with a multiple of the period of inherent periodicity. Qualitative analysis of 3 fibers with discharge rates above 130 spikes/s suggests that this

may indeed be the case (see Fig. 8, right, and caption). However, we do not have sufficient data to conclusively test this hypothesis.

Linearity did not hold as well for MTFs estimated from responses of fibers with discharge rates below 130 spikes/s. Because these responses were significantly rectified (e.g. Fig. 3), it is possible that linearity would hold with even smaller modulation depths. However, responses to smaller modulation would require longer data collection times to accurately estimate the shape of the period histogram.

The average MTF filter estimated from the responses of fibers with discharge rates during the unmodulated DPT above 130 spikes/s was at least approximately consistent with the average modulation gain found for the three sinusoidal modulation frequencies investigated in Litvak *et al.* (2002a). The estimated MTF showed a bandpass characteristic (pass band 400-600 Hz). In the pass band, modulation gain was extremely high (near 100). Outside the pass band, the response dropped off at about 6 dB/octave. The gentle drop-off on both the low and high-frequency sides implies that with the DPT, ANFs can represent the formant frequencies in their temporal discharge patterns up to at least 1000 Hz. The neural mechanisms responsible for the bandpass characteristic are unclear. The high-frequency drop off is broadly consistent with low pass filtering in the neural membrane. Adaptation of electric responses (Dynes and Delgutte, 1992; Killian, 1994; Litvak et al., 2002a) may lead to the drop in the magnitude of the MTF for frequencies below 10 Hz (Hammond et al., 1996). Another mechanism for the low-frequency drop off may be a partial excitation block in response to high-frequency stimulation. For example, Litvak et al. (2001) noted that down modulation of a 5 kpps carrier often led to an increase in firing rate. The mechanism for the excitation block was hypothesized to be the interaction between the dynamics of the sodium and channels and high-frequency stimulation. One possibility is that the sodium inactivation gate is never able to recover with high-frequency stimulation.

C. Stochastic threshold model

In an earlier paper (Litvak et al., 2002a), we showed that the stochastic threshold model (STM) quantitatively predicts some features of single ANF responses to sinusoidal modulations of a DPT, including growth of synchrony and discharge rate with modulation depth, dependence of threshold on discharge rate during the DPT, and

dynamic range. In order to extend the STM model to arbitrary modulators, the model's underlying MTF filter needs to be specified. Because responses to sinusoidal modulators were only measured at three modulation frequencies, the data did not allow us to estimate the underlying MTF filter with sufficient resolution.

In the present paper, we estimated the underlying MTF filter directly from responses to complex modulations. This approach assumes that (1) the system behaves linearly and that (2) the measured MTF accurately reflects the shape of the underlying MTF. We tested the linearity assumption and found that it held for a modulation depth of 1% in fibers that had discharge rates during the unmodulated DPT above 130 spikes/s. To test the second assumption, we measured the MTF of the STM and compared it to its underlying MTF. We found that, for fibers with high spontaneous discharge rates, the measured MTFs from model fibers resembled the underlying MTF. This suggests that, if the STM is applicable to modeling the responses of ANFs to complex modulations of the DPTs, then the underlying MTF can indeed be estimated from the measured MTF.

With the underlying MTF of Fig. 9, the STM model could approximately predict the shape of the period histogram of responses to vowel modulators. At the modulation depth of 2.5%, the correlation coefficient between measured and predicted period histograms is 0.89 on the average. Prediction is poorer for larger modulation depths. Responses of ANFs to stimuli with larger modulation depths are strongly influenced by the fiber's refractory properties. By adjusting the refractory characteristics of the model to better fit those of ANFs after adaptation to the DPT, it might be possible to make the STM model better approximate the responses at larger modulation depths.

Our results suggest that, with the estimated MTF filter, and at least for small modulation depths, the STM can accurately predict the discharge probability of ANFs during arbitrarily modulator waveforms. The predictive capability of the STM model for small modulation depths makes it a powerful tool for investigating the psychophysics of electric hearing. In particular, the predictions of the STM can be useful in interpreting psychophysical responses of human cochlear implant subjects to modulations of the DPT. The STM could also be useful for modeling responses of neurons in the central nervous system to electric stimulation, where a model of activity in the auditory nerve inputs is necessary.

Chapter 6

Summary and Discussion

I. Recapitulation of key results

Cochlear implants are devices that electrically stimulate the auditory nerve with the goal of restoring hearing in those with severe and profound hearing impairment. Rubinstein et al. (1999b) suggested that the representation of electric stimulus waveforms in auditory-nerve fiber (ANF) temporal discharge patterns might be improved by introducing a sustained, high-rate, desynchronizing pulse train (DPT). He argued that activity in response to the DPT may mimic spontaneous activity in a healthy ear. In the first study (Chapter 2), we investigated responses of auditory-nerve fibers to relatively brief (250 ms) modulated and unmodulated high-rate electric pulse trains presented through an intracochlear electrode in acutely-deafened, anesthetized cats. Over a narrow (1-2 dB) range of pulse train levels, discharge rates were in the range appropriate for spontaneous activity, which is below 150 spikes/s (Liberman, 1978). Given that the thresholds of ANFs at a given cochlear location could vary by as much as 12-20 dB (van den Honert and Stypulkowski, 1987a), these results suggest that for any single level of the DPT, only a small fraction of fibers would respond with discharge rates that are appropriate for spontaneous activity.

Responses to pulse trains showed variability from presentation to presentation, but differed from spontaneous activity in the shape of the envelope of the interval histogram (IH) for pulse rates above 4.8 kpps (kilo pulses per second). These IHs had a prominent mode near 5 ms that was followed by a long tail. These non-exponential interval histograms were only observed for responses to pulse trains with pulse rates above 2.4 kpps. Responses to modulated biphasic pulse trains resembled responses to tones in intact ears for small (<10%) modulation depths, suggesting that acoustic-like responses to

sinusoidal stimuli might be obtained with a DPT. However, responses resembling those to pure tones were only observed over a narrow range of levels and modulation depths.

Overall, this study suggested that a DPT may be effective over a narrow range of levels and modulation depths. However, it left open the question of whether responses to the DPT would more resemble spontaneous activity after several minutes of DPT stimulation. In particular, it seemed probable that adaptation in discharge rate may preferentially reduce the high initial discharge rates of the more sensitive fibers.

Next, we investigated responses of single ANFs from deafened cats to 5 kpps pulse trains of 10 minute duration (Chapter 3). All responses showed adaptation during the first 1-2 minutes of DPT stimulation. While most (60%) of the neurons showed a sustained response during the entire 10 minute stimulus, some (40%) responded only transiently for a minute or so. The range of discharge rates after two minutes of stimulation resembled that for normal spontaneous activity (Liberman, 1978), although about 3% of the fibers responded with rates that exceeded spontaneous by more than 50 spikes/s. The sustained discharge rates depended, to some extent, on the level of the DPT relative to the fiber's threshold. DPTs less than 2 dB above the fiber's threshold rarely evoked a sustained response. However, when the DPT level exceeded fiber threshold by more than 2 dB, the DPT level relative to threshold accounted for only 11% of the variance in the sustained discharge rates from fiber to fiber.

We described temporal discharge patterns of sustained DPT responders. For 20% of the fibers, the (envelope of the) interval histogram was statistically indistinguishable from exponential for intervals longer than the refractory period. These histograms resembled those for spontaneous activity in a normally-functioning ear (Kiang et al., 1965). We referred to fibers that exhibited an exponential interval histogram as Type IA. The majority of fibers had non-exponential period histograms. For 40% of the fibers (Type IB), the interval histogram had a large mode below 20 ms, followed by a long exponential tail. For the other 40% of the fibers (Type II), interval histograms included a mode beyond 20 ms. 50% of Type IB fibers and 100% of Type II fibers exhibited serial correlations between successive intervals. Recordings from 4 pairs of fibers showed no

evidence of correlated activity, suggesting that DPT responses are statistically independent across fibers.

In Chapter 4, we described responses to sinusoidally modulated electric pulse trains in the presence of a long-duration DPT. We modeled a stimulus composed of a DPT and a modulated pulse train as a modulated DPT with low modulation depth. We found that the ANF responses to sinusoidal modulations depended on whether the response to the unmodulated DPT was sustained or transient. For sustained DPT responders, the responses to a sinusoidally modulated DPT resembled acoustic responses to tones over a 15-25 dB range of modulation depths. These responses were stochastic, occurring randomly at one, two, or more stimulus periods. For modulation depths below about 2.5%, the period histogram resembled the sinusoidal modulator in shape, suggesting excellent coding of modulation waveform. In contrast, fibers that responded only transiently to the DPT had smaller dynamic ranges than those for sustained DPT responders; however, their dynamic ranges were still larger than the dynamic range observed for responses to electric stimulation without the DPT. Period histograms of responses of transient DPT responders did not resemble the sinusoidal modulation waveform.

A simple stochastic threshold model that incorporated Gaussian noise and membrane refractory properties predicted the dependence of responses to the modulation on the discharge rate during the unmodulated DPT for both the sustained and the transient DPT responders.

For sustained DPT responders, the representation of sinusoidal modulation depended somewhat on whether the responses of the fiber during the unmodulated DPT exhibited an exponential (Type IA) or a non-exponential (Type IB or II) interval histogram. Type IA fibers were generally more sensitive to modulation. In addition, interval histograms of these fibers resembled those for acoustic responses to tones. In contrast, responses of non-exponential fibers to modulators with low modulation depth contained periodicities that were related both to the period of the modulator and to the intrinsic periodicities

exhibited by responses during the unmodulated DPT, as revealed by the interval histogram.

In Chapter 5, we reported responses of ANFs to filtered-vowel modulators to test the hypothesis that with a DPT, ANFs can encode complex modulations in their temporal discharge patterns. As with sinusoidal modulators, we found that responses to the vowel modulators depend greatly on whether the response to the unmodulated DPT was sustained or transient. Responses of sustained DPT responders showed complex period histograms with modes that roughly corresponded to peaks in the modulator waveforms. In contrast, transient DPT responders primarily responded to the vowel fundamental. For modulation depths below 2%, period histograms of responses of ANFs with discharge rates during the unmodulated DPT above 130 spikes/s had spectra that were nearly a linear transformation of the modulator spectrum. We characterized this linear transformation by a modulation transfer function (MTF). The MTF was a band-pass function of frequency, with a broad plateau between 300 and 600 Hz. When this MTF was incorporated into the stochastic threshold model, the model was able to predict the period histogram with good accuracy for fibers with discharge rates during the unmodulated DPT between 10 to 300 spikes/s for modulation depth of 2.5% (average peak correlation coefficient of 0.89[4]). The model was less accurate in estimating period histograms for larger modulation depths.

II. Limitations of this study

A. Choice of the animal model

One must be cautious in applying the results of this study, which are based on the recording of neural activity from auditory-nerve fibers in anesthetized cats, to human cochlear implantees. Human and cat auditory nerves are morphologically different (Nadol, 1988). The differences include (1) shorter peripheral processes in cat as compared to human, and (2) unmyelinated cell body in humans, but not in cats (for Type I fibers). These peripheral differences may be of little consequence for responses to a

[4] The peak correlation coefficient was defined as the maximum correlation coefficient of the period histograms based on the recorded and model responses. We allowed delays up to ±0.4 ms to allow for cross-fiber differences in conduction delay.

DPT when its level is substantially above threshold. Responses to strongly suprathreshold stimuli are thought to originate on the central processes (van den Honert and Stypulkowski, 1984; Miller et al., 1999). While the central processes are more difficult to excite in humans than in cats (Rattay et al., 2001), only stimulation of central processes may be possible for implantees that have lost most of the peripheral dendrites due to slow retrograde degeneration beginning at the onset of deafness (Spoendin, 1975). The morphology of the cat and the human auditory nerve is more similar centrally. For example, the average diameter of the central myelinated axons of auditory-nerve fibers is 2.5 μm in both man and cat (Liberman and Oliver, 1984; Spoendlin and Schrott, 1989). However, we cannot rule out that the human auditory nerve may differ from cats' in ways that are more subtle than those reported by morphological studies.

The dimensions of the cochlear structures are different for the cat and the human (Nadol, 1988). These differences may change the potential distributions elicited by an intracochlear stimulating electrode, which in turn may affect the neural response properties by affecting the shape of the activation function (Rattay, 1990). However, the effects of these differences are likely to be greatest at locations close to the electrode, at the peripheral processes, and smaller for central processes, where responses to a relatively high-level DPT may originate.

We used a drug protocol to acutely deafen our cats several hours prior to single-unit recordings. It is unclear whether the auditory nerve of an acutely deafened cat is a good model of the auditory nerve in human cochlear implantees, who may have been deaf for as long as 50 years. Morphologically, long-term deafness leads to retrograde degeneration of the cochlear nerve (Spoendin, 1975). Because the DPT may stimulate most fibers centrally, it is likely that the responses to the DPT would be affected by the duration of deafness. However, to confirm this, the study would need to be repeated in cats deafened for at least several months.

B. Residual hearing

We were not able to achieve complete deafening in all animals. As patients with residual hearing get implanted in greater number, a significant number of future

implantees may have some remaining hair cells. The presence of hair cells might therefore not be a problem for extrapolating the results of the present study to these patients. However, it may complicate modeling of ANF responses to the DPT, because some responses may be affected by hair cell excitation. We attempted to control for this by monitoring residual hearing in the non-implanted ear. We showed in Chapter 3 that basic responses to the DPT do not depend strongly on the status of the residual hearing. As another control, we also excluded fibers with spontaneous activity, which is also suggestive of an intact hair cell[5].

C. Uncertainty on matching stimulus levels for cats and humans

A major difficulty in extrapolating the population data to human implantees is that it is unknown how the levels of the DPT used in this study correlate with the levels that would produce comfortable sensations for human listeners. For each animal, we quantified the DPT level relative to the growth curve of the EAP. In several cases, the level of the DPT was near saturation of the evoked compound action potential response. However, in humans, there is rarely a hint of saturation of the ECAP near maximum comfortable levels (Brown et al., 1994). In general, therefore, a DPT presented on a single electrode at a level that is within a comfortable range may excite fewer fibers than a DPT at the levels used in our study.

We have demonstrated that a key factor to predicting responses of ANFs to DPT modulators is the discharge rate during the unmodulated DPT. Fibers that respond to the DPT should have similar response properties to modulation in humans and in cats. In addition, because the DPT evokes activity that is spontaneous-like, it may not be heard. It may therefore be possible to slowly increase the DPT over several minutes to a level that exceeds what would have been the maximum comfortable loudness had it been applied suddenly. While this can be done for experimental purposes for up to 10 minutes, sustained stimulation at that level would probably be unsafe.

[5] However, presence of spontaneous activity does not necessarily indicate an intact hair cell (Shepherd and Javel, 1997).

III. Neural mechanisms

A. Adaptation

A ubiquitous response property of auditory-nerve fibers to high-rate electric pulse trains is the slow decrease of discharge rate with time. A time-dependent decrease in the driven neural response is usually referred to as adaptation (Chimento and Schreiner, 1990; 1991; 1992; Delgutte, 1980). Adrian (1926) first reported adaptation in sensory neurons innervating frog muscles. Although mechanisms of adaptation may differ in different neural systems, phenomologically, adaptation is nearly universal.

In response to 5 kpps electric pulse trains, ANFs adapt on at least two different time scales. First, there is a relatively fast adaptation on a time scale of 100 ms (Chapter 2). This adaptation can be observed in responses to high-rate pulse trains whose level is between threshold and saturation. We demonstrated that this faster adaptation depends not only on the driven activity, but also on the stimulus used to evoke the activity. In particular, when the initial discharge rates are normalized, adaptation is greater for 5 kpps pulse trains than for 1 kpps pulse trains. There also a slower adaptation with a time constant on the order of several minutes (Chapter 3) that can be observed for both near-threshold and high-level stimulation. We demonstrated (Chapters 4 and 5) that, in the adapted state, the membrane is capable of firing at rates as high as 600 spikes/s when a modulated stimulus is added. Our ECAP results suggest that there may be additional adaptation occuring on the scale of hours (Chapter 3). Our results suggest that the latter adaptation may occur exclusively in response to high-level stimulation.

Both the fast and the slow adaptation may reflect changes of neural homeostasis during sustained stimulation. Possible mechanisms for such changes are varied and include changes in intra-cellular or extra-cellular concentrations of potassium, sodium or calcium, or down-regulation of ion channels. Some neurons have special ionic mechanisms that are activated during sustained stimulation to reduce sensitivity. Examples of these mechanisms include sodium-activated potassium channels (Foehring et al., 1989; Sanchez-Vives et al., 2000), calcium-activated potassium channels (Sanchez-Vives et al., 2000), or activation of electrogenic pumps (Gustafsson and Wigstrom, 1983). It is unclear which combination of mechanisms is responsible for each form of

ANF adaptation. Intracellular recordings near the site of spike initiation may be necessary to resolve this issue.

B. Increase in rate at onset of down modulation

We showed (Chapter 2) that 10% sinusoidal downwards modulation of a 5 kpps pulse train tends to increase the response discharge rate. This observation is surprising, because the root-mean-squared level of the stimulus is decreased by 0.9 dB by the modulation, while the peak level remains constant. This phenomenon was also exhibited by a Hodgkin-Huxley (Hodgkin and Huxley, 1952) model. In that model, the high-rate stimulation prevented recovery of the sodium inactivation variable. However, partial recovery could occur during the modulation troughs. The mechanism for the increase of discharge rate with down modulation may be similar for ANFs.

The increase in discharge rate with down modulation demonstrates dramatically that a 5 kpps pulse train is not very efficient in evoking neural responses. This appears consistent with the adaptation data which shows that neural resources could be drained at least in the short term even while the pulse train does not evoke spikes.

C. Non-exponential interval histograms

Rubinstein et al. (1999b), using a biophysical model, predicted that responses of ANFs to 5 kpps electric pulse trains should have Poisson-like, exponential interspike interval distributions, like normal spontaneous activity. For short electric stimuli, nearly 50% of responses to high-rate stimulation exhibited non-exponential interval histograms (Chapter 2, Fig. 7). However, in response to longer (10 minute) 5 kpps pulse trains, about 80% of the neurons had non-exponential histograms (Chapter 3). The difference between the two percentages may be because longer recordings allowed us to estimate the interval histogram with greater precision and for longer intervals. For example, while most interval histograms in Chapter 2 contained on the order of 200 intervals, some histograms analyzed in Chapter 3 had as many as 20000 intervals. Therefore, the test applied in Chapter 3 is likely to be more sensitive than that in Chapter 2.

The precision of the interval histograms computed in Chapter 3 allowed us to distinguish between two types of non-exponential histograms. In one (Type IB), the

histogram could be described by a mode below 20 ms, and a long exponential tail. Type II histogram had a non-exponential tail. We associated Type II responses with bursting.

While responses to 200 ms, 5 kpps pulse trains showed non-exponential behavior primarily in neurons with discharge rates below 100 spikes/s (Chapter 2, Fig. 7), responses to sustained pulse trains were non-exponential primarily for neurons with higher discharge rates (Chapter 3, Fig. 8). It is unclear whether differences in test sensitivity may account for these differences, or whether different mechanisms may operate during short and sustained stimulation to produce non-exponential histograms.

We discussed possible neural mechanisms underlying non-exponential interval histograms in the Discussion sections of Chapters 2 and 3. We suggested that sub-threshold oscillations that are observed in some models of neural membranes may account for the preferred firing periods seen in some ANF responses to the DPT (Schneidman et al., 1998). Therefore, a biophysical model of the type proposed by Rubinstein et al. (1999b) which properly represents the neural dynamics of ANFs may account for Type IB interval histograms. To account for Type II histograms, other, slower-acting mechanisms may be required.

D. Dependence of modulation sensitivity on modulation frequency

We demonstrated that the majority of sustained DPT responders were most sensitive to 420 Hz sinusoidal modulation (Chapters 4, 5) than for 100 Hz and 830 Hz modulators. The mechanism for the drop in sensitivity at both the high and low frequency side is unclear. One possibility is that the decrease in sensitivity for higher modulation frequencies may be related to the low-pass filtering of the neural membrane. Membrane filtering has been estimated by measuring threshold to an electric sinusoid. For frequencies above 200 Hz, ANF threshold to electric sinusoids increases with frequency at a rate from 3 dB/octave to 6 dB/octave (Dynes and Delgutte, 1992). This rate of threshold decrease is broadly consistent with that reported in this study for sinusoidal modulations, which is below 6 dB/octave.

We hypothesize that the decrease in sensitivity at low frequencies may be related to the phenomenon of excitation block discussed in Section B above. Because excitation block may occur only in response to high-rate pulse trains, it may not be seen in

dependence of threshold on sinusoid frequency, which shows a minimum near a lower frequency of 100 Hz (Hartmann and Klinke, 1990).

IV. Comments on strategies that incorporate the DPT

In Chapters 3 and 4, we discussed specific proposals for future cochlear implant strategies that may incorporate a DPT. In both chapters, we described a strategy in which a moderate-level DPT would be presented on multiple electrodes, and where small modulations of the DPT would be used to encode sounds. We further suggested that the envelope detector be entirely eliminated from the DPT-enhanced CIS strategy. Instead, negative phases of the filtered waveforms could be represented as decreases in DPT level.

Here, we mention a few important points that have been omitted from discussions in previous chapters.

A. Adding onset emphasis

One of the most salient features of representation of speech on the normal auditory nerve is peaks at amplitude onsets in each frequency band (Delgutte, 1997; Young et al., 1983). The mechanism for the emphasis is thought to be rapid adaptation of neurotransmitter release by the inner hair cells (Westerman and Smith, 1984). While electrically-stimulated ANFs adapt somewhat in discharge rate (Killian, 1994; Delgutte and Cariani, 1992), the time course of adaptation to electric stimuli differs from ANF adaptation to sounds. The acoustic responses adapt in at least three stages: the rapid adaptation stage (few milliseconds), short-term adaptation (tens of milliseconds) and long-term adaptation (tens of seconds) (Ruggero, 1992). Geurts and Wouters (1999) incorporated a form of rapid adaptation into the CIS waveforms. They found modest improvements in consonant recognition when rapid adaptation was incorporated into CIS processors. Using the results of this thesis, it may be possible to design a DPT-enhanced strategy that produces ANF responses that better resemble the short-term adaptation seen in an acoustically stimulated ear than the CIS waveforms used by Geurts and Wouters (1999). Incorporating a model of rapid adaptation into a DPT-enhanced CIS strategy may lead to further improvements in understanding of speech by cochlear implant listeners.

B. Determining maximal safe level of DPT stimulation

The maximal safe level for the DPT can be determined by recording evoked responses prior to and immediately following several minutes of DPT stimulation. Temporary threshold shifts after short DPT stimulation may be indicative of sustained thresholds shifts or neural damage for longer-duration stimuli. However, a precise relationship between neural damage and temporary threshold shifts due to a short DPT remains to be determined.

C. Proper representation of stimulus level

All cochlear implant strategies must address the question of how to properly compress the 100 dB range of acoustical hearing into the 20 dB range of electric stimulation in which the responses resemble those to sounds. One possibility is to represent an increase in loudness by spreading the excitation to adjacent electrodes. To a degree, this may occur already in systems that use band-pass filters with gradual skirts. However, Another possibility is to use stimuli which evoke higher cross-fiber synchrony for the same peak level. For example, with a DPT, a waveform with narrow peaks may evoke higher cross-fiber synchrony than a waveform in which the peaks are broader[6]. Increase in inter-fiber correlation may be a possible cue for loudness in the healthy ear (Carney, 1994). Psychophysical experiments are necessary to determine whether manipulation of "peakedness" together with modulation depth could change perceived loudness sufficiently to produce the entire range of loudness sensations.

V. Future work

Our results demonstrate that with a DPT, it may be possible to produce responses on the auditory nerve of a deaf cochlea that include both the fine-time structure and envelope information in their temporal discharge patterns. This novel capability offers psychophysicists a method to directly assess the role of fine temporal information in ANF discharges in a variety of auditory tasks.

A. Can information in the temporal discharge patterns of ANFs be utilized for discrimination?

[6] However, without the DPT, both the peaked and non-peaked stimuli may produce highly synchronous responses.

Improved speech recognition is an important goal in cochlear implant design. In the acoustic domain, speech sounds can be distinguished by their spectrum, and correspondingly, by their temporal waveforms. If more information about the temporal waveform can be transmitted to cochlear implant patients, then their speech-reception performance may improve. We conclusively showed in this thesis that responses to small complex modulations of the DPT contain a great deal of information about the complex modulator waveform in their temporal discharge patterns. However, it remains to be seen whether the central processor is capable of making use of the information in these temporal discharge patterns.

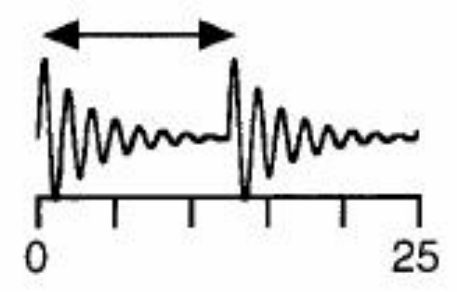


Figure 1. Two periods of a single-formant vowel (adopted from Cariani and Delgutte (1996a)).

One possible experiment is to test whether subjects can distinguish sounds based on their fine-time structure. One possibility is to use single-formant vowels (Rhode, 1998; Cariani and Delgutte, 1996a) (Fig. 1). This stimulus is completely determined by three parameters: the fundamental frequency (large arrow), the formant frequency (distance between smaller peaks), and bandwidth, which determines the rate at which the envelope decreases. The fundamental frequency could be fixed at 100 Hz, while the first formant frequency could be varied continuously. The stimulus could be presented on the most apical electrode both as a modulator of a sustained DPT, and, as a control, in isolation. In both cases, the subject's task would be to distinguish single-formant vowels based on the formant frequency in a forced-choice test. To minimize the possible effects of temporal patterns on loudness, modulation depth (or in the control case, level) of the stimulus should be roved. We predict that because complex temporal patterns are poorly represented in electrically stimulated ANFs, for large bandwidths subjects will have trouble distinguishing vowels based on formants without a DPT. However, distinctions should improve when the DPT is added.

In subsequent experiments, one could explore whether discrimination of single-formant vowels could be improved through training (both with and without a DPT).

Finally, it would be of interest to determine whether this training improves the reception of speech.

B. Does fine-time structure contribute to perception of virtual pitch?

Pitch is defined as "that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale" (American Standards Association, 1960). Normal-hearing listeners report that complex tones containing frequency components that are harmonics of a low-frequency fundamental have a low pitch that corresponds to the frequency of the fundamental. The low pitch of a complex tone is reported even when the complex tone does not contain energy at the fundamental. In this case, the low pitch is referred to as "virtual pitch." Whether the virtual pitch is heard by first hearing out the harmonics, and then centrally synthesizing the harmonic pattern into a low-pitch sensation (e.g. (Goldstein, 1973)), or whether the ear detects pitch by detecting periodicities in the temporal discharge patterns of ANFs (e.g. (Schouten, 1940; Cariani and Delgutte, 1996a; Cariani and Delgutte, 1996b; Meddis and Hewitt, 1991)) is a matter of active debate.

Phase locking of ANFs is more accurate to electric sinusoids than to acoustic tones for frequencies up to 10 kHz (Hartmann et al., 1982; Dynes and Delgutte, 1992). However, discrimination of the rate of the electric pulse by cochlear implantees is less accurate than discrimination of pitch of simple and complex tones in normal-hearing subjects. In addition, pitch scaling experiments indicate that pitch sensation saturates for frequencies above about 500 Hz (Eddington et al., 1978; Shannon, 1983; Dorman et al., 1994). These observations have been used as arguments against temporal pitch theories (e.g. Viemeister et al. (1999)).

Although these findings conclusively disprove *some* temporal pitch theories, they by no means argue that temporal structure is not important for pitch. One possibility is that a combination of temporal and spatial cues may be essential for proper perception of virtual pitch. Alternatively, temporal analyzers of pitch in the central nervous system could be critically dependent on the discharge patterns being acoustic-like. Responses to electric pulse trains are hyper-synchronized, and for trains with rates higher than 300 pps, produce higher discharge rates than responses to sound in a healthy ear. For complex

electric modulators, only the fundamental, and not the fine-time structure, is represented in the temporal discharge patterns.

The hypothesis that temporal discharge patterns must be acoustic-like to evoke a robust pitch sensation can be tested by comparing pitch perception of electric pulse trains and complex modulations with perception of pitch evoked by small (<5%) modulations of a high-rate DPT. If the DPT evokes temporal discharge patterns that are sufficiently similar to those evoked in a healthy ear, then pitch discrimination might be improved by the DPT. In addition, the sinusoidal frequency at which pitch sensation saturates should be increased with the DPT. Finally, if fine-time structure is important for evoking virtual pitch sensation, then pitch discrimination should be more accurate, and pitch should be more salient for complex modulators of the DPT than for either low-frequency pulse trains, or complex modulators alone.

C. Does proper coding of fine-time structure allow for more binaural advantages for bilateral implantees?

Binaural hearing is important for hearing out sounds of interest among competing sound sources. At least for sounds containing significant components below 1000 Hz, binaural processing depends critically on the ability of the auditory system to analyze small timing differences between the two ears (Wightman and Kistler, 1992). It is hoped that advantages of binaural hearing can be conferred to bilateral cochlear-implant patients. A key issue, however, is how to encode stimuli so that patients obtain a maximal benefit.

Sensitivity of cochlear implant patients to interaural time differences (ITD) of electric pulse trains is poorer than that for normal-hearing listeners (van Hoesel et al., 1993; Van Hoesel and Clark, 1995; van Hoesel and Clark, 1997; Lawson et al., 1998). This is true even if electrode location and level have been optimized in the two ears to give best ITD sensitivity (Long et al., 2001). One possibility is that the poor performance of cochlear implant patients is due to differences in the representation between electric and acoustic pulse trains in ANF responses. In particular, the hypersynchronized (Hartmann et al., 1984) and extremely high discharge rate (Moxon, 1967; Javel et al., 1987) responses evoked by electric pulse trains may “swamp” the finely tuned central detectors that

analyze small temporal differences in the responses from the two ears. Since ANF responses to small (<2%) modulations of a sustained high-rate DPT have lower synchrony, and evoke smaller discharge rates than electric pulses presented in isolation, this hypothesis would predict that binaural sensitivity should be improved by the DPT. Studies to test this hypothesis are already underway in our laboratory.

Overall, this thesis shows that a DPT, when used in a combination with a strategy that includes the fine-time structure sound information, may improve performance of cochlear implantees. In addition, a DPT may become a valuable research tool in investigating psychophysics and physiology of hearing.

Bibliography

Abbas, P. J., C. A. Miller, J. T. Rubinstein, and A. J. Matsuoka. 1999. Tenth Quarterly Progress Report: The Neurophysiological Effects of Simulated Auditory Prosthesis Simulation. University of Iowa, Iowa City, Iowa.

Adrian, E. D. 1926. The impulses produced by sensory nerve endings. *J. Physiol.* 61:49-72.

Assmann, P. F., and Q. Summerfield. 1990. Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. *J Acoust Soc Am.* 88:680-97.

American Standards Association 1960. Acoustical Terminology S1, 1-1960. American Standards Association, New York.

Battmer, R. D., Y. Zilberman, P. Haake, and T. Lenarz. 1999. Simultaneous Analog Stimulation (SAS)--Continuous Interleaved Sampler (CIS) pilot comparison study in Europe. *Ann Otol Rhinol Laryngol Suppl.* 177:69-73.

Bekesy, G. v. 1960. Experiments in Hearing. McGraw-Hill, New York.

Brown, C. J., P. J. Abbas, H. Fryauf-Bertschy, D. Kelsay, and B. J. Gantz. 1994. Intraoperative and postoperative electrically evoked auditory brain stem responses in nucleus cochlear implant users: implications for the fitting process. *Ear Hear.* 15:168-76.

Bruce, I. C., L. S. Irlicht, M. W. White, S. J. O'Leary, S. Dynes, E. Javel, and G. M. Clark. 1999. A stochastic model of the electrically stimulated auditory nerve: pulse- train response. *IEEE Trans Biomed Eng.* 46:630-7.

Cariani, P. A., and B. Delgutte. 1996a. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol.* 76:1698-716.

Cariani, P. A., and B. Delgutte. 1996b. Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J Neurophysiol.* 76:1717-34.

Carney, L. H. 1994. Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness. *Hear Res.* 76:31-44.

Chimento, T. C., and C. E. Schreiner. 1990. Time course of adaptation and recovery from adaptation in the cat auditory-nerve neurophonic. *J Acoust Soc Am.* 88:857-64.

Chimento, T. C., and C. E. Schreiner. 1991. Adaptation and recovery from adaptation in single fiber responses of the cat auditory nerve. *J Acoust Soc Am.* 90:263-73.

Chimento, T. C., and C. E. Schreiner. 1992. Adaptation and recovery from adaptation of the auditory nerve neurophonic (ANN) using long duration tones. *Hear Res.* 62:131-41.

Collins, J. J., C. C. Chow, and T. T. Imhoff. 1995. Stochastic resonance without tuning. *Nature.* 376:236-8.

Cox, D. R. 1967. Renewal theory. Methuen, London,.

Delgutte, B. 1980. Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers. *J Acoust Soc Am.* 68:843-57.

Delgutte, B. 1982. Some correlates of phonetic distinctions at the level of the auditory nerve. *In The Representation of Speech in the Peripheral Auditory System*. R. C. a. B. Granström, editor. Elsevier, Amsterdam. 131-150.

Delgutte, B. 1984. Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds. *J. Acoust. Soc. Am.* 75(3):879-886.

Delgutte, B. 1997. Auditory Neural Processing of Speech. *In* The Handbook of Phonetic Sciences. W. J. Hardcastle and J. Laver , editors. Blackwell Handbooks in Linguistics.

Delgutte, B., and P. Cariani. 1992. Coding of the pitch of harmonic and inharmonic complex tones in the interspike intervals of auditory nerve fibers. *In* The Processing of Speech. M. E. H. Schouten, editor. Mouton-DeGruyer, Berlin. 37-45.

Delgutte, B., B. M. Hammond, and P. A. Cariani. 1999. Neural coding of the temporal envelope of speech. *In* Listening to Speech. S. Greenberg and W. Ainsworth, editors.

Delgutte, B., and N. Y. S. Kiang. 1984. Speech coding in the auditory nerve: I. Vowel-like sounds. *J. Acoust. Soc. Am.* 75(3):866-878.

Djourno, A., and C. Eyriès. 1957. Prothèse auditive par excitation électrique a distance du nerf sensoriel a l'aid d'un bobinage inclus a demeure. *Presse Medicale.* 65:1417.

Dorman, M. F., M. Smith, L. Smith, and J. L. Parkin. 1994. The pitch of electrically presented sinusoids. *J Acoust Soc Am.* 95:1677-9.

Duckert, L. G., and J. M. Miller. 1982. Acute morphological changes in guinea pig cochlea following electrical stimulation. A preliminary scanning electron microscope study. *Ann Otol Rhinol Laryngol.* 91:33-40.

Dynes, S. B., and B. Delgutte. 1992. Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. *Hear Res.* 58:79-90.

Dynes, S. C. 1995. Discharge Characteristics of Auditory Nerve Fibers for Pulsatile Electrical Stimuli. Ph.D. Thesis. MIT, Cambridge, MA.

Eddington, D., W. Dobelle, D. Brackmann, M. Mladejousky, and J. Parkin. 1978. Auditory prosthesis research with multiple channel intracochlear stimulation in man. *Ann. Oto. Rhino. Laryn.* 87 Suppl.:1-39.

Eddington, D. K., W. R. Rabinowitz, J. Tierney, V. Noel, and M. Whearty. 1997. Speech Processors for Auditory Prostheses, 8th Quarterly Progress Report, NIH Contract N01-DC-6-2100.

Efron, B., and R. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, New York.

Favre, E., and M. Pelizzone. 1993. Channel interactions in patients using the Ineraid multichannel cochlear implant. *Hear Res.* 66:150-6.

Foehring, R. C., P. C. Schwindt, and W. E. Crill. 1989. Norepinephrine selectively reduces slow Ca2+- and Na+-mediated K+ currents in cat neocortical neurons. *J Neurophysiol.* 61:245-56.

Frankenhaeuser, B., and A. F. Huxley. 1964. The action potential in the myelinated nerve fibre of *xenopus laevis* as computed on the basis of voltage clamp data. *J. Physiol.* 171:302-315.

Friesen, L. M., R. V. Shannon, D. Baskent, and X. Wang. 2001. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *Journal of Acoustical Society of America.* 110:1150-1163.

Gaumond, R. P. 1980. Studies of the stimulus and recovery dependence of cat cochlear nerve fiber discharge probability. Ph.D. Thesis. Washington University, St. Louis, MD.

Geurts, L., and J. Wouters. 1999. Enhancing the speech envelope of continuous interleaved sampling processors for cochlear implants. *J Acoust Soc Am.* 105:2476-84.

Gfeller, K., A. Christ, J. F. Knutson, S. Witt, K. T. Murray, and R. S. Tyler. 2000. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J Am Acad Audiol.* 11:390-406.

Goldstein, J. L. 1973. An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am.* 54:1496-516.

Good, P. I. 1994. Permutation tests : a practical guide to resampling methods for testing hypotheses. Springer-Verlag, New York.

Gustafsson, B., and H. Wigstrom. 1983. Hyperpolarization following long-lasting tetanic activation of hippocampal pyramidal cells. *Brain Res.* 275:159-63.

Guttman, R., and R. Barnhill. 1970. Oscillation and repetitive firing in squid axons. Comparison of experiments with computations. *J Gen Physiol.* 55:104-18.

Guttman, R., S. Lewis, and J. Rinzel. 1980. Control of repetitive firing in squid axon membrane as a model for a neuroneoscillator. *J Physiol (Lond).* 305:377-95.

Hammond, B. M., W. M. Rabinowitz, and B. Delgutte. 1996. Modulation transfer functions of auditory-nerve fibers: Measurements and use in predicting the neural response to speech. *In* Assoc. Res. Otolaryngol. 78.

Hartmann, R., and R. Klinke. 1990. Impulse patterns of auditory nerve fibres to extra- and intracochlear electrical stimulation. *Acta Otolaryngol Suppl.* 469:128-34.

Hartmann, R., G. Topp, and R. Klinke. 1982. Comparison of auditory single fiber responses during acoustic and electric stimulation of the intact cat cochlea. *Arch Otorhinolaryngol.* 234:187-8.

Hartmann, R., G. Topp, and R. Klinke. 1984. Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea. *Hear Res.* 13:47-62.

Hill, A. V. 1936. Excitation and accomodation in nerve. *Proc. R. Soc. B.* 119:305-355.

Hodgkin, A. L., and A. F. Huxley. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117:500-544.

House, W. F., and J. Urban, " 82:504-517. 1973. Long term results of electrode implantation and electronic stimulation of the cochlea in man. *Ann. Otol. Rhinol. Laryngol.* 82:504-517.

Javel, E. 1989. Acoustic and Electrical Encoding of Temporal Information. *In* Cochlear Implants: Models of the Electrically Stimulated Ear. J. M. Miller and F. A. Spelman, editors. Springer-Verlag, New York, NY. 249-299.

Javel, E., Y. C. Tong, R. K. Shepherd, and G. M. Clark. 1987. Responses of cat auditory nerve fibers to biphasic electrical current pulses. *Ann. Otol., Rhinol., Laryngol.* 96 Suppl. 128:26-30.

Johnson, D. H. 1980. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am.* 68:1115-22.

Johnson, D. H. 1996. Point process models of single-neuron discharges. *J Comput Neurosci.* 3:275-99.

Johnson, D. H., and N. Y. S. Kiang. 1976. Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers. *Biophysical Journal.* 16:719-734.

Joris, P. X., L. H. Carney, P. H. Smith, and T. C. Yin. 1994. Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol.* 71:1022-36.

Kelly, O. E., D. H. Johnson, B. Delgutte, and P. Cariani. 1993. Factors affecting the fractal character of auditory nerve activity. *In* ARO MidWinter Meeting.

Kelly, O. E., D. H. Johnson, B. Delgutte, and P. Cariani. 1996. Fractal noise strength in auditory-nerve fiber recordings. *J Acoust Soc Am.* 99:2210-20.

Kiang, N., E. Moxon, and R. Levine. 1970a. Auditory-nerve activity in cats with normal and abnormal cochleas. *In* Sensorineural hearing loss. Ciba Found Symp. 241-273.

Kiang, N. Y., and E. C. Moxon. 1972. Physiological considerations in artificial stimulation of the inner ear. *Ann Otol Rhinol Laryngol.* 81:714-30.

Kiang, N. Y.-S. 1965. Stimulus coding in the auditory nerve and cochlear nucleus. *Acta Otolaryngol.* 59:186-200.

Kiang, N. Y. S., E. C. Moxon, and R. A. Levine. 1970b. Auditory nerve activity in cats with normal and abnormal cochleas. *In* Sensorineural Hearing Loss. G. E. W. Wolstenholme and J. Knight, editors. Churchill, London. 241-268.

Kiang, N. Y. S., T. Watanabe, E. C. Thomas, and L. F. Clark. 1965. Discharge Patterns of Single FIbers in the Cat's Auditory Nerve. The MIT Press, Cambridge, MA.

Killian, M. J. P. 1994. **Excitability of the Electrically Stimulated Auditory Nerve**. Ph.D. University of Utrecht, Utrecht.

Klatt, D. H. 1980. Software for cascade/parallel formant synthesizer. *J. Acoust.Soc. Am.* 67:971-995.

Knauth, M., R. Hartmann, and R. Klinke. 1994. Discharge pattern in the auditory nerve evoked by vowel stimuli: a comparison between acoustical and electrical stimulation. *Hear Res.* 74:247-58.

Laneau, J., and J. Wouters. 2001. **Fractal effects and models of electrically stimulated nerve fibers**. *In* Proceedings of the International Conference on NeuroInformatics. in press, Wenen, Austria.

Lawson, D. T., B. S. Wilson, M. Zerbi, C. van den Honert, C. C. Finley, J. C. Farmer, Jr., J. T. McElveen, Jr., and P. A. Roush. 1998. Bilateral cochlear implants controlled by a single speech processor. *Am J Otol.* 19:758-61.

Leake, P. A., A. L. Kuntz, C. M. Moore, and P. L. Chambers. 1997. Cochlear pathology induced by aminoglycoside ototoxicity during postnatal maturation in cats. *Hear Res.* 113:117-32.

Levitt, H. 1971. Transformed Up-Down Methods in Psychoacoustics. *J Acoust Soc Am.* 49(2):467-477.

Liberman, M. C. 1978. Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am.* 63:442-55.

Liberman, M. C., and N. Y. Kiang. 1978. Acoustic trauma in cats. Cochlear pathology and auditory-nerve activity. *Acta Otolaryngol Suppl.* 358:1-63.

Liberman, M. C., and M. E. Oliver. 1984. Morphometry of intracellularly labeled neurons of the auditory nerve: Correlations with functional properties. *J. Comp. Neurol.* 223:163-176.

Litvak, L. M., B. Delgutte, and D. K. Eddington. 2001. Auditory nerve fiber responses to electric stimulation: modulated and unmodulated pulse trains. *JASA.* Submitted.

Litvak, L. M., B. Delgutte, and D. K. Eddington. 2002a. Responses of auditory nerve fibers to sustained high frequency electric stimulation: II. Sinusoidal Modulators. *Manuscript, to be submitted to JASA.* None.

Litvak, L. M., B. Delgutte, and D. K. Eddington. 2002b. Responses of auditory nerve fibers to sustained high frequency electric stimulation: III. Vowel Modulators. *Manuscript, to be submitted to JASA.* None.

Litvak, L. M., Z. Smith, B. Delgutte, and D. K. Eddington. 2002c. Responses of auditory nerve fibers to sustained high frequency electric stimulation: I. Unmodulated pulse trains. *Manuscript, to be submitted to JASA.* None.

Long, C., D. Eddington, H. Colburn, and W. Rabinowitz. 2001. Sensitivity to interaural time difference as a function of interaural electrode position in a cochlear implant user. *In* Conference on Implantable Auditory Prostheses. 20.

Lowen, S. B., and M. C. Teich. 1992. Auditory-nerve action potentials form a nonrenewal point process over short as well as long time scales. *J Acoust Soc Am.* 92:803-6.

Lowen, S. B., and M. C. Teich. 2000. Toward fractal coding in auditory prostheses. *In* Cochlear Implants. W. A. e. al., editor. Thieme Medical Publishers, New York. 57-59.

McKinney, M. F. a. D., B. (1998). "," Abstracts of the 1998 549. 1998. Correlates of the subjective octave in auditory-nerve fiber responses: Effect of phase-locking and refractoriness. *In* Midwinter Meeting of the Association for Research in Otolaryngology, Florida. 549.

Meddis, R., and M. J. Hewitt. 1991. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. *Journal of Acoustical Society of America.* 89:2866-2882.

Menzel, D. H. 1960. Fundamental Formulas of Physics. Dover, New York.

Miller, C. A., P. J. Abbas, B. K. Robinson, J. T. Rubinstein, and A. J. Matsuoka. 1999. Electrically evoked single-fiber action potentials from cat: responses to monopolar, monophasic stimulation. *Hearing Research.* 130:197-218.

Miller, M. I., and M. B. Sachs. 1984. Representation of voice pitch in discharge patterns of auditory-nerve fibers. *Hear Res.* 14:257-79.

Miller, M. I., and J. Wang. 1993. A new stochastic model for auditory-nerve discharge. *J Acoust Soc Am.* 94:2093-107.

Moller, A. R. 1970. The use of correlation analysis in processing neuroelectric data. *In* Progress in Brain Research, Computers and Brains. J. P. Schade and J. Smith, editors. Elsevier, Amsterdam. 87-99.

Moller, A. R. 1973. Statistical evaluation of the dynamic properties of cochlear nucleus units using stimuli modulated with pseudorandom noise. *Brain Res.* 57:443-456.

Moller, A. R., and A. Rees. 1986. Dynamic properties of the responses of single neurons in the inferior colliculus of the rat. *Hear Res.* 24:203-15.

Moxon, E. C. 1967. Electric Stimulation of the cat's cochlea: a study of discharge rates in single auditory nerve fibers. Master of Science. MIT, Cambridge, MA.

Moxon, E. C. 1971. Neural and mechanical responses to electrical stimulation of the cat's inner ear. MIT, Cambridge.

Nadol, J. B., Jr. 1988. Comparative anatomy of the cochlea and auditory nerve in mammals. *Hear Res.* 34:253-66.

Parkins, C. W. 1989. Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys. *Hearing Research.* 41:137-168.

Peterson, G. E., and H. L. Barney. 1952. Control methods used in a study of vowels. *Journal of the Acoustical Society of America.* 24:175-184.

Rattay, F. 1990. Electrical Nerve Stimulation: Theory, Experiments, and Applications. Springer-Verlag/Wien, Vienna, Austria.

Rattay, F., R. N. Leao, and H. Felix. 2001. A model of the electrically excited human cochlear neuron. II. Influence of the three-dimensional cochlear structure on neural excitability. *Hear Res.* 153:64-79.

Rhode, W. S. 1998. Neural encoding of single-formant stimuli in the ventral cochlear nucleus of the chinchilla. *Hear Res.* 117:39-56.

Rose, J. E., J. R. Brugge, D. J. Anderson, and J. E. Hind. 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.* 30:769-793.

Rubinstein, J. T. 1991. Analytical theory for extracellular electrical stimulation of nerve with focal electrodes. II. Passive myelinated axon. *Biophys J.* 60:538-55.

Rubinstein, J. T., P. J. Abbas, and C. A. Miller. 1998. The Neurophysiological Effects of Simulated Auditory Prosthesis Simulation: 8th Quarterly Progress Report. University of Iowa, Iowa City, Iowa.

Rubinstein, J. T., C. A. Miller, P. J. Abbas, and B. S. Wilson. 1999a. Emulating physiologic firing patterns of auditory neurons with electrical stimulation. *In* Association for Research in Otolaryngology Midwinter Meeting, St. Petersburg Beach, Florida.

Rubinstein, J. T., B. S. Wilson, C. C. Finley, and P. J. Abbas. 1999b. Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res.* 127:108-18.

Ruggero, M. 1992. Physiology and coding of sound in the auditory nerve. *In The Mammalian Auditory Pathway: Neurophysiology.* R. F. AN Popper, editor. Springer-Verlag, New-York. 34-93.

Ruggero, M. A., N. C. Rich, L. Robles, and B. G. Shivapuja. 1990. Middle-ear response in the chinchilla and its relationship to mechanics at the base of the cochlea. *J Acoust Soc Am.* 87:1612-29.

Sachs, M. B., and P. J. Abbas. 1974. Rate versus level functions for auditory-nerve fibers in cats: tone- burst stimuli. *J Acoust Soc Am.* 56:1835-47.

Sachs, M. B., H. F. Voigt, and E. D. Young. 1983. Auditory nerve representation of vowels in background noise. *J. Neurophysiol.* 50:27-45.

Sachs, M. B., R. L. Winslow, and C. C. Blackburn. 1988. Representation of Speech in the auditory periphery. *In* Auditory Function: Neurobiological Bases of Hearing. G. M. Edelman, W. E. Gall, and W. M. Cowman, editors. Wiley, New York. 747-774.

Sachs, M. B., and E. D. Young. 1979. Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate. *J. Acoust. Soc. Am.* 66:470-479.

Sanchez-Vives, M. V., L. G. Nowak, and D. A. McCormick. 2000. Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. *J Neurosci.* 20:4286-99.

Schneidman, E., B. Freedman, and I. Segev. 1998. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput.* 10:1679-703.

Schouten, J. F. 1940. The residue and the mechanism of hearing. *Proc. K. Ned. Akad. Wet.* 43:991-999.

Shannon, R. V. 1983. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear Res.* 11:157-89.

Shannon, R. V. 1992. Temporal modulation transfer functions in patients with cochlear implants. *J. Acoust. Soc. Am.* 91:2156-2164.

Shannon, R. V., F. G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid. 1995. Speech recognition with primarily temporal cues. *Science.* 270:303-4.

Shepherd, R. K. 1986. Cochlear Prosthesis: Safety Investigations. Ph.D. Thesis. University of Melbourne, Melbourne.

Shepherd, R. K., and E. Javel. 1997. Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status. *Hear Res.* 108:112-44.

Simons, F. B. 1964. Electrical Stimulation of the Auditory Nerve in man. *Arch Otolaryngol.* 79:559-567.

Smith, Z., B. Delgutte, and A. Oxenham. 2002. Auditory chimeras. *Nature.* Submitted.

Spoendin, H. 1975. Retrograde degeneration of the cochlear nerve. *Acta Otolaryngol.* 79:266-75.

Spoendlin, H., and A. Schrott. 1989. Analysis of the human auditory nerve. *Hearing Research.* 43:25-38.

Taylor, M. M., and C. D. Creelman. 1967. PEST: Efficient estimates on probability functions. *J. Acoust. Soc. Am.* 41:782-787.

Teich, M. C. 1989. Fractal character of the auditory neural spike train. *IEEE Trans Biomed Eng.* 36:150-160.

Teich, M. C., and S. M. Khanna. 1985. Pulse-number distribution for the neural spike train in the cat's auditory nerve. *J. Acoust. Soc. Am.* 77(3):1110-1128.

Tukey, J. W. 1977. Exploratory data analysis. Addison-Wesley Pub. Co., Reading.

Tykocinski, M., R. K. Shepherd, and G. M. Clark. 1995a. Electrophysiologic effects following acute intracochlear direct current stimulation of the guinea pig cochlea. *Ann Otol Rhinol Laryngol Suppl.* 166:68-71.

Tykocinski, M., R. K. Shepherd, and G. M. Clark. 1995b. Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates. *Hear Res.* 88:124-42.

Tykocinski, M., R. K. Shepherd, and G. M. Clark. 1997. Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates. II. Comparison of fixed amplitude with amplitude modulated stimuli. *Hear Res.* 112:147-57.

van den Honert, C., and P. H. Stypulkowski. 1984. Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings. *Hear Res.* 14:225-43.

van den Honert, C., and P. H. Stypulkowski. 1987a. Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. *Hear Res.* 29:195-206.

van den Honert, C., and P. H. Stypulkowski. 1987b. Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hear Res.* 29:207-22.

Van Hoesel, R. J., and G. M. Clark. 1995. Fusion and lateralization study with two binaural cochlear implant patients. *Ann Otol Rhinol Laryngol Suppl.* 166:233-5.

van Hoesel, R. J., and G. M. Clark. 1997. Psychophysical studies with two binaural cochlear implant subjects. *J Acoust Soc Am.* 102:495-507.

van Hoesel, R. J., Y. C. Tong, R. D. Hollow, and G. M. Clark. 1993. Psychophysical and speech perception studies: a case report on a binaural cochlear implant subject. *J Acoust Soc Am.* 94:3178-89.

Viemeister, N. F., G. Donaldson, and D. A. Nelson. 1999. Comparison between the basic psychophysics of acoustic and electric hearing. *In* Conference on Implantable Auditory Prostheses, Asilomar, CA.

Volta, A. 1800. On the Electricity Excited by Mere Contact of Conductiong Substances of Different Kinds. *Trans. Roy. Acad. of Phil.* 90:403-431.

Walsh, B. T., J. B. Miller, R. R. Gacek, and N. Y. S. Kiang. 1972. Spontaneous activity in the eighth cranial nerve of the cat. *Intern. J. Neuroscience.* 3:221-236.

Westerman, L. A., and R. L. Smith. 1984. Rapid and short-term adaptation in auditory nerve responses. *Hearing Res.* 15:249-260.

Wiesenfeld, K., and F. Moss. 1995. Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. *Nature.* 373:33-6.

Wightman, F. L., and D. J. Kistler. 1992. The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am.* 91:1648-61.

Wilson, B. 1993. Chapter 2: Signal Processing. *In* Cochlear implants : audiological foundations. R. S. Tyler, editor. Singular Pub. Group, San Diego, Calif. 35-84.

Wilson, B., C. C. Finley, D. T. Lawson, and M. Zebri. 1997. Temporal Representations With Cochlear Implants. *American Journal of Otology*. 18, Suppl. 6:S30-S34.

Wilson, B., D. Lawson, M. Zerbi, C. Finley, and C. van den Honert. 1998. Speech Processors for Auditory Prostheses. Research Triangle Institute, Research Triangle Park, NC.

Wilson, B. S., C. C. Finley, J. C. Farmer, D. T. Lawson, B. A. Weber, R. D. Wolford, P. D. Kenan, M. W. White, M. M. Merzenich, and Schin. 1987. Comparative studies of speech processing strategies for cochlear implants. *To be pub. in Laryngoscope*.

Wilson, B. S., C. C. Finley, D. T. Lawson, R. D. Wolford, D. K. Eddington, and W. M. Rabinowitz. 1991. Better speech recognition with cochlear implants. *Nature.* 352:236-238.

Xu, S. A., R. K. Shepherd, Y. Chen, and G. M. Clark. 1993. Profound hearing loss in the cat following the single co-administration of kanamycin and ethacrynic acid. *Hear Res.* 70:205-15.

Young, E. D., J. A. Costalupes, and D. J. Gibson. 1983. Representation of acoustic stimuli in the presence of background sounds: adaptation in the auditory nerve and cochlear nucleus.

Young, E. D., and M. B. Sachs. 1979. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory nerve fibers. *J. Acoust. Soc. Am.* 66:1381-1403.