Coincidence Detection in the Cochlear Nucleus: Implications for the Coding of Pitch

by

Grace I. Wang

B.S. Electrical Engineering, University of Illinois, Urbana-Champaign, 2005
S.M. Electrical Engineering, Massachusetts Institute of Technology, 2007

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Signature of author:
.....
Department of Electrical Engineering and Computer Science
May 19, 2011

Certified by:
.....
Bertrand Delgutte, Ph.D.
Professor of Otology and Laryngology, Harvard Medical School
Thesis Supervisor

Accepted by:
.....
Leslie A. Kolodziejski, Ph.D.
Chair of the Committee on Graduate Students

Coincidence Detection in the Cochlear Nucleus:
Implications for the Coding of Pitch

By

Grace I. Wang

B.S. Electrical Engineering, University of Illinois, Urbana-Champaign, 2005
S.M. Electrical Engineering, Massachusetts Institute of Technology, 2007

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science
at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Abstract

The *spatio-temporal pattern* in the auditory nerve (AN), i.e. the temporal pattern of AN fiber activity across the tonotopic axis, provides cues to important features in sounds such as pitch, loudness, and spatial location. These spatio-temporal cues may be extracted by central neurons in the cochlear nucleus (CN) that receive inputs from AN fibers innervating different cochlear regions and are sensitive to their relative timing. One possible mechanism for this extraction is cross-frequency coincidence detection (CD), in which a central neuron converts the degree of cross-frequency coincidence in the AN into a rate response by preferentially firing when its AN inputs across the tonotopic axis discharge in synchrony.

We implemented a CD model receiving AN inputs from varying extents of the tonotopic axis, and compared responses of model CD cells with those of single units recorded in the CN of the anesthetized cat. We used Huffman stimuli, which have flat magnitude spectra and a single phase transition, to systematically manipulate the relative timing across AN fibers and to evaluate the sensitivity of model CD cells and CN units to the spatio-temporal pattern of AN discharges. Using a maximum likelihood approach, we found that certain unit types (primary-like-with-notch and some phase lockers) had responses consistent with cross-frequency CD cell. Some of these CN units provide input to neurons in a binaural circuit that process cues for sound localization and are sensitive to interaural level differences. A possible functional role of a cross-frequency CD mechanism in the CN is to increase the dynamic range of these binaural neurons. However, many other CN units had responses more consistent with AN fibers than with CD cells.

We hypothesized that CN units resembling cross-frequency CD cells (as determined by their responses to Huffman stimuli) would convert spatio-temporal cues to pitch in the AN into rate cues that are robust with level. We found that, in response to harmonic complex tones, cross-frequency CD cells and some CN units (primary-like-with-notch

and choppers) maintained robust rate cues at high levels compared to AN fibers, suggesting that at least some CN neurons extend the dynamic range of rate representations of pitch beyond that found in AN fibers. However, there was no obvious correlation between robust rate cues in individual CN units and similarity to cross-frequency CD cells as determined by responses to Huffman stimuli. It is likely that a model including more realistic inputs, membrane channels, and spiking mechanism, or other mechanisms such as lateral inhibition or spatial and temporal summation over spatially distributed inputs would provide insight into the neural mechanisms that give rise to the robust rate cues observed in some CN units.

Thesis Supervisor:
Bertrand Delgutte, Ph.D., Professor of Otology and Laryngology, Harvard Medical School

Acknowledgments

First and foremost, I owe my deepest gratitude to my advisor Bertrand for his invaluable assistance, support, and guidance. His attention to detail and eagerness to learn have set an example I hope to match someday.

I would like to also thank my thesis committee members Denny Freeman, John Guinan, and Al Oppenheim for sharing their expertise and steering me in the right direction.

This thesis would not have been possible without software help from Ken Hancock, expert surgical preparations from Connie Miller, and the friendship and support of the members of the Eaton Peabody Laboratory (EPL), especially fellow iglooers Yoojin Chung, Mitch Day, Sasha Devore, Nik Francis, Wendy Gu, Harrison Lin, SR Prakash, Michael Slama, and Bo Wen.

Finally, I am extremely grateful to my parents, my brother Nick, and my sister Rita for their love, and to John for helping me to look forward to tomorrow.

Table of Contents

CHAPTER 1

General Introduction

As sound enters the inner ear, the cochlea performs mechanical frequency analysis that maps frequency to place along the cochlear axis. In response to a low frequency pure tone (<5 kHz), peripheral auditory fibers in the auditory nerve tend to fire at a specific phase of the stimulus, a phenomenon called *phase locking*. The temporal pattern of activity across the population of these fibers, or the *spatio-temporal pattern* of discharges, ("spatio" from the tonotopic map, and "temporal" from phase locking) contains cues to important features in sounds, such as loudness, pitch, and spatial location. It is currently unclear how central neurons make use of these spatio-temporal cues at the periphery. One possible mechanism to extract these cues is cross-frequency coincidence detection (CD), in which a central neuron preferentially fires when its fiber inputs across the tonotopic axis discharge in synchrony, essentially converting the peripheral spatio-temporal cues into a rate response. This thesis seeks to understand whether central neurons perform a cross-frequency CD operation, and if so, what implications this would have for the neural mechanisms responsible for pitch extraction.

Peripheral Auditory Processing

The inner ear breaks up the wideband acoustic signals into a series of narrowband signals that are converted into electrical signals by hair cells. The auditory filters of the cochlea

consist of overlapping band-pass filters that are tonotopically mapped (Greenwood 1961, 1990). This frequency-to-place mapping of the basilar membrane is due to variation in its stiffness along the length of the cochlea. Its narrow and stiff base vibrates maximally for high frequencies, whereas its wider and more compliant apex is most responsive for low frequencies (von Bekesy 1960).

A healthy cat cochlea contains approximately 2600 inner hair cells (Retzius 1884), each of which is innervated by about 20 auditory nerve (AN) fibers. The tonotopically arranged inner hair cells transduce their mechanical inputs (from the vibration of the basilar membrane) into electrical outputs (to generate action potentials in AN fibers) and do so maximally near places of mechanical resonance along the basilar membrane (Pickles 1988). Each fiber has a frequency selectivity that is similar to the mechanical selectivity of the point that it innervates along the basilar membrane, and consequently each fiber is tuned to a particular characteristic frequency (CF). The frequency-to-place map in the cochlea is the foundation for a place code for frequency maintained up to the primary auditory cortex (A1).

AN fibers tend to phase lock to a specific phase of a low frequency pure tone stimulus. Due to limitations presented by the mechanics of the inner hair cells and their synapses, phase locking breaks down for frequencies above 4 to 5 kHz in cat (Johnson 1980, Palmer and Russell 1986). For low frequency tones, phase locking is the foundation for a temporal code for sound frequency.

While the frequency information of the stimulus is coded in AN fibers with either a place code, a temporal code, or a combination of both, information regarding the level of the stimulus may be found in the fiber firing rates. Fibers with high spontaneous

activity have low thresholds and narrow dynamic ranges, but low spontaneous rate (SR) fibers have higher thresholds and are able to provide useful level information over a wider range (Liberman 1978, Evans and Palmer 1980). However every fiber (even with low SR) still has a limited dynamic range compared to that associated with normal perceptual hearing. Information containing the intensity of the stimulus may be coded in fibers with different thresholds at different levels or in spatio-temporal patterns.

Pitch

A *harmonic complex tone* consists of sinusoidal sounds at frequencies which are all integer multiples of a common fundamental frequency (F0), corresponding to the *pitch* of the complex tone. Pitch is a subjective percept of a sound that allows ordering from low to high on a frequency scale. Many natural sounds such as human speech, animal vocalizations, and the sounds of most musical instruments contain harmonic complex tones, which evoke a strong pitch sensation at their fundamental frequency. Pitch plays a functional role in speech, such as for speaker identification (Nolan 2002), portrayal of emotions, and lexical distinctions in tonal languages (Cutler and Chen 1997). Differences in pitch can serve as cues for sound source segregation (Scheffers 1979, Assmann and Summerfield 1990). The pitch evoked by a harmonic complex tone is usually the same as the pitch of a pure tone at F0, even when the complex tone component at F0 is not physically present in the stimulus (Schouten 1940). This phenomenon, known as the "pitch of the missing fundamental," occurs for F0s up to 1400 Hz (Moore 1973b), though more recent data show that the phenomenon may also occur at higher frequencies (Oxenham et al. 2011). In addition to humans, the phenomenon of the

pitch of the missing fundamental has also been shown to occur in cats (Heffner and Whitfield 1976).

Consecutive components of a harmonic complex tone are said to be "resolved" when their frequencies are spaced far enough apart so they fall in the pass bands of different auditory filters. When two or more harmonics fall within the pass band of a single filter, they are said to be "unresolved." Because the bandwidths of the cochlear filters increase with increasing center frequency (Kiang et al. 1965a, Shera et al. 2002), low order harmonics are better resolved than high order harmonics. Psychophysical studies have shown that humans are able to resolve the first 6-10 harmonics (Bernstein and Oxenham 2003, Plomp 1964). A stimulus containing resolved harmonics elicits a stronger pitch percept than one consisting entirely of unresolved harmonics (Carlyon 1998, Bernstein and Oxenham 2005). Pitch perception degrades in both hearing impaired listeners (Bernstein 2006) and cochlear implant users (Moore and Carlyon 2005), prompting the question of what physiological mechanisms account for these degradations. Though pitch phenomena have been studied fairly extensively psychophysically and many theories for its coding have been proposed, the neural mechanisms for its extraction are still poorly understood.

The peripheral auditory system generates essentially two types of cues to the pitch of harmonic complex tones. The first is a spatial cue that depends on the frequency-to-place mapping and frequency selectivity in the cochlea. The second cue generated by the peripheral auditory system is a temporal cue that depends on neural phase locking to the waveform periodicity. Traditional neural models to estimate the pitch of a stimulus have typically relied on either purely spatial or purely temporal representations. However,

neither type of model is adequate to account for all of the observed psychophysical phenomena.

One way to derive pitch is to determine the frequencies of individual components and then select a "best-fitting" F0 from internally stored harmonic templates. The rate-place profile of AN fibers (the discharge rate plotted as a function of CF) has local maxima at frequencies corresponding to resolved harmonics. Thus, an advantage of rate-place models is that they depend on harmonic resolvability for pitch estimates. However, at high stimulus levels, these models predict a degraded pitch due to firing rate saturation (Cedolin and Delgutte 2010), while pitch discrimination performance remains relatively robust (Bernstein and Oxenham 2005). Furthermore, high CF fibers have relatively high Q-factors (CF/bandwidth) compared to low CF fibers, such that rate-place models predict better pitch strength at higher CFs, which is inconsistent with the traditional upper limit of the "pitch of the missing fundamental."

Models based on temporal representations predict robust pitch with level and are also able to predict the upper frequency limit of the "pitch of the missing fundamental" due to the degradation of phase locking at higher CFs. Despite these consistencies with psychophysical studies, weaknesses of the temporal models include failing to predict stronger pitch for resolved harmonics and underestimating the pitch strength of pure tones. Moreover, the temporal models require long neural delay lines or intrinsic oscillators, for which there is little physiological evidence. Finally, users of cochlear implants, who have poor frequency selectivity and likely receive mostly temporal cues to pitch, tend to perform poorly with pitch tasks.

The marmoset auditory cortex contains neurons near the anterolateral border of A1 that respond selectively to pitch stimuli and show responses consistent with the psychophysical properties of pitch (Bendor and Wang 2005, 2010). This location is consistent with a pitch selective area found in humans (Penagos et al. 2004). Specifically, these so-called "pitch neurons" respond to stimuli that have the same F0 but different spectra, such as pure tones and harmonic complex tones with missing fundamental. In addition, these neurons also show increased discharge rates in response to stimuli with increasing pitch salience. These findings suggest that separate spatial and temporal pitch processing mechanisms may converge in or prior to the auditory cortex.

Convincing biological data are needed to either confirm or refute the spatial or temporal models. Recently Cedolin and Delgutte (2010) gave physiological evidence for a spatio-temporal representation of the pitch of harmonic complex tones with missing F0s in the cat AN. The cochlear traveling wave for a pure tone stimulus slows down considerably at the place of resonance along the basilar membrane and thus has rapid phase accumulation around this location (Pfeiffer and Molnar 1970, Pfeiffer and Kim 1975, Lyon and Shamma 1996, Shamma 1985a). For low-CF, phase locking AN fibers, this phase change is reflected in the relative timing of AN spikes at different cochlear locations and generates a spatio-temporal cue to the frequency of the pure tone. In principle, this cue could be extracted by central neurons that receive inputs from neighboring cochlear locations and are sensitive to the relative timing of these inputs (Shamma 1985b). These central neurons could then provide inputs to a harmonic template mechanism.

Unlike the rate representations for pitch, which degrade with level, the spatio-temporal representation for pitch remained relatively robust with level, which is more consistent with the psychophysical observation that pitch strength does not degrade with increases in level. Moreover, the spatio-temporal representation depends on neural phase locking and is weak for higher CF fibers where phase locking is poor. This is consistent with the existence of an upper limit for the pitch of a complex tone with missing F0. The spatio-temporal representation is thus consistent with key trends in human psychophysics of pitch. By requiring the presence of both phase locking and resolved harmonics, the spatio-temporal representation of pitch in the AN combines the advantages of the two traditional temporal and spectral theories while overcoming some of their limitations (Cedolin and Delgutte 2010). However, it is not known whether higher level processes in the brain are making use of these spatio-temporal cues. It is possible that they are extracted by central neurons sensitive to the relative timing of spikes from AN fibers innervating neighboring regions of the cochlea. A promising site for exploring this idea is in the cochlear nucleus (CN), where some neurons display the desired spectral and temporal characteristics.

Cochlear Nucleus Cell Types

All AN fibers convey information about the acoustic stimulus encoded in their firing patterns directly to neurons in the CN, which is the first site of neural computation along the ascending central auditory pathway. Each of the three subdivisions of the CN (antero-ventral CN (AVCN), postero-ventral CN (PVCN), and dorsal CN (DCN)) contains distinct cell types, differing in morphologies, synaptic inputs, membrane channels, cytochemistry, and output projections. In addition, AN fiber endings in the CN

have a wide variety of sizes and shapes ranging from small bouton endings to large endbulbs (Kane 1973, Ryugo and Sento 1991).

Extensive work has been done to understand the physiological mechanisms and anatomical structures of cells in the CN, and studies have correlated single unit response types with morphological characteristics (Rhode et al. 1983ab, Bourk 1976, Blackburn and Sachs 1989, Young et al. 1988, Osen 1969). A widely used approach to the classification of CN units is based on the shape of the post-stimulus-time histogram (PSTH) in response to short tone bursts at CF (STBCF) (Kiang et al. 1965b, Pfeiffer 1966).

Spherical bushy cells are located in the anterior part of the AVCN (Osen 1969, Cant and Morest 1979). Their AN inputs end in large endbulb synapses (only a few inputs per cell), which result in nearly one-to-one synaptic transmission with short response latencies. Responses to STBCFs of spherical bushy cells typically resemble those of AN fibers and are thus called *primary-like*, characterized by a high discharge rate at the stimulus onset, followed by a gradual decrease to a steady rate during the remainder of the tone burst.

Globular bushy cells predominantly appear in the posterior part of the AVCN and anterior part of the PVCN (Osen 1969, Tolbert and Morest 1982a). On average, 15-23 (Spirou et al. 2005) AN inputs converge onto the soma of each globular bushy cell in the form of small endbulbs (Liberman 1991, Brawer and Morest 1975, Osen 1969, 1970, Smith and Rhode 1987, Tolbert and Morest 1982ab). These cells fire in response to STBCFs with a *primary-like-with-notch* pattern, which is characterized by a high onset peak, followed by a short notch of little to no activity for 0.5 to 2 ms, and finally a

sustained response throughout the remainder of the tone. If globular bushy cells behave as coincidence detectors (CD), the high onset response could indicate a high probability that many of the cells' inputs are firing at the stimulus onset. A refractory period property prohibits action potentials from firing within around 0.7 ms of one another and could cause the notch of no activity following the onset.

Octopus cells are found in the caudal region of the PVCN and receive many small bouton endings from many AN fibers (Kane 1973). These cells are usually associated with an *onset* pattern in response to STBCFs, characterized by a short latency, sharply timed response at the beginning of the stimulus followed by little to no activity throughout the remainder of the tone burst. Subcategories of onset cells exist, and there is likely a continuum between these and primary-like-with-notch units. CD has been proposed to explain the onset responses of octopus cells (Oertel et al. 2000).

Stellate cells are located throughout the CN but tend to be more concentrated in the posterior part of the AVCN and the PVCN (Osen 1969, Rhode et al. 1983b, Smith and Rhode 1989). Stellates receive many AN inputs in the form of small bouton endings terminated at some distance from the soma and are likely to receive spatially distributed inputs from many AN fibers. In contrast with the irregular AN firing pattern, stellate cells respond to STBCFs with *chopping* patterns featuring high discharge regularity (Cant 1981, Tolbert and Morest 1982ab, Rhode et al. 1983b). Choppers can be further categorized as transient choppers, which fire regularly only at the beginning of the tone, and sustained choppers, which fire regularly throughout the duration of the tone burst. These cells might perform temporal integration, in which AN input activity is accumulated across the many inputs and over time (Wu and Oertel 1984). When

threshold is reached, the neuron fires and the membrane potential resets to its resting value (van Gisbergen et al. 1975). This summation has been shown to produce regular output patterns for irregular input patterns (Molnar and Pfeiffer 1968).

Pyramidal or fusiform cells are predominantly found in the DCN (Rhode et al. 1983a). In anesthetized preparations, these cells are associated with *pauser* (onset response followed by long pause (5-10 ms) and gradual buildup of firing) or *build-up* (like pauser but without onset response) patterns in response to STBCFs. Pyramidal cells have inhibitory as well as excitatory inputs (Zhang and Oertel 1993, Young 1984, Oertel and Young 2004). Anesthesia strongly influences the response patterns of DCN cells (Anderson and Young 2004), while there is little such evidence in the VCN (May et al. 1998). The effects of anesthetic state complicate the interpretation of DCN cell responses and make them more difficult to model. Furthermore, these cells tend to poorly phase lock to the fine structure, indicating that they may not be sensitive to spatio-temporal cues in the AN. For these reasons, the target of this thesis is to study responses of VCN cells, so pauser and build-up units were not examined in detail.

While most CN units can be classified into one of these ideal groups, there also exist other response types that do not belong to any of the categories described above. These "unusual" units can display characteristics not found in typical units, such as long latencies or complicated onset responses.

Spatio-Temporal Processing

In order to extract the spatio-temporal pitch cues present in the AN, central neurons would first need to receive inputs from AN fibers tuned to different frequencies. In addition, they would also need to prefer a specific temporal relationship among its inputs,

though the details of this relationship are currently unclear. One possibility is cross-frequency coincidence detection (CD), in which a neuron would fire more when its AN inputs (across different CFs) discharge in synchrony and less when there are spatio-temporal discontinuities.

It is generally accepted that the binaural auditory system uses CD to compare the timing between the two ears to encode interaural time differences, a major cue to sound localization (Jeffress 1948). Models for binaural processing have relied heavily on a CD mechanism (Colburn 1996), and physiological studies provide strong evidence for binaural CD in a variety of mammalian species, including dog (Goldberg and Brown 1969), cat (Yin and Chan 1990, Rose et al. 1966), rabbit (Batra et al. 1997), gerbil (Spitzer and Semple 1995), and kangaroo rat (Moushegian et al. 1975).

While much work has been done to investigate the role of CD in binaural processing, less is known about the possible role of *monaural* CD. Previous work in monaural cross-frequency CD has focused on the coding of sound level. Normal hearing human listeners can discriminate small changes in sound level (~1 dB) over a huge range of 0-120 dB SPL (Florentine et al. 1987). However, the firing rates of AN fibers typically only change over a relatively narrow range (~35 dB). This discrepancy between behavioral performance and neurophysiology is known as the *dynamic range problem* (Evans 1981, Viemeister 1988ab). A mechanism that might at least partially explain this discrepancy is monaural cross-frequency CD, which would extend the neural dynamic range by using the temporal information across AN fibers with different CFs (Carney 1994). This mechanism has been shown to give robust performance across a wide range of sound level that is consistent with psychophysical observations (Heinz et al. 2001).

Some CN neurons have responses consistent with monaural cross-frequency CD, as indicated by sensitivity to changes in stimulus phase (Carney 1990) and enhanced synchronization to low-CF tones (Joris et al. 1994ab).

The aim of this thesis is to assess the sensitivity of CN neurons to the spatio-temporal pattern of their AN inputs and to evaluate the implications of this processing on the coding of pitch. Using standard single-unit electrophysiological recording techniques, we first determine whether individual CN units resemble cross-frequency CDs. We then examine responses of these units to harmonic complex tones to evaluate whether cross-frequency coincidence in the CN leads to extraction of spatio-temporal pitch cues. We find that there is a small subset of CN neurons that have responses consistent with model CD cells. Another subset of CN neurons has robust rate representations of pitch at high levels, though the two subsets are mostly non-overlapping. Thus, our method for determining whether individual units perform a cross-frequency CD mechanism ultimately could not predict robust rate cues in response to harmonic complex tones.

To measure spatio-temporal sensitivity, we use complex all-pass transient stimuli called *Huffman sequences*, designed to systematically manipulate the timing of AN fibers at neighboring cochlear locations without drastically changing their overall discharge rates (Carney 1990). If a CN unit receives inputs from AN fibers across different CFs and is sensitive to the relative timing of these inputs, it should respond differently to Huffman stimuli differing in their phase pattern. In Chapter 2, we compare responses of model AN fibers (Zilany and Bruce 2006) and real fibers recorded from the anesthetized cat to Huffman sequences. As expected, manipulating the phase pattern in the Huffman

sequences results in systematic changes to the AN's spatio-temporal pattern in both model and real fibers while changes in firing rate are minimal. In Chapters 3 and 4, we characterize responses to Huffman sequences and pitch stimuli, respectively, of cross-frequency CD model cells operating on AN activity from varying extents of the tonotopic axis for comparison with responses of single CN units. Using a maximum likelihood approach, we find that primary-like-with-notch (Pri-N) units in the CN have responses to Huffman sequences that are consistent with a cross-frequency CD mechanism operating on AN activity. In response to pitch stimuli, the firing rates of AN fibers tend to saturate at high levels, thus degrading the rate representation of pitch. In contrast, we find robust rate cues to pitch in model CD cells as well as in CN Pri-N and chopper units at high levels. However, it is unclear whether the robust rate cues found in some CN units arise from a CD mechanism because there was no clear correlation between a CN unit's rate robustness and resemblance to model CD cells as determined by responses to Huffman stimuli. It is possible that the CD model used in this work is too simple to capture the complex circuitry in the CN. It is also possible that other mechanisms that extract spatio-temporal cues are present in CN neurons, such as lateral inhibition or spatial and temporal summation over spatially distributed inputs. Regardless of whether a CD mechanism gives rise to the these CN responses, our finding that some CN units maintain robust rate representations of pitch at high levels is relevant for understanding how pitch is encoded in the neural pathways.

CHAPTER 2

Responses of Auditory Nerve Fibers to All-Pass Stimuli with Phase Manipulations

Abstract

The *spatio-temporal* pattern in the auditory nerve (AN), or the temporal pattern of activity across AN fibers tuned to different frequencies, contains cues to important features in sounds, such as pitch, loudness, and spatial location. In order to evaluate whether central neurons extract these spatio-temporal cues, we must first find a stimulus that systematically manipulates the temporal pattern across AN fibers without changing their discharge rates.

Huffman sequences have flat magnitude spectra and a 2π phase transition, and have been previously used to manipulate the relative timing of spikes across neighboring AN fibers (Carney 1990). By comparing responses of model AN fibers (Zilany and Bruce 2006) with those of real fibers to Huffman sequences, we used scaling invariance to infer the true spatio-temporal pattern across an array of fibers with different CFs from the responses of a single fiber to stimuli with varying phase transition frequencies.

We found that a Huffman sequence with sharp phase transition resulted in small early responses when the transition frequency was near the fiber's CF. The early responses increased when the transition frequency was further from CF. For the later response peaks, a sharp transition Huffman stimulus excites more coincidentally across CF than a stimulus with a broad phase transition. This result was consistent across stimulus levels and transition frequencies, and held for both real fibers as well as model predictions. As expected, AN fibers showed little sensitivity to the phase transition bandwidth in that overall rate responses to broad and sharp transition stimuli were similar across the population. Overall, we verified that these stimuli can be used to systematically manipulate the spatio-temporal pattern in the AN without greatly changing their discharge rates.

Introduction

The spatio-temporal pattern in the auditory nerve (AN), i.e. the pattern of AN fiber activity across the tonotopic axis, provides cues to important features in natural sounds. For sounds that contain spectral peaks such as speech formants and harmonics, the cochlear traveling wave slows down at the corresponding points on the basilar membrane, producing rapid phase accumulation at these locations. These rapid phase shifts are represented as steep discontinuities in the spatio-temporal pattern of AN discharges (Shamma 1985b, Cedolin and Delgutte 2010) and may serve as cues to the frequencies of spectral peaks in sounds.

In principle, these phase cues can be extracted by a network of central neurons that receive AN inputs tuned to different characteristic frequencies (CFs) and are sensitive to their relative timing. One possible mechanism to extract these phase cues is a cross-frequency coincidence detector (CD) circuit, where a central neuron would prefer to fire when its AN inputs across CF discharge in synchrony. Our goal is to evaluate whether a cross-frequency CD mechanism exists in cochlear nucleus (CN) neurons by examining their responses to different manipulations of the AN spatio-temporal pattern.

In order to evaluate the spatio-temporal sensitivity of central neurons, we must first find a stimulus that manipulates the spatio-temporal pattern in the AN without drastically varying the pattern of discharge rate across CF. Complex stimuli have previously been used to manipulate the relative timing across low-CF AN fibers (Carney 1990). "Huffman sequences" have a flat magnitude spectrum and a single phase transition at a transition frequency (FT), thus exciting all cochlear regions with nearly the same energy while controlling the frequency location of rapid phase shift (Huffman 1962,

Patterson and Green 1970, Patterson et al. 1969). In theory, varying the slope and frequency of the phase transition should systematically change the spatio-temporal pattern of AN fibers with CFs near FT with minimum changes in firing rates.

Carney (1990) compared responses of CN neurons and AN fibers to Huffman sequences to evaluate spatio-temporal sensitivity in the CN. She argued that a Huffman stimulus with a broader phase transition would excite AN fibers more coincidentally across CF than a stimulus with a sharper phase transition. For a linear time invariant system, the more gradual phase changes in a stimulus with a broad phase transition should introduce less group delay across adjacent frequencies and would produce more cross-CF synchrony at the output. However, we find that while the relative group delay introduced by the Huffman sequences affects the response envelope, the fine structure in the response pattern is affected by phase delays. Because a CD mechanism operates on a short time scale, the degree of find structure synchrony across CF matters much more than envelope synchrony. For these reasons, the fine structure of the AN spatio-temporal response patterns must be considered when interpreting the responses to Huffman stimuli.

In this chapter, we compare responses of model AN fibers (Zilany and Bruce 2006) with real AN fibers to Huffman sequences. Both model fibers and real AN fibers showed similar systematic changes in their spatio-temporal response patterns to changes in the phase transition width of the Huffman sequence. We find that the fine structure in the responses of AN fibers to Huffman sequences are more complicated than previously thought, and depend on relative phase delays. In contrast to Carney's hypothesis, we find that it is the stimulus with the sharper phase transition that excites AN fibers more coincidentally across CF, especially for the late portion of the response. In the early

response, we find that the response to the sharp transition stimulus is small when FT is near CF, and increases as FT is moved further from CF, while the early response to the broad transition stimulus changes less. These results serve as a reference for interpreting responses of CD model neurons and CN neurons (Chapter 3).

Methods

Neurophysiology

The methods for recording from single fibers in the AN of anesthetized cats have been well established (Kiang et al. 1965a). We chose the anesthetized cat as our animal model to facilitate comparison to previous studies. Surgery and preparation are similar to those described by Cedolin and Delgutte (2010). Cats are anesthetized with Dial (75 mg/kg) or Nembutal (40 mg/kg) in urethane, with supplementary doses given to maintain an areflexic state. After removal of the posterior portion of the skull, the cerebellum is retracted to expose the AN. The tympanic bullae and middle-ear cavities are opened to expose the round window. Throughout the experiment the cat is given injections of dexamethasone (0.26 mg/kg) to prevent brain swelling and Ringer's solution (50 mL/day) to prevent dehydration.

The cat is placed on a vibration-isolated table in an electrically-shielded, sound-proof chamber. A silver electrode is positioned near the round window to record the AN compound action potential (CAP) in response to click stimuli in order to assess the condition and stability of cochlear function. At the beginning of each recording session, the CAP threshold to a click was typically near -60 dB re 1V. If the threshold increased by more than 10 dB from this value, the experiment was terminated.

Sound is delivered to the cat's ear through a closed acoustic assembly driven by an electrodynamic speaker (Realistic 40-1377). The acoustic system is calibrated to allow accurate control over the sound pressure level at the tympanic membrane. Stimuli are generated by a 24-bit digital-to-analog converter using a sampling rate of 20 kHz.

Stimuli are digitally filtered to compensate for the transfer characteristics of the acoustic system.

Action potentials are recorded with glass micropipettes filled with 2 M KCl. The electrode is inserted into the dorsal surface of the AN and mechanically advanced with a micropositioner (Kopf 650). The recorded signal is bandpass filtered (1-3 kHz) and input into a software spike detector. The signal is initially sampled at 100 kHz resolution. Then the times when the waveform crosses a user-set threshold are linearly interpolated, and spike times are saved for offline processing.

A click at -40 dBre1V is used to search for single units. Once a fiber is isolated, its frequency tuning curve is measured by a tracking algorithm (Kiang et al. 1970) to determine its CF and threshold at CF. The neuron's spontaneous activity is then measured over 20 seconds. The responses to Huffman sequences are then studied.

Auditory Nerve Model

Many attempts have been made to create computational models to mimic the behavior of AN fibers in the mammalian cochlea. We use one of the latest models in a chain of auditory peripheral models (Carney and Yin 1988, Carney 1993, Zhang et al. 2001, Bruce et al. 2003, Zilany and Bruce 2006, 2007, 2009) that have been built to be consistent with a wide range of physiological data in the cat. The model (Zilany and Bruce 2006) receives as input an arbitrary stimulus waveform, and outputs spike times across AN fibers with different CFs.

A block diagram of the model is shown in Figure 2.1. The model consists of a middle ear filter followed by three parallel pathways representing cochlear processing: component 1 and 2 (C1 and C2) signal paths and a wideband feed-forward control path

with level-dependent gain and bandwidth. The model distinguishes between two forms of excitation (C1 and C2) between basilar membrane motion and inner hair cell potential. C1, which dominates at low levels, contains a narrowband chirp filter that accounts for the instantaneous frequency glides and best frequency shifts observed in physiological data (Tan and Carney 2003). C2 is more broadly tuned, dominates C1 at high levels, and is out of phase with C1.

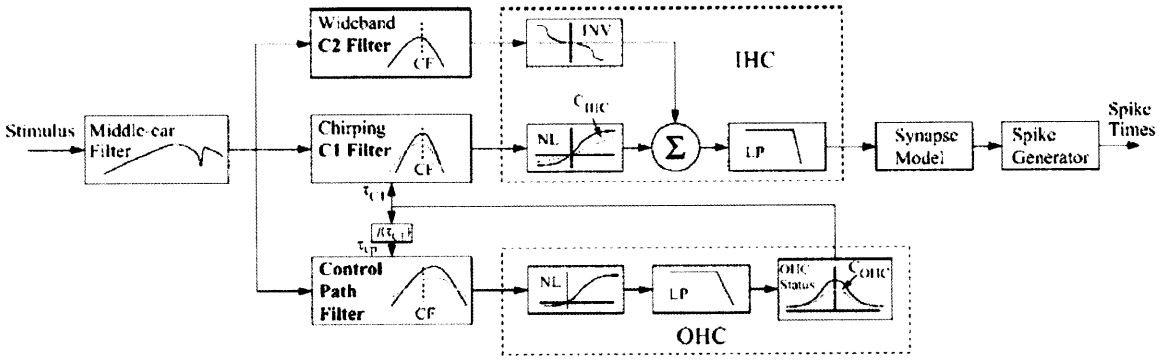


Figure 2.1: Block diagram of auditory nerve model (Zilany and Bruce 2006). See Methods for description.

The interactions between C1 and C2 (within the inner hair cell (IHC) block) simulate a nonlinear phenomenon called peak splitting (Kiang 1990), where AN responses show one peak per cycle for low frequency tones at low levels (phase locking), two or three smaller peaks per cycle at high levels, and at even higher levels, one peak per cycle 180° out of phase with the response at low levels. The control path represents the active cochlear processes and produces two-tone suppression (Ruggero et al. 1992), where the response to a near-CF tone is decreased in the presence of another tone.

The non-linear synapse model is a time-varying three store diffusion model (Westerman and Smith 1988), which sets the synapse gain between the IHC and the AN

fiber. Finally, the output of the synapse drives the spike generator, which produces spike times using a renewal process that includes refractoriness.

Discharge patterns across the population of model AN fibers will be used to test for spatio-temporal manipulation using Huffman sequences, and to interpret physiological AN responses.

Stimuli: Huffman Sequences

"Huffman sequences" (Huffman 1962, Patterson and Green 1970, Carlyon and Shamma 2003) have a flat magnitude spectrum and a single 2π phase transition centered at a transition frequency FT. To generate these stimuli, a unit sample was input into a digital filter with transfer function of the form

$$H(z) = \frac{1 - a_1 z^{-1} + a_2 z^{-2}}{1 - b_1 z^{-1} + b_2 z^{-2}}$$

where

$$a_1 = \frac{2\cos\theta}{r} \;,\; a_2 = \frac{1}{r^2} \;,\; b_1 = 2r\cos\theta \;,\; b_2 = r^2 \;,\; \text{and } \theta = \frac{2\pi FT}{FS}.$$

FT is the transition frequency, and FS is the sampling rate (20 kHz).

Huffman sequences have a set of conjugate pair poles within the unit circle and a set of corresponding conjugate pair zeros outside the unit circle. The parameter "r" is the distance from the origin to either pole and is inversely related to the phase transition bandwidth. Moving the poles and zeros closer to the unit circle results in a sharper phase transition and a longer group delay. This is illustrated in Figure 2.2A, which shows the magnitude and phase spectra of Huffman stimuli with sampling frequency 20 kHz and FT 1 kHz, and shows that a stimulus with a greater r-value (r=0.95, in black) has a sharper

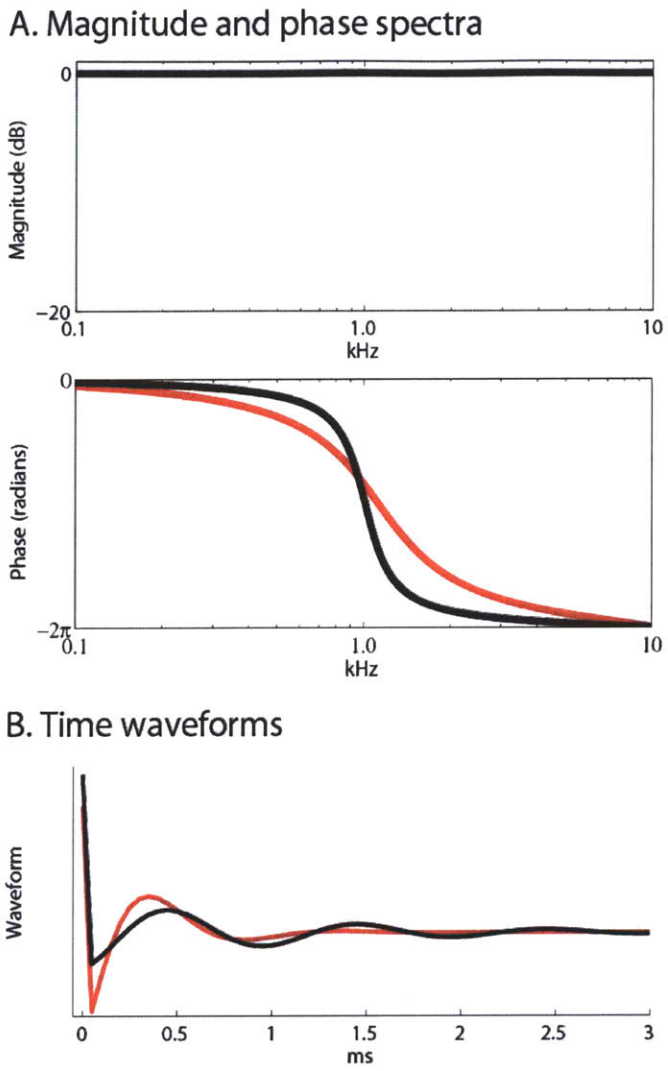


Figure 2.2: Huffman sequences with FT 1000 Hz and r-values 0.85 (red) and 0.95 (black). **A.** Constant magnitude spectrum and single phase shift centered at FT with variable phase transition bandwidth controlled by the parameter "r". Red: r = 0.85, shallow phase transition. Black: r = 0.95, sharp phase transition. **B.** Time waveforms of Huffman stimuli.

phase transition than the stimulus with a smaller r-value (r=0.85, in red). The temporal waveforms of these two Huffman sequences are shown in Figure 2.2B.

Over the range of FT investigated, the -π /2 to -3π/2 phase transition bandwidths are 1017 Hz (when r = 0.85, red) or 326 Hz (when r = 0.95, black). We also used an r-value of 0.90, which had a phase transition bandwidth of 666 Hz (not shown). Carney (1990) used smaller FS (6-10 kHz) and also co-varied FS and FT to keep the number of samples per waveform fixed at 100. As a result, her stimuli had smaller bandwidths than ours for a given r (503 Hz vs.1017 Hz when r=0.85, 330 Hz vs.666 Hz when r=0.90, and

161 Hz vs.326 Hz when r=0.95). To facilitate comparison, we found the r-values that match her phase transition bandwidths for a 20 kHz stimulus (r=0.92, 0.95, and 0.98) and used these r-values in a subset of our fibers (28/143).

Our Huffman stimuli decayed to nearly zero within a few milliseconds, but at least a 30 ms inter-stimulus interval was used to ensure recovery between responses to successive stimuli. Stimulus level is defined in dB re the rms sound pressure over the 30 ms interval, thus underestimating the effective stimulus level in the temporal region where the stimulus has significant energy. If, rather than defining level over a fixed time interval, we had defined the level over the time interval that encompasses 99% of the energy of the stimulus waveform, this raised the stimulus level by roughly 15 dB.

For fibers with low spontaneous rates, the threshold level was found by listening for an audible response just above spontaneous activity from the electrode signal. For fibers with high spontaneous rates, a quick rate-level function was sometimes measured to find the threshold level. The stimulus levels were chosen to be near threshold and increased in 10 or 15 dB increments. At least 3 stimulus levels were presented for each fiber. Level was interleaved randomly across trials to minimize any effect of the order of sound levels in the stimulus presentation.

Each 300 ms trial consisted of a sequence of either 6 stimuli (3 ascending and 3 descending r-values (0.85, 0.9, 0.95, 0.95, 0.9, 0.85) with 50 ms inter-stimulus interval, or 10 stimuli (0.85, 0.9, 0.92, 0.95, 0.98, 0.98, 0.95, 0.92, 0.9, 0.85) with 30 ms inter-stimulus interval. Spike times from the same r-value stimulus from the ascending and descending parts of the sequence were combined. This was done to minimize any possible effect of the order in which the stimuli were presented, though we observed no

systematic differences between the ascending and descending parts. Responses to these sequences were recorded for at least 100 trials (and typically between 200-300 trials) with no interruption.

In all fibers (n=143), we measured responses to a Huffman stimulus in which FT was set at the CF of the fiber. In addition, for a subset of fibers (44/143), FT was varied from a half octave below CF to a half octave above CF in 1/6 octave steps. We used scaling invariance to infer responses of a population of AN fibers to a Huffman stimulus with fixed FT from responses of a single fiber to stimuli with varying FT (see Results for a detailed explanation).

Analysis

We compared "virtual" spatio-temporal patterns (from single real fiber responses to varying FT Huffman stimuli) with model virtual spatio-temporal patterns and true spatio-temporal patterns (from model fiber responses across CF to a fixed FT stimulus). Post stimulus time histograms (PSTHs) for each CF (or FT) were constructed (with bin width 0.1 ms) to build a (virtual) spatio-temporal pattern. Response peaks were found by detecting all local maxima in each PSTH that exceeded 5 standard deviations above baseline activity measured during the last 10 ms of each inter-stimulus interval (long after both stimulus and response had decayed). Examples of virtual spatio-temporal response patterns and their peaks are shown in Figures 2.7 and 2.9. These virtual spatio-temporal response patterns have had the non-scaling cochlear delay corresponding to the location of the CF (same delay in seconds across y-axes in Figures 2.7 and 2.9) subtracted from their x-axes. Then, the non-scaling cochlear delays corresponding to the locations of the

virtual CFs were added to each x-axis (different delay in seconds across virtual-CF axis in Figure 2.11C).

Sets of peaks across CF (or FT) were grouped together based on a matching algorithm that finds the most likely alignment, accounting for temporal offsets and warping. This is similar to the image registration problem in computer vision (Zitova and Flusser 2003). Specifically, for each pair of adjacent CFs (or FTs), the peak times for one were held fixed, while the peak times for the other were iterated over to find the most likely matching pairs between CF (or FT). This process was repeated until all adjacent pairs were aligned, and peaks were grouped into sets across CF (or FT). Peak sets were discarded if there were fewer than 3 peaks in the set, or if there was no peak in the set when FT was at CF, for either the broad or sharp phase transition stimulus. For real fiber responses, each peak set across FT was then assigned a number corresponding to the nearest model peak in time.

For each peak set across CF (or FT), line segments (at most 3 segments) were fit (least mean squares) to the peak locations. The slopes of these line segments (in frequency/time) quantify the degree of coincidence across CF (or FT) that each stimulus elicited (Figure 2.11D). The slopes were then compared for different r-values to evaluate the relative degree of coincidence for each phase transition width (Figure 2.12).

Results

Model AN fiber responses to Huffman stimuli

General characteristics

To test whether changing the slope of the phase transition of Huffman sequences manipulated the spatio-temporal pattern of discharges of the AN, we used an AN model (Zilany and Bruce 2006) to simulate the discharge patterns of a population of AN fibers. Figure 2.3 plots the response of an array of model fibers with varying CFs (500 to 2000 Hz with 1/25 octave spacing) to two Huffman sequences with different phase transition bandwidths: broad transition (r = 0.85) in panel A and sharp transition (r = 0.95) in panel B. The transition frequency (FT) was set to 1000 Hz, and the stimulus level was 60 dB SPL. In each plot, pixel intensity indicates the model neural response, i.e. the output of the inner hair cell synapse. The non-scaling cochlear delay (here defined as the time to the first peak in response to a 50 dB SPL condensation click stimulus) was subtracted for each CF (a different delay for each value on the y-axis) to facilitate comparison with normalized virtual spatio-temporal patterns described later.

The earliest peaks in response to Huffman sequences for both broad and sharp phase transitions line up nearly perfectly at time zero. Since the cochlear delay has been subtracted in the plot, this means that the first response peak to Huffman stimuli lines up with the first response peak to a click stimulus. Subsequent peaks in firing probability tend to occur at 1/CF intervals, as expected for a broadband stimulus such as a Huffman sequences. For example, the response patterns show 2 peaks per ms when CF is 2 kHz

and 1 peak every 2 ms when CF is 500 Hz. However, there are some slight deviations from this, especially in the response to the sharp transition stimulus.

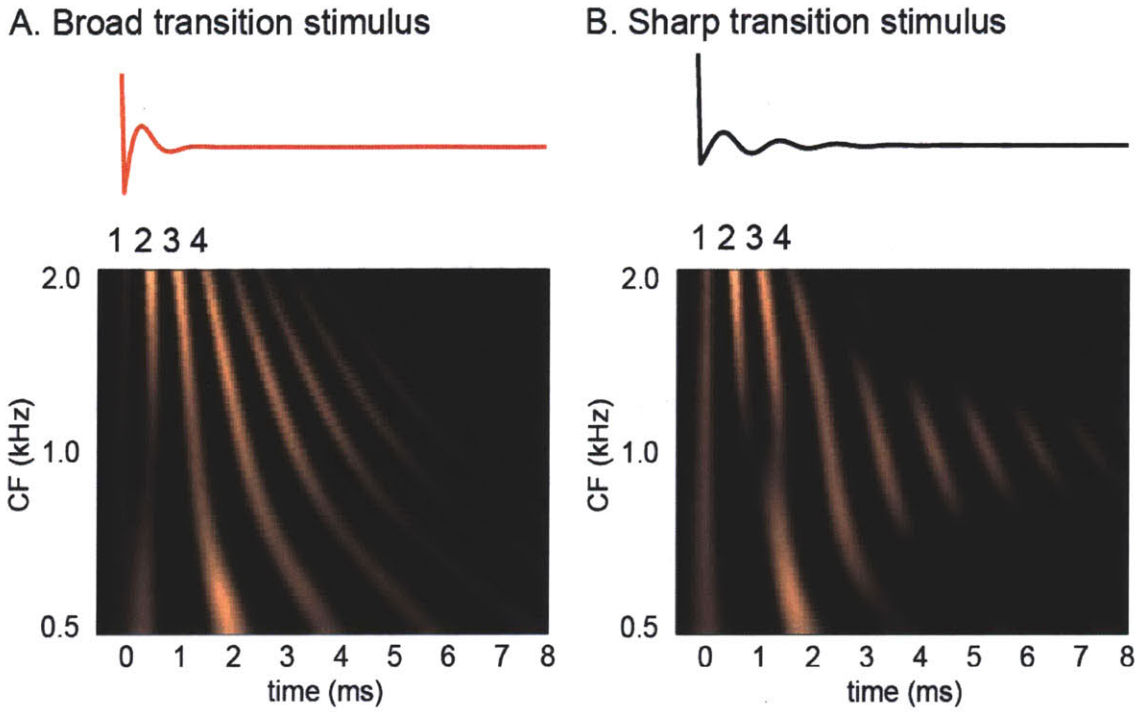


Figure 2.3: Array of model AN responses to two Huffman sequences with FT 1000 Hz at 60 dB SPL. The y-axis is the CF of the fibers, and the x-axis is time in ms. **A.** Broad phase transition (r = 0.85). **B.** Sharp phase transition (r = 0.95).

Effect of transition bandwidth

To quantify changes in the response patterns to the two transition widths, the patterns such as in Figure 2.3 were first reduced to their local maxima. Each panel in Figure 2.4 shows these maxima plotted together for the broad transition width in red and the sharp transition width in black. The relative area of each dot corresponds to the height of a local maximum in the response. Figure 2.4B shows the local maxima taken from the patterns in Figure 2.3 (at 60 dB SPL), while panel A shows the pattern at a lower stimulus level (40 dB SPL), and panel C shows the pattern at a higher level (80 dB SPL).

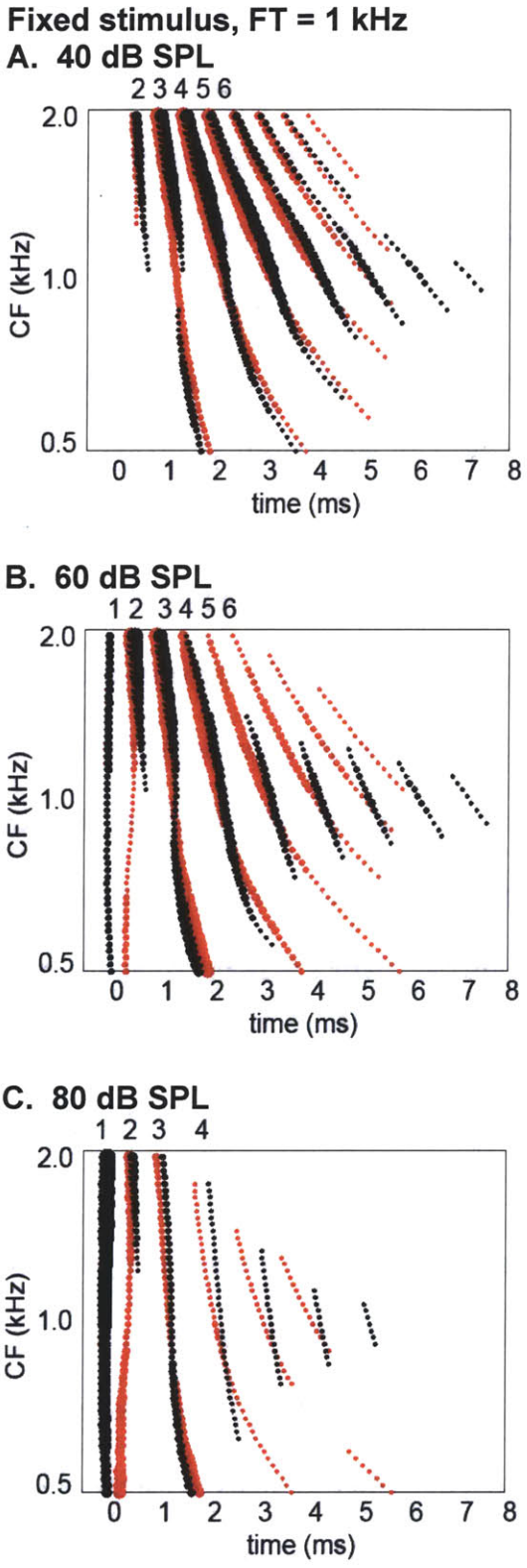


Figure 2.4: **A.** Array of model AN responses to two Huffman sequences with fixed FT of 1000 Hz (broad phase transition in red, sharp phase transition in black). The y-axis is the CF and ranges from 0.5 (one octave below FT) to 2 kHz (one octave above FT). The x-axis is time normalized by FT. Level is 40 dB SPL. **B.** Same as in A, but at 60 dB SPL. **C.** Same as in A, but at 80 dB SPL.

As level increases, the earlier peak heights increase while later peak heights decrease. For example, Peak 1 is not observed at 40 dB, has roughly the same size as later peaks at 60 dB, and is very prominent at 80 dB. On the other hand, Peak 4 is prominent at 40 and 60 dB, but is smaller than earlier peaks at 80 dB. Furthermore, the peak alignment across CF for *both* phase transition widths gets greater at higher levels, consistent with the widening of the cochlear filters.

Despite this level dependence, the pattern of differences in the relative response for the two transition widths remains the same across all three levels in Figure 2.4A-C. Peak 1, when it appears, is nearly vertical for both sharp and broad phase transitions, indicating nearly perfect cross-frequency coincidence. Peak 2 diverges for the two transition widths. As CF decreases away from 2 kHz, the red peaks (broader transition) occur earlier while the black peaks (sharper transition) occur later. Furthermore, the Peak 2 amplitude decreases for the sharp transition stimulus as CF decreases, and the peak ultimately disappears when the CF is below the FT of 1 kHz. For the broad transition stimulus, Peak 2 displays similar behavior as for the sharp transition stimulus at the lowest level, but at 60 dB, the broad transition peak does not disappear when CF is below FT. At 80 dB, the broad transition Peak 2 height very slightly decreases when CF is near FT, but increases again when CF is below FT.

For Peak 3, at 40 and 60 dB, the broad transition response remains fairly constant across CF, while the sharp transition response is larger when CF is away from FT and smaller when CF is near FT. The region of CFs for which the Peak 3 height to the sharp transition stimulus is smaller than that to the broad transition stimulus is narrower at 60 dB than it is at 40 dB. At 80 dB, the peak heights for both transition widths are roughly

constant across CF. The peak heights of Peaks 4 and later tend to be largest when CF is near FT. The heights of these late peaks when CF is away from FT for the sharp transition stimulus tend to be smaller than for the broad transition stimulus.

Peaks 3 and later all have approximately the same temporal relationship between the two transition widths across all levels. When CF is above FT, responses to the sharp transition stimulus occur slightly later than responses to the broad stimulus. As shown in Figure 2.2, for frequencies above FT, the sharp transition stimulus has a more negative phase than the broad transition stimulus, resulting in an increased lag for the sharper phase transition at high frequencies. This trend reverses when CF is below FT, when sharp transition responses occur slightly earlier than broad responses, as a result of the decreased lag introduced with the sharper phase transition at low frequencies. Responses to the two stimuli only line up in time (have the same lag) when the FT is near the CF. As a result, the stimulus with sharp transition bandwidth excites more coincidentally across CF than the broad transition stimulus. These results are consistent with model AN responses in Carney's Figure 2 (1990) obtained from a simpler model (Carney and Yin 1988) that used linear cochlear filters (from revcor analysis) rather than the 3 path nonlinear model of Zilany and Bruce (2006).

Varying the r-value between and beyond the two values shown here (0.9, 0.92, and 0.98) also resulted in systematic and monotonic changes to the spatio-temporal pattern of the AN model (not shown).

Cochlear Scaling Invariance

To physiologically measure the response pattern to Huffman stimuli across an array of real AN fibers would be experimentally challenging because it would require finely and

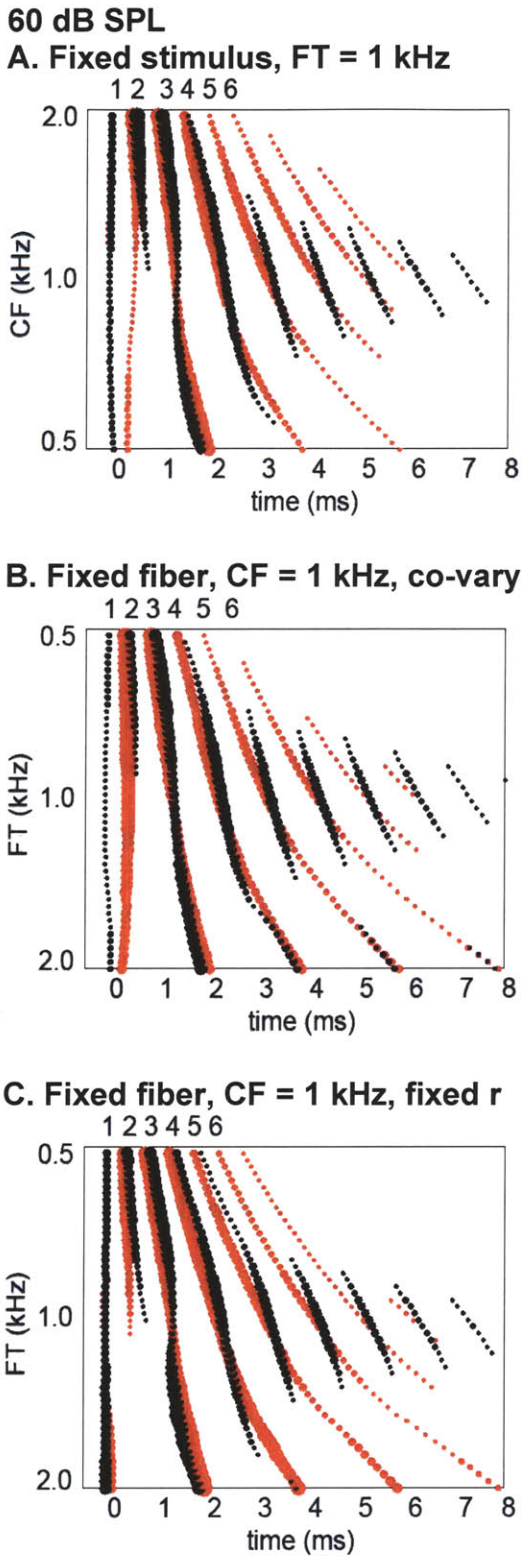


Figure 2.5: A. Array of model AN responses to two Huffman sequences with fixed FT of 1000 Hz (broad phase transition in red, sharp phase transition in black). The y-axis is the CF and ranges from 0.5 (one octave below FT) to 2 kHz (one octave above FT). The x-axis is time normalized by FT. Level is 40 dB SPL. **B.** Responses of a single model fiber with CF 1 kHz to Huffman stimuli scaled in time such that the r-value and FT co-vary. FT is shown on the y-axis and inverted. **C.** Responses of a single model fiber with CF 1 kHz to Huffman stimuli scaled such that the r-value is fixed, and FT varies.

uniformly sampling from AN fibers with differing CF. To get around this challenge, we apply the principle of local scaling invariance in cochlear mechanics (Zweig 1976), which states that in a completely invariant cochlea, the basilar membrane impulse response $h(t, CF)$ at location CF and time t is the same as $h(t\beta, \frac{CF}{\beta})$, where β is an arbitrary constant. This means that the shape of the impulse response h at location CF is the same as a time-scaled (by β) version of the response at a different cochlear location $\frac{CF}{\beta}$.

We define β as the ratio between the fiber CF and the FT of the Huffman stimulus. It then follows that the basilar membrane response at the cochlear location corresponding to the fiber's CF has the same shape as the basilar membrane response at the location corresponding to the stimulus FT. Furthermore, the two responses are scaled (in time by β) versions of each other. Therefore, scaling the Huffman stimulus in time and recording the response at a fixed location allows us to infer the response to a fixed Huffman stimulus at differing cochlear locations. However, this principle only holds locally (Shera and Guinan 2003, van der Heijden and Joris 2006), and FT should be near CF in order to apply scaling invariance.

In order to scale the Huffman stimuli in time, both the FT and the r-value (which sets the decay of the stimulus envelope) must be co-varied. For simplicity we only varied FT while keeping the r-value fixed in our recordings from real AN fibers. We use the AN model to address differences in the two scaling methods.

In Figure 2.5, we compare the response pattern of an array of model fibers (with different CFs) responding to a fixed FT stimulus (panel A, repeated from Figure 2.4B)

with the response patterns of a single model fiber responding to multiple Huffman stimuli. Panel B shows responses when the r-value and FT are co-varied, and panel C shows responses when only FT is varied. The y-axes in panels B and C have been inverted such that the normalized frequency (CF/FT) is the same as in panel A.

The patterns in all panels are very similar, and most of the description in the previous section for panel A (also Figure 2.4B) also describe panels B and C. Importantly, the relative changes in the response patterns to the different transition widths across all panels are the same, in that the earlier response peaks are more synchronous than the later peaks, and that the sharp transition stimulus excites more coincidentally across CF and FT for Peaks 3 and later.

However, there are some minor differences to note between the patterns. Peak 2 in panel A for the broad transition stimulus is small in height and is large in panel B. Furthermore, Peak 2 appears when CF is below FT in panels A and B but disappears completely in panel C for the same normalized frequency range. Furthermore, there are some differences in the way the peak height changes along the y-axis. For example, the height of Peak 1 decreases at lower CFs in panel A, but increases at higher FTs in panel C. Despite these differences, the nearly identical patterns across all three panels support the use of scaling invariance to infer the response of an array of fibers to a single Huffman stimulus from the response of a single fiber to Huffman stimuli scaled in time by either co-varying the r-value and FT, or by only varying FT.

To quantify the similarity between these patterns, we computed correlation coefficients between the two response patterns over a certain range of CF/FT. Correlations are shown as a function of frequency range (in octaves) in Figure 2.6 for

different stimulus levels. Not surprisingly when scaling invariance was rigorously applied (by co-varying the r-value and FT, shown in open circles), the correlation coefficients were better (ranging between 0.97 and 0.99) than when only the FT was varied (ranging between 0.94 and 0.99, shown in x symbols). The patterns are less correlated when a wider range of frequencies is considered, suggesting that the patterns

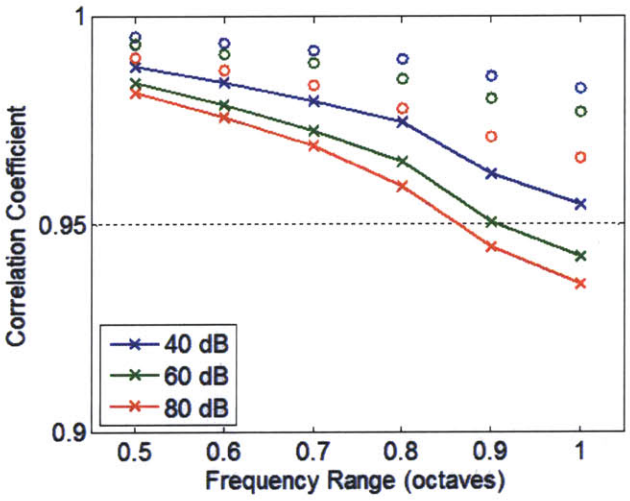


Figure 2.6: Correlation coefficients between model patterns for: x = varying CF (FT = 1000 Hz) and varying FT (CF = 1000 Hz, r fixed), and o = varying CF (FT = 1000 Hz) and co-varying FT and r (CF = 1000 Hz).

look less similar far from 1000 Hz than they do near 1000 Hz. This is consistent with the notion of the locality of scaling invariance, or that scaling invariance is only valid over a relatively narrow range of frequencies. Furthermore, the patterns are somewhat less correlated at high levels, possibly due to cochlear nonlinearities that may have different effects across the length of the cochlea at high sound levels. Despite these changes, the correlation between the patterns for both methods of scaling are high over the frequency range considered, justifying our approach in our experiments. The correlations were similarly high when the fixed frequency (CF for single fiber, or FT for fiber array) was varied from 250 Hz to 2500 Hz (not shown).

Physiological AN fiber responses to Huffman Sequences

We limited our analysis to low CF fibers within the range of phase locking, such that the relative timing of discharges across CF could be manipulated. We recorded from 143 AN fibers (with CFs 107 – 2565 Hz) in 11 cats. Of these, 92 fibers had high spontaneous rate (SR, >18 spikes/s), 43 had medium SR (0.5 to 18 spikes/s), and 8 had low SR (<0.5 spikes/s). In all fibers, we recorded the response to Huffman stimuli with FT set to the CF for at least 3 phase transition widths (r = 0.85, 0.90, 0.95). In a subset of our fibers (44/143), we varied the FT from a half octave below to a half octave above the fiber's CF in 1/6 octave steps. To facilitate comparison to Carney's (1990) study, we also recorded responses of 28 fibers to Huffman stimuli with r-values adjusted to match the bandwidths of her stimuli (0.92 and 0.98).

In this section, we will first describe an example fiber's responses when FT is varied, and verify that these responses match with those of a model fiber with the same parameters. Then we will show population responses of fibers (from one cat) when the FT is varied above and below CF, and again verify the results with model fiber responses. Finally, we use a slope metric to quantify the degree of peak alignment across FT, which we interpret as cross-CF coincidence.

Example responses when the transition frequency is varied

As described earlier, cochlear scaling invariance allows us to approximately infer the response of an array of fibers with different CFs to a stimulus with a single FT from the responses of a single fiber to varying FT stimuli. Post stimulus time histograms (PSTHs) of an AN fiber are shown (CF 532 in Figures 2.7AC) to Huffman sequences in red (broad

transition) and black (sharp transition) with FT on the y-axis and time in cycles (normalized by the FT and cochlear delay, or the time to the first peak in response to a click stimulus, subtracted) on the x-axis. FT was varied from a half octave below CF to a half octave above CF in 1/6 octave steps. Panels A and C show the raw response PSTHs (at 80 dB SPL and 100 dB SPL respectively), while panels B and D plot dots corresponding to peaks in the PSTHs, where the relative size of the dot corresponds to the relative height of the peak. Line segments were fit to sets of peaks across FT (see methods) and are plotted as dotted lines. The numbers above the plots in Figure 2.7 correspond to the model peaks that most closely match in time.

For comparison, responses of a 532 Hz model fiber to the same stimuli are shown in Figure 2.8. The model levels used were below the stimulus level used for the real fibers because this was found to provide a better match (see Discussion). The physiological response in Figure 2.7 has temporally sharper peaks compared to the model responses in Figure 2.8. This was typical for low CF fibers (CF<800 Hz). In higher CF fibers, the peak widths are better matched between physiological and model fibers (see Figure 2.9). The model fiber responses also tended to last longer, ringing for several more cycles than observed in real AN fibers. This was possibly a consequence of using lower stimulus levels in the model. However, increasing the model level above 80 dB SPL resulted in the peak splitting phenomenon (described in Methods, Kiang 1990), which was never observed in real fibers, even at the highest stimulus levels tested (110 dB SPL).

Despite the poor match in peak width for this low CF fiber, the overall pattern is similar at both levels between the model and real fibers. Peak 1, when it occurs for

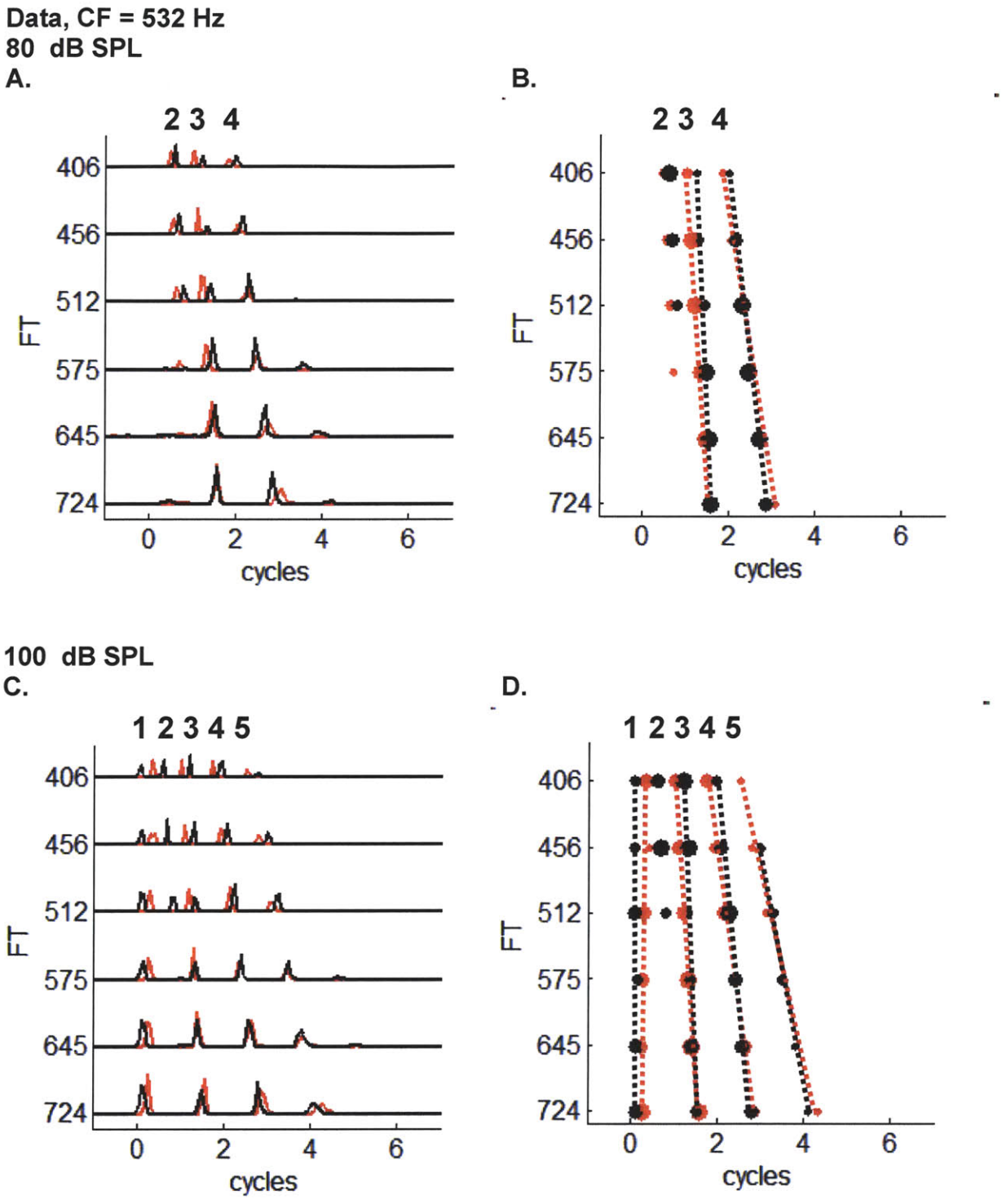


Figure 2.7: Responses of an AN fiber with CF 532 to Huffman stimuli with varying FT on the y-axis and time normalized by FT on the x-axis. Responses to stimuli with broad phase transition in red, sharp phase transition in black. **A.** Real fiber response PSTHs at 80 dB SPL. **B.** Peak locations of response from panel A. **C.** PSTHs at 100 dB SPL. **D.** Peak locations of response from panel C.

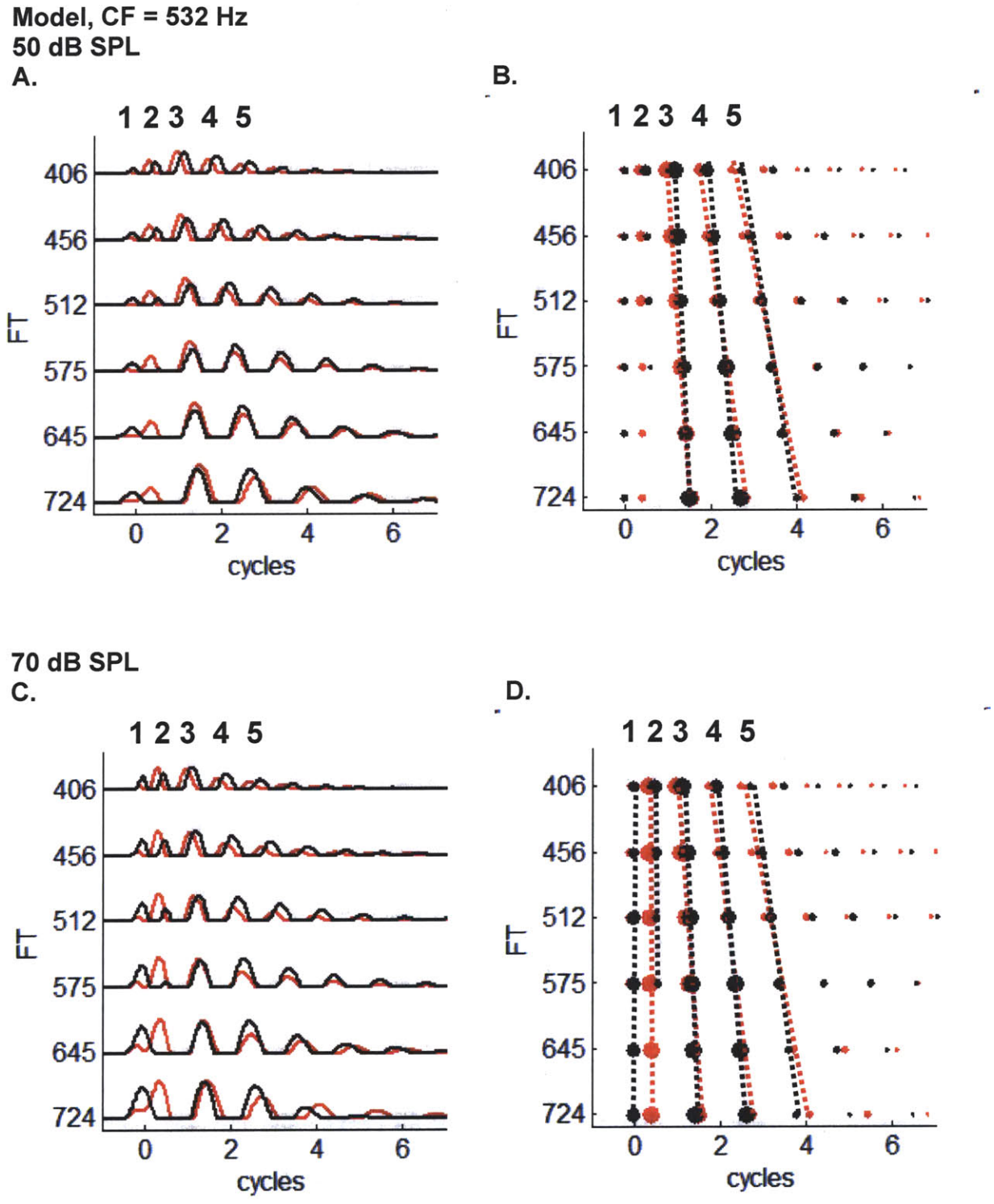


Figure 2.8: Responses of a model AN fiber with CF 532 Hz, responding to Huffman stimuli with varying FT on the y-axis and time normalized by FT on the x-axis. **A.** Response PSTHs at 50 dB SPL. **B.** Peak locations of response from panel A. **C.** PSTHs at 70 dB SPL. **D.** Peak locations of response from panel C.

the higher level, in response to the sharp transition stimulus is nearly vertical, and occurs earlier at higher FTs for the broad transition stimulus. For both model and real fibers, Peak 2 at both levels for the sharp transition stimulus is only observed when FT is below CF, with the peak height decreasing when FT is near CF. A similar result is observed for the real fiber's response to the broad transition stimulus at the lower level, though a Peak 2 response is observed when FT is above CF (for FT=575 Hz). At the higher level, we did not observe a Peak 2 in the real fiber's response to the broad transition stimulus. The model response for Peak 2 differs from the data in that Peak 2 for the broad transition stimulus occurs slightly earlier as FT increases, and the Peak 2 height does not decrease with increasing FT as much as observed in the data. These observations are consistent with the real fiber's Peak 1 at the higher level, which occurs earlier as FT increases for the broad transition stimulus.

For all later peaks, the temporal relationship between the two transition widths for both model and data are the same. When FT is below CF, the responses to the sharp transition stimulus occur slightly later than the broad responses. When FT is above CF, this trend reverses direction, causing the responses to be overall more coincident across FT for the sharp transition stimulus.

Figure 2.9 shows the responses of another fiber with CF 938 Hz to Huffman stimuli at 75 dB (panels AB) and 85 dB (panels CD). The late peaks in these responses are similar to those of the lower-CF fiber in that for Peaks 3 and later, the sharp transition stimulus excites more coincidentally across FT than the broad transition stimulus. We observe some differences between the two fibers (in Figures 2.7 and 2.9) in the early response (Peaks 1 and 2). For the low-CF fiber, Peak 2 to the broad transition stimulus

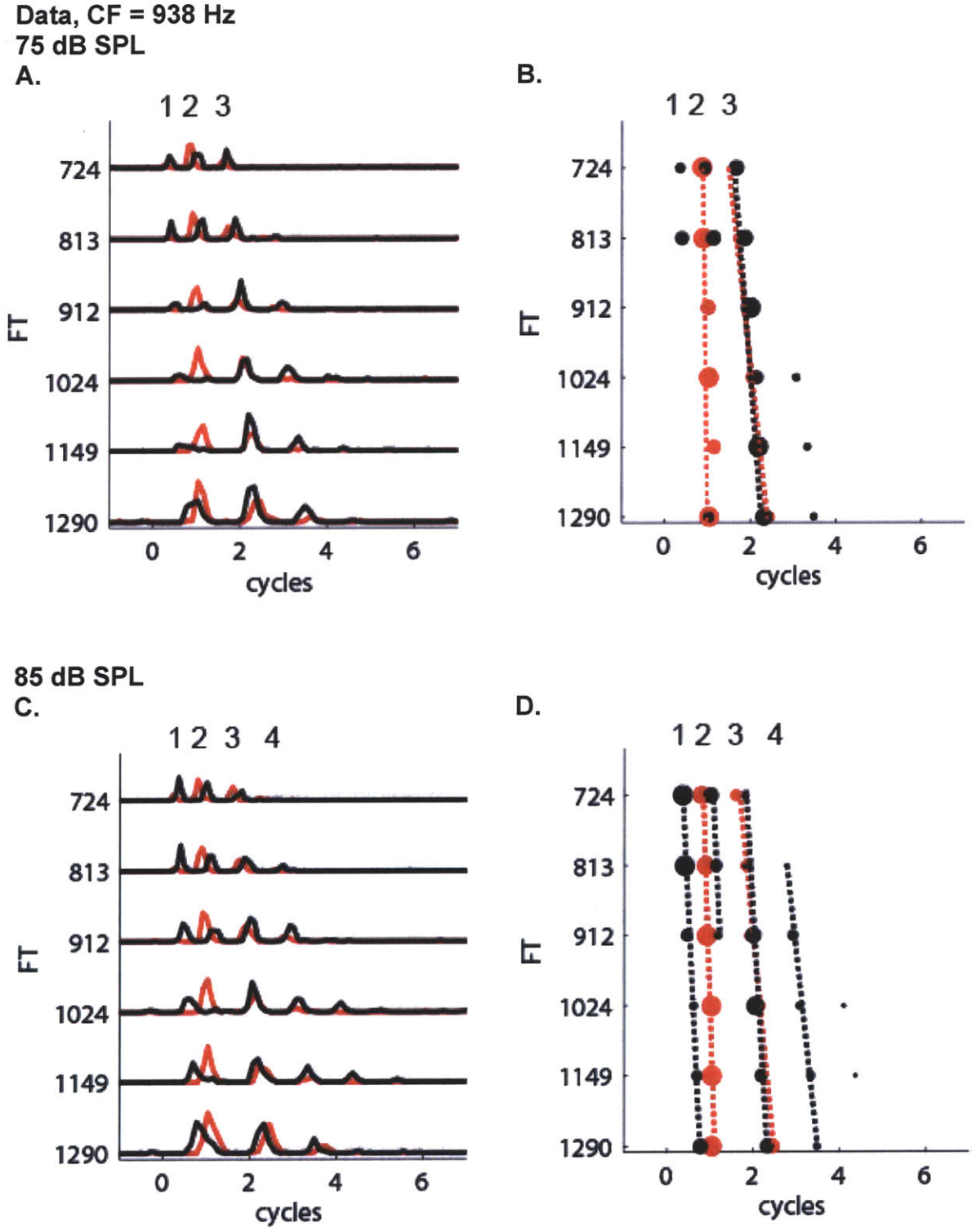


Figure 2.9: Responses of an AN fiber with CF 938. As in figure 2.6.

either decreases in height as FT increases (lower level) or does not occur (higher level). In contrast, for the higher-CF fiber, the height of Peak 2 to the broad transition stimulus remains roughly constant across all FT at both levels. Furthermore, the low-CF Peak 1 height remains constant across all FT when it appears at the higher level for both transition widths. In contrast, for the higher-CF fiber, the Peak 1 response to the sharp transition stimulus is smaller when FT is near CF, and is larger when FT is away from CF. This effect was observed in many fibers in our population across all CFs, and tended to be stronger at lower levels compared to higher levels.

Population responses when the transition frequency is varied

Figure 2.10 plots the peaks of responses (local maxima in the response PSTHs at least 5 standard deviations above spontaneous activity) of AN fibers from one cat to Huffman stimuli for CF>FT ($1.2 < CF/FT < 1.8$, mean 1.4, panel A), CF=FT (panel B), and CF<FT ($0.5 < CF/FT < 0.9$, mean 0.8, panel C). In each panel, CF is on the y-axis, and time normalized by 1/CF is on the x-axis. Physiological AN fiber response peaks are plotted as solid dots while model AN fiber peaks are plotted as open squares. As before, broad transition stimulus responses are shown in red, while sharp transition stimulus responses are shown in black. Because the peak times did not change much with level (earlier peaks appeared as level increased, but increasing level did not change the times of the later peaks), data for all stimulus levels recorded are plotted together (50 to 110 dB SPL, mean 71 dB SPL). Model level was fixed at 70 dB SPL, and the model CF/FT was fixed to the mean in the data (1.4 in panel A, 1.0 in panel B, 0.8 in panel C).

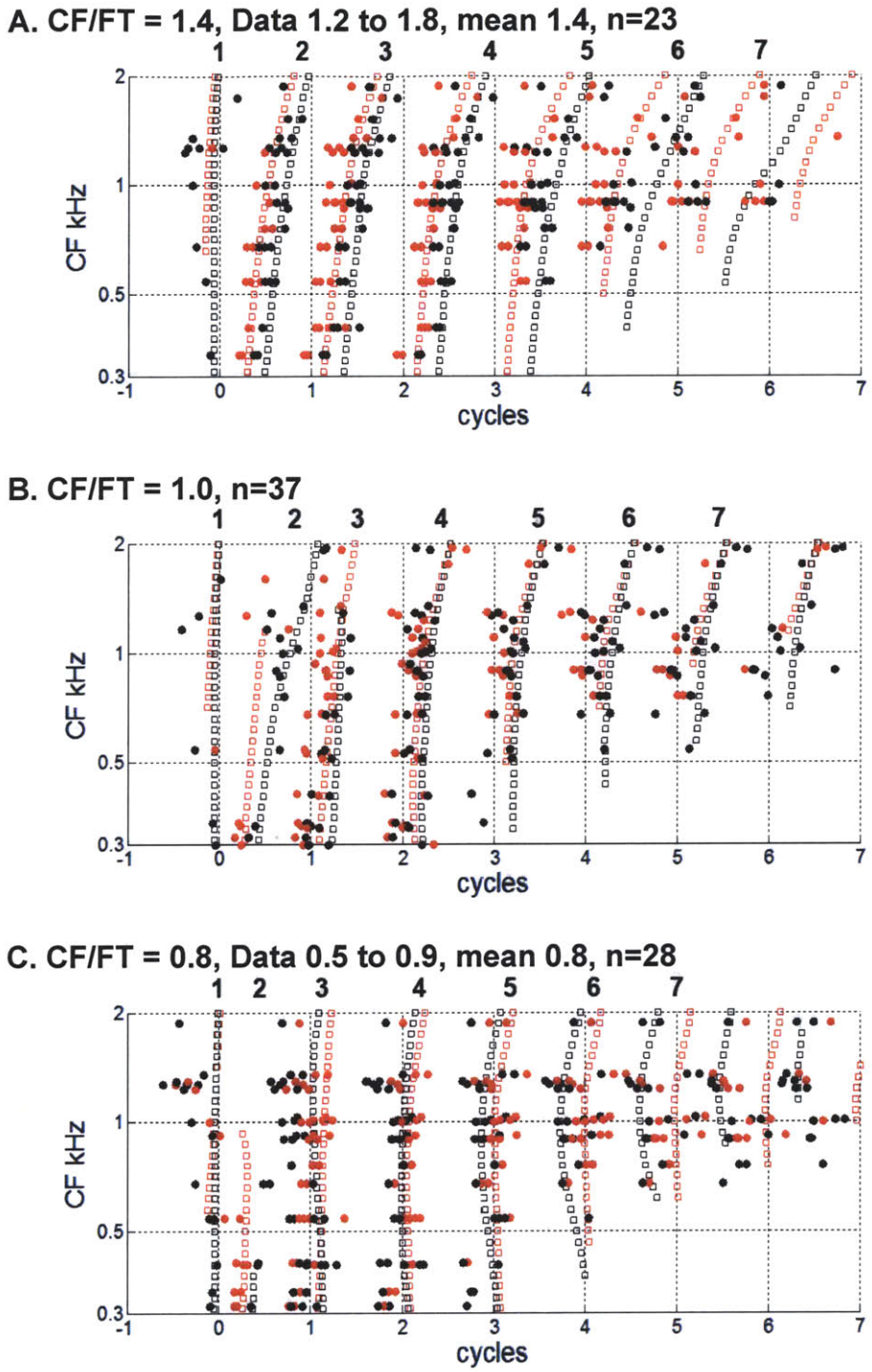


Figure 2.10: Response peaks of AN fibers (n=37) to Huffman stimuli from one cat across all stimulus levels. CF is on the y-axis and time normalized by CF on the x-axis. Peaks of fiber responses are plotted in red (broad phase transition) and black (sharp phase transition) symbols. Real fiber responses are plotted in solid dots, and model fiber responses (computed at 70 dB SPL) are plotted in open squares. **A.** CF/FT = 1.4, **B.** CF/FT = 1.0, **C.** CF/FT = 0.8.

Though there is variability among the fiber responses, the peak times are roughly consistent with those predicted by the model. The population response is also consistent with trends observed in the examples in Figures 2.7 and 2.9. In both the physiological and model patterns, when CF is above FT (Figure 2.10A), for Peaks 2 and later, the peaks in response to the broad transition stimulus (red) tend to occur about 0.2 cycles earlier than the peaks in response to the sharp transition stimulus (black). When CF is set equal to the FT (panel B), the responses to the two transition widths are closer in time. When CF falls below FT (panel C), the response to the sharp transition stimulus occurs slightly earlier than for the broad transition stimulus for Peaks 3 and later.

Furthermore, the real fibers rarely display a peak at time 0 (corresponding to model Peak 1), and typically only do so at the highest recorded levels (>90 dB SPL), where responses have increased onset responses and shorter latencies. Real fibers with CF below 1 kHz tended to display their earliest peak at times corresponding to the model Peak 2 or 3. Peak times of fibers with CFs above 1 kHz were more variable and less consistently matched with model peak times. However the relative timings to the different phase transition widths were still consistent with the model for these higher CF fibers.

The scatter in the measured responses across different fibers justifies our experimental design based on scaling invariance to quantitatively address the dependence on transition bandwidth.

Quantifying cross-frequency coincidence

We fit line segments to sets of peaks in the response patterns across FT and use the slopes of the fitted lines when FT was near CF to quantify the degree of cross-FT coincidence between stimuli with different phase transition bandwidths. For real fiber responses, each peak set across FT was assigned a number corresponding to the nearest model peak in time. In order to predict the degree of cross-*CF* coincidence, we used scaling invariance to normalize and shift the axes of our response patterns such that they represent our prediction for the true spatio-temporal pattern across *CF*.

Figure 2.11A shows the responses of a 999 Hz fiber to varying FT Huffman stimuli on the y-axis, with the time axis normalized by FT. The response pattern when the cochlear delay corresponding to the fiber's CF (999 Hz) was subtracted is shown in panel B. We applied scaling invariance by normalizing both frequency and time axes by FT and assuming the resulting pattern mimics that of an array of fibers responding to a fixed FT stimulus. Here, we use "virtual-FT" to refer to the FT of the hypothetical fixed FT stimulus, whose FT matches the true CF of the fiber. We also use "virtual-CF" to refer to the CFs of the hypothetical array of fibers, and set virtual-CF to be the normalized frequency (CF/FT) times the virtual-FT, or $\frac{CF^2}{FT}$. Then the cochlear delays corresponding to these virtual-CF locations were added to each time axis (Figure 2.11C). Slopes across virtual-CF were then measured from the resulting times of the response peaks (Figure 2.11D). For all three peaks shown, the sharp transition stimulus resulted in a higher absolute slope than the broad transition stimulus.

The fit of the lines to the data was quite good (across 125 recordings in 44 fibers, the mean error was 0.030 cycles). Slopes were always negative because the peak times always occurred later for higher FT (lower virtual-CF). For simplicity, we use the

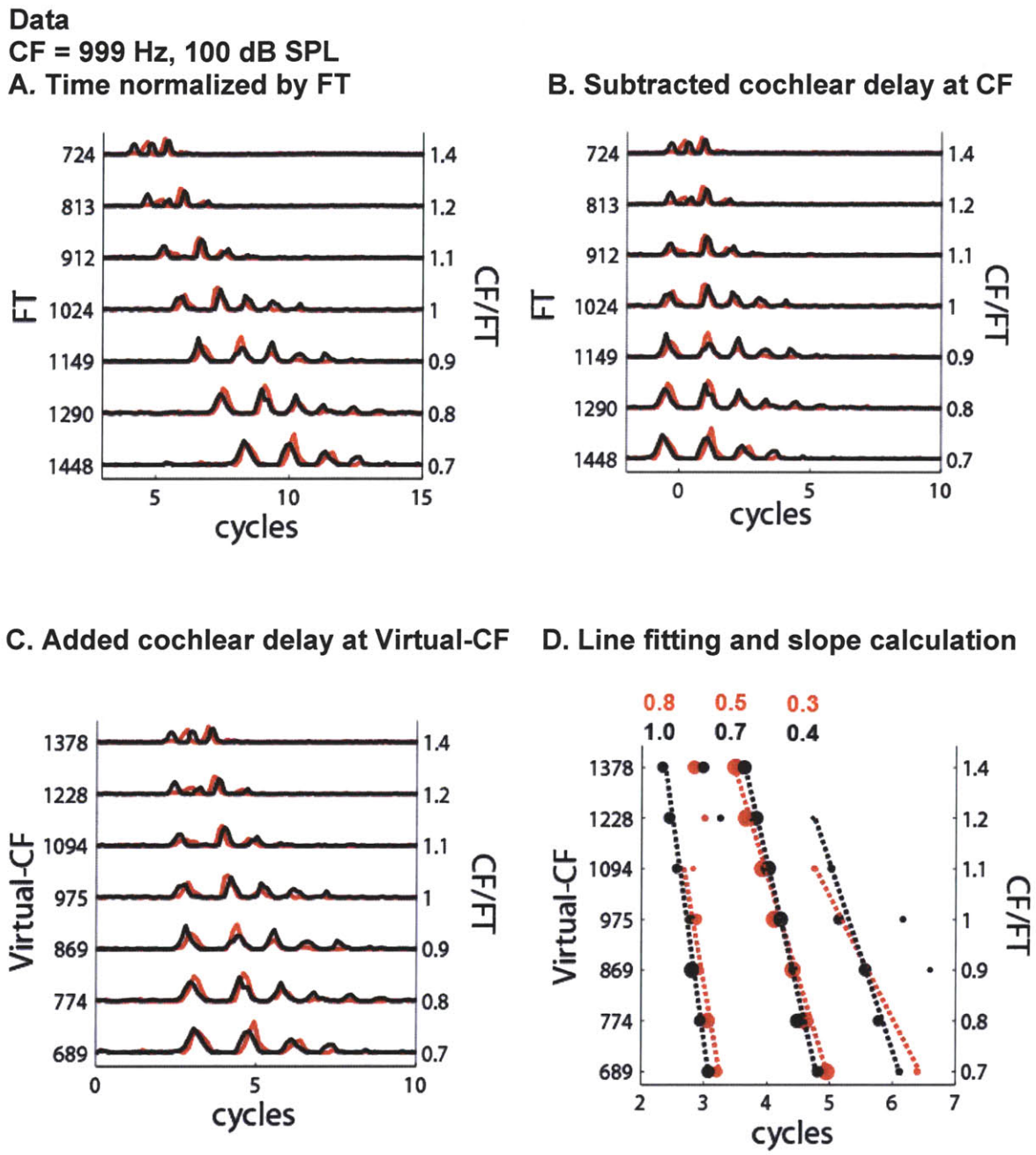


Figure 2.11: Responses of an AN fiber with CF 999 Hz to Huffman stimuli with varying FT on the y-axis and time on the x-axis. Responses to stimuli with broad phase transition in red, sharp phase transition in black. **A.** Time axis normalized by the FT of the true stimulus. **B.** Cochlear delay at the true CF subtracted from the time axis (same delay in seconds, but different number of cycles of FT for each row), and normalized by the FT of the true stimulus, as in Figures 2.7-2.9. **C.** Y-axis has been updated to reflect the "Virtual-CF," or CF^2/FT. Cochlear delays corresponding to "Virtual-CFs" (different delay in seconds for each row) have been added to the time axes. **D.** Line segments are fitted to the peaks in the pattern in panel C, and the absolute slopes when CF is near FT are shown above each peak set.

absolute values of the slopes to characterize the degree of cross-CF coincidence in the response patterns. Figure 2.12 shows scatter plots of the absolute slopes for the sharp transition (y-axis) against absolute slopes for the broad transition (x-axis) for Peaks 1, 2, 3, and 4+ (different panels). For each recording, the weighted average slopes (in the rightmost panels) were computed by summing each slope for Peaks 3 and later, weighted and normalized by the peak height when CF is near FT. The weighted average slopes represent the overall degree of coincidence throughout the late portions of the responses. Top panels plot slopes measured from model fibers across CF responding to a fixed FT stimulus. Bottom panels show slopes measured from real AN fiber responses for varying FT stimuli. Slopes measured from single model fibers responding to varying FT stimuli closely matched those measured from multiple model fibers across CF responding to a fixed FT stimulus (not shown).

A slope of one means that the response peak has shifted to the right one cycle of FT (in time) for a shift of one in normalized frequency CF/FT. A higher slope means greater cross-frequency coincidence (more vertical). A point above the identity line (indicated by the black solid line) means that the sharp phase transition stimulus excited more coincidentally across frequency than the broad transition stimulus.

We find that the model tended to over-estimate the degree of cross-CF coincidence compared to that found in the data. Furthermore, the data are somewhat more variable than the model, especially for Peaks 1 and 2. Despite these differences, both the data and the model show similar trends. For both transition widths in the model and data, the responses are less coincident (absolute slope decreases) for the later peaks

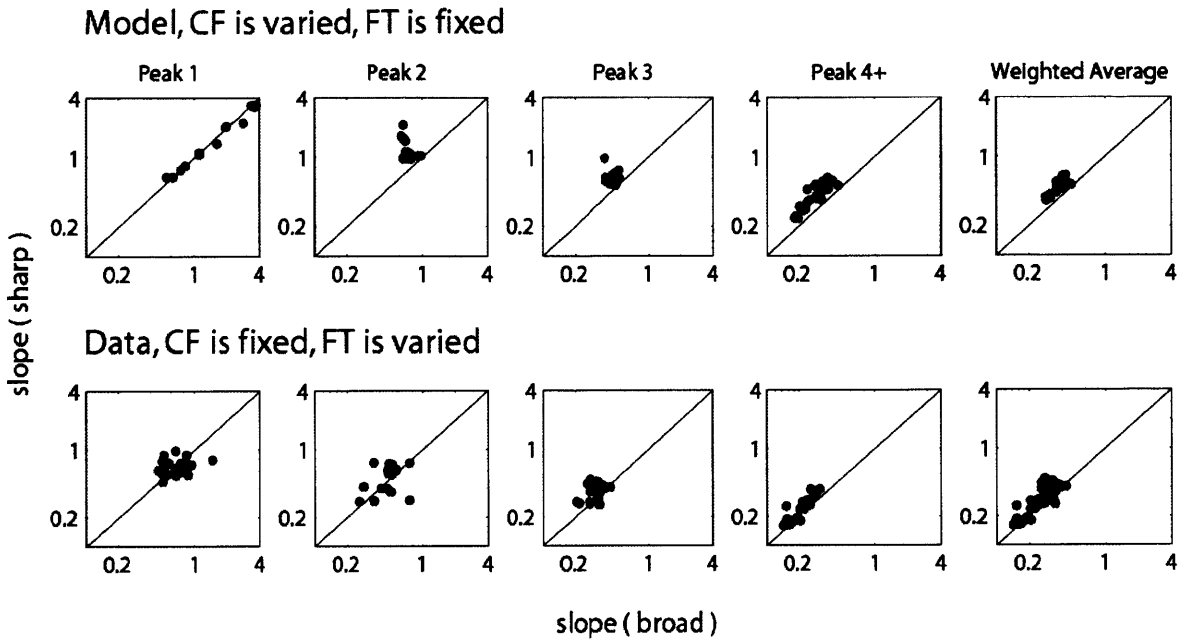


Figure 2.12: Absolute slopes for sharp transition (y-axis) and broad transition (x-axis). Top panels plot slopes measured from model fibers across CF responding to a fixed FT stimulus. Bottom panels plot slopes measured from real AN fibers responding to varying FT stimuli. Peak numbers are evaluated by reference to the number of the closest model peak. Rightmost panels plot average slopes for each recording, weighted by relative peak heights.

than they are for the earlier peaks. Model Peak 1 has nearly the same slope for both phase transition widths, while data Peak 1 shows more scatter. All the model slopes and most of the real fiber slopes for Peak 2 and later lay above the identity line, indicating that the sharp transition excited more coincidentally across frequency. In the real AN responses, several points for Peaks 2 and later lie below the identity line, but are still fairly close to the identity line. For both model and data, slopes for peak numbers greater than 4 tended to decrease further, and the points that have smaller slopes (around 0.2) in panels labeled "Peak 4+" corresponded to these later peaks.

Weighted average slopes in the model always indicated stronger cross-CF coincidence for the sharp transition stimulus (top right panel in Figure 2.12). The same trend was observed in the data, though to a lesser extent, as a small number of points lie

below the identity line in the bottom right panel of Figure 2.12. Furthermore, the overall estimated degree of cross-virtual-CF coincidence in the data was smaller than that predicted by the model for both transition widths.

In Figure 2.12, the data are pooled across all CFs and levels. Despite the differences in the early response peaks of the two fiber examples shown in Figures 2.7 and 2.9, there was no significant dependence on CF. The slopes tended to increase with increases in level, consistent with widening of the cochlear filters at higher levels, or the higher weighting of earlier peaks. However, there was no obvious trend for the slope *difference* (between the two transition widths) with level in both model and data (not shown).

Finally, slopes measured from a small number of fibers for which responses to stimuli with bandwidths matched to Carney's stimuli (1990) were also recorded displayed similar characteristics in that the early peaks elicited a higher degree of cross-CF (or FT) coincidence than the later peaks. In addition, in the later responses, the sharper transition stimulus excited more synchronously than the broader transition stimulus (not shown).

Discussion

Comparison with Carney's (1990) study

Carney (1990) hypothesized that the broad transition Huffman stimulus would excite more coincidentally across CF because it has a smaller group delay (smaller inverse slope of the phase) at FT compared to the sharp stimulus. In her Figure 2C, she plots model response peaks across CF (based on a simple linear cochlear filter bank derived from revcor analysis, Carney and Yin 1988) to broad and sharp transition stimuli (similar to our Figure 2.4). In her figure, the broad stimulus appears to excite more coincidentally for the earliest peak (occurring between 4-5 ms) when CF = FT. However the peak heights are low compared to the later peaks, where the sharp stimulus excites more coincidentally across CF. Carney limited her analysis to responses at low stimulus levels, since at high levels the responses could become dominated by the longer ringing for the sharp transition stimulus.

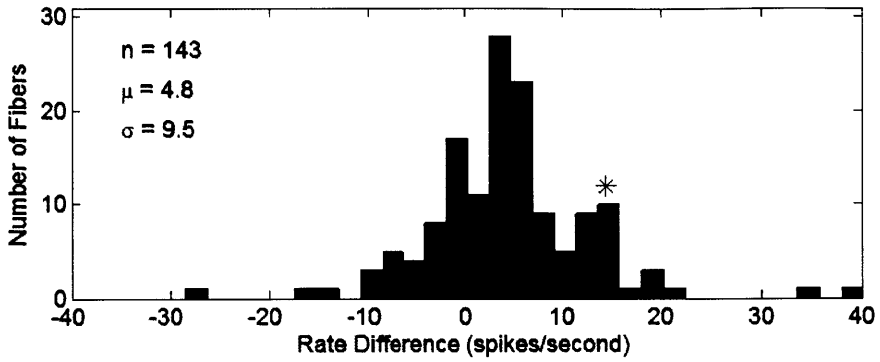
The responses of the updated AN model (Zilany and Bruce 2006) in Figure 2.4B resemble those of the older model (Figure 2C in Carney 1990) in that the Peak 2 height for the sharp transition stimulus is large when CF is greater than FT, decreases when CF is near FT, and ultimately disappears when CF is below FT. Furthermore, the Peak 3 height for the sharp transition stimulus is small when CF is near FT, and increases as CF is further from FT, while the Peak 3 height for the broad transition stimulus is comparatively more constant across CF. Finally, Peaks 4 and later all show greater cross-frequency coincidence for the sharp transition stimulus as compared to the broad transition stimulus. These observations are also consistent with responses recorded from real fibers.

Carney used the overall firing rate difference (RD, between the broad transition stimulus and the sharp transition stimulus, or $RD = R_{BROAD} - R_{SHARP}$ over the entire duration of the response) as a metric for sensitivity to the phase transition width. We repeated her analysis for our data and show a histogram of the maximum RD near threshold (level that elicited response just above baseline activity) across our population of AN fibers in Figure 2.13A. For our entire sample of 143 fibers, the mean RD was 4.8 spikes/second with a standard deviation of 9.5. The mean RD for Carney's population was 14.4 spikes/second with a standard deviation of 8.2. A t-test indicated that our distributions were significantly different ($p<1e-23$).

It is possible that the difference between our RD distribution and hers result from the use of different phase transition bandwidths: for the same Huffman r-values, our stimuli had wider bandwidths than hers. In a smaller number (n=28) of fibers, we also recorded responses to r-values that corresponded to the same bandwidths as Carney's stimuli (r = 0.92 and =.98). However, the mean RD for this smaller population was 4.0 spikes/second (histogram shown in Figure 2.13B). A t-test again indicated a significant difference between this distribution and Carney's ($p<1e-6$), so it is unlikely that the difference between our data and Carney's is attributed to the use of different bandwidths.

Another difference between our analysis and Carney's is that we limited our data to only those corresponding to when FT was set at CF, since this was where the larger RDs were observed. Carney also included responses to FT off CF and stated that these responses had smaller RDs than when the FT was on CF, indicating that the mean RD if she had limited her data to only FT at CF was likely to be even greater than shown in her average data.

A. RD = $R_{0.85} - R_{0.95}$



B. RD = $R_{0.92} - R_{0.98}$

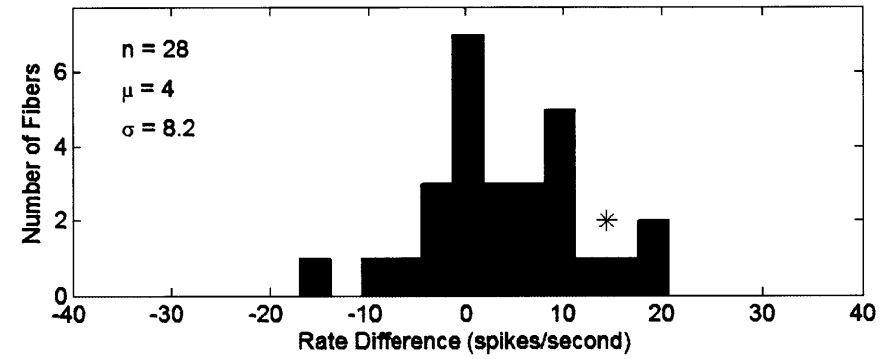


Figure 2.13: Maximum rate difference (spikes/second) between broad and sharp transition widths when FT = CF (near threshold level, lowest level for which response exceeded 5 standard deviations above baseline activity). Positive rate difference means preference for the broad transition stimulus. An asterisk indicates the mean rate difference of Carney's (1990) fibers. **A.** Rate difference between r-values of 0.85 and 0.95. **B.** Rate difference between r-values of 0.92 and 0.98.

It is also possible that our stimulus levels were higher than the levels that Carney used. RD tended to decrease with increases in level due to longer ringing of the sharp transition stimulus, causing prolonged responses to the sharp transition stimulus at higher levels. However we set our levels low enough to elicit responses just above spontaneous activity to ensure our stimulus levels were not too high. Furthermore, limiting our data set from Figure 2.13 to only include data which had the lowest thresholds to Huffman

stimuli did not significantly increase the mean RD (Panel A: n=26, mean=4.3 spikes/second, Panel B: n=16, mean=4.7 spikes/second)

A final possibility is that the differences between our findings and Carney's result from differences in stimulus generation. Carney states that her broad transition stimulus has a slightly higher spectral level than the sharp transition stimulus (~2 dB), possibly contributing to the preference of her fibers for the broad transition stimulus. For our data, a 2 dB increase near threshold resulted in an average increase in spike rate of 2.9 spikes/second. This may account for some, but not all of the 9.6 spikes/second difference between our mean RD and Carney's mean RD.

Carney used the RD in the AN to provide a baseline for comparison to the RD in CN neurons that may perform cross-frequency coincidence detection. CN neurons were said to be sensitive if RD exceeded 2 standard deviations away from the mean of the AN population. She found that compared to some CN neurons, the sensitivity to phase transition bandwidth in the AN was insignificant. Since our AN fibers were even less sensitive to the phase transition than Carney's AN population (possibly resulting from differences in stimulus generation), it is important for us to compare responses of our CN neurons to those of our AN fibers. In Chapter 3 of this thesis, we use a similar rate metric for evaluating spatio-temporal sensitivity of CN neurons, but equipped with a better understanding of the patterns elicited by these stimuli across the AN.

Model vs.Data

One notable difference between our model and data results was that the model fibers displayed responses at lower stimulus levels than real fibers. The mean level threshold

for Huffman stimuli (lowest level to elicit a PSTH response at least 5 standard deviations above activity during silent interval between subsequent stimuli) in real fibers was typically around 70 dB SPL, while model fibers already displayed responses to stimuli presented at 40 dB SPL. Furthermore, model responses lasted longer, "ringing" for up to 8 cycles, while real fibers never displayed such long responses, rarely lasting for longer than 5 cycles. Another difference was observed in low CF fibers (<800 Hz), where the model underestimated the degree of phase locking (compare Figures 2.7 and 2.8).

One possible explanation for these differences is that we use the output of the inner hair cell synapse as an approximation for the PSTH of the model fibers. The model's spike generator introduces refractoriness after the inner hair cell synapse, and might affect the peak width, peak height, or the duration of the response. However, simulated results using model PSTHs closely matched with those of the synapse output, so the spike generator is unlikely to play a role in explaining the differences observed between model and real fibers.

Another possibility is that the hearing thresholds in our animals were not ideal. Higher thresholds in the real fibers would explain the need to go to higher stimulus levels in order to evoke responses. However, cochlear function was closely monitored during experiments with the use of an electrode at the round window. The experiment was terminated if the click-evoked CAP threshold increased by more than 10 dB since the beginning of the recording session. Furthermore, we discarded data recorded from fibers with thresholds that exceeded 3 standard deviations from the mean of a standard database of AN fibers (Liberman 1978). The thresholds of our resulting set of fibers matched

those of the model fiber thresholds, which were measured using the same tracking algorithm used for measuring tuning curves of real fibers (Kiang et al. 1970).

Summary

In this chapter, we compared responses of model fibers and real AN fibers to Huffman sequences, which have an all-pass magnitude spectrum and a single phase transition. We found that these stimuli can be used to systematically manipulate the spatio-temporal pattern of responses in the AN across CF while causing minimum changes in firing rate. We used cochlear scaling invariance to infer the true spatio-temporal pattern across an array of fibers with different CFs from the responses of a single fiber to stimuli with varying phase transition frequencies. We found that regardless of the phase transition width, the early responses to Huffman sequences were more synchronous than the late responses. In addition, we found that at low levels, the early response peaks to the stimulus with the sharp phase transition tend to be smaller than the response to the broad transition stimulus when the transition frequency FT is near CF. The early response to the sharp transition stimulus increases when FT is away from CF, while the early response to the broad transition stimulus remains roughly constant for different FT. Furthermore, for the late response peaks, the sharp transition Huffman stimulus excites more coincidentally across CF than the stimulus with the broad phase transition. These results were consistent across stimulus levels and FTs, and held for both real fibers as well as model predictions. Finally, as expected, the firing rates of AN fibers showed little sensitivity to the phase transition bandwidth in that overall rate responses to broad and sharp transition stimuli were similar across the population.

CHAPTER 3

Sensitivity of Cochlear Nucleus Neurons to Spatio-Temporal Changes in Auditory Nerve Activity

Abstract

Patterns of auditory nerve (AN) activity contain spatio-temporal phase cues to the pitch of harmonic complex tones that are more consistent with psychophysics than traditional rate-place or temporal cues (Cedolin and Delgutte 2010). These spatio-temporal cues might be extracted in the cochlear nucleus (CN) by a cross-CF coincidence detection (CD) mechanism, where a CN neuron would fire preferentially when its AN inputs tuned to different frequencies discharge in synchrony.

We implemented a CD model receiving AN inputs from varying extents of the tonotopic axis. We used Huffman stimuli, which have flat magnitude spectra and a 2π phase transition, to manipulate the relative timing of spikes across neighboring AN fibers. Responses of CD model cells tended to be shorter than those of AN fibers. As the input range of CFs to the CD cell widened, CD cells increasingly preferred the stimulus with sharp transition, and the average peak width in the CD response tended to increase.

We used a maximum likelihood approach to quantitatively assess the similarity between CN units and CD cells, and found the parameters of the CD model that were most likely to give rise to responses observed in the CN. While many CN units had responses more similar to those of AN fibers rather than CD cells, certain unit types (primary-like-with-notch and some phase lockers) had responses consistent with cross-frequency CD cells.

Introduction

In Chapter 2, we used "Huffman sequences" (Carney 1990) to systematically manipulate the relative timing of activity across auditory nerve (AN) fibers tuned to low characteristic frequencies (CFs). In this chapter, we present responses of units in the cochlear nucleus (CN) to these stimuli. Huffman sequences have a flat magnitude spectrum and a single phase transition at a transition frequency (FT). We showed that by changing the slope of the phase transition, we could systematically alter the pattern of activity across the AN without greatly changing the discharge rate across CF. Specifically, we showed that a stimulus with a sharper phase transition resulted in a better temporal alignment of AN activity across different CFs (or stronger cross-CF coincidence) in the late response compared with responses to a stimulus with broader phase transition. In addition, regardless of the phase transition, the early responses were more aligned across CF than the later responses. Furthermore, while the height of the early response to the broad transition stimulus remained roughly constant across CF, the early response to the sharp transition stimulus was small when FT was near CF and increased when FT was away from CF.

The phase cues in the AN, or the degree of cross-CF coincidence, could be extracted centrally by a mechanism that is sensitive to the relative timing across AN fibers tuned to different CFs. One way to extract these cues is a network of cross-frequency coincidence detector (CD) cells, which produce more spikes when the activity of AN fibers with different CFs is temporally aligned, and produce fewer spikes when AN activity is less aligned.

All AN fibers provide direct inputs onto neurons in the CN, making the CN an obligatory relay in the ascending auditory pathway. Moreover, the CN contains many distinct cell types differing in their morphology, synaptic inputs, and output projections. Several attempts have been made to describe the nature of the convergence of AN fibers onto CN cells both anatomically and physiologically. There is anatomical evidence indicating that certain types of cells in the CN receive inputs from multiple AN fibers, and these studies have estimated the number of afferent inputs for various CN cell types, which can range from one to hundreds, depending on the cell type (less than 5 inputs per spherical bushy cell (Liberman 1991, Sento and Ryugo 1989), 15-25 inputs per globular bushy cell (Spirou et al. 2005, Liberman 1991), 5-8 inputs per stellate cell (Liberman 1991), and more than 60 inputs per octopus cell (Liberman 1993, Cao and Oertel 2010). There is also physiological evidence for coincidence detection in the CN, based on enhanced onset responses and enhanced phase locking to pure tones (Joris et al. 1994ab, Joris and Smith 2008).

Less is known about the range of tonotopic locations of the input AN fibers and their relationship to the CF of the CN cell. At one extreme, spherical bushy cells in the CN receive only one or two AN excitatory inputs with large synapses (Cant and Morest 1979, Sento and Ryugo 1989, Ryugo and Sento 1991). Due to the small number of convergent inputs, the responses of these cells often resemble those of AN fibers and are unlikely to be sensitive to changes in the AN's spatio-temporal pattern. At the other extreme, CN octopus cells are thought to perform coincidence detection on AN fibers across a very wide range of CFs, because of their broad tuning curves and temporally

precise onset responses (Oertel et al. 2000). However, the range of CFs of inputs from the AN for any CN cell type has not been explored systematically.

In this chapter, we study whether cross-frequency CD occurs in the CN by comparing the responses to Huffman sequences of model CD cells operating on AN activity to those of CN units. We use a CD mechanism to model responses of a cell that prefers to fire when its inputs discharge in synchrony. We systematically vary CD model parameters to find a set of parameters, if it exists, that best matches CN unit responses. We find that while responses of most CN units resemble those of AN fibers more than those of CD cells, some CN unit types, likely corresponding to globular bushy cells, have responses consistent with cross-frequency CD cells.

Methods

Neurophysiology and Stimuli

The methods for recording from single units in the cochlear nucleus (CN) are largely as described in the classic studies of Bourk (1976), Pfeiffer (1966), and Godfrey et al. (1975). The surgery, anesthesia, and acoustic stimulation system are the same as for recording from single auditory nerve (AN) fibers described in Chapter 2. The only differences between the AN and CN preparations are the exposure, electrodes, and search stimuli. Rather than retracting the cerebellum as was done to expose the surface of the AN, the cerebellum was partially aspirated to expose the surface of the dorsal CN and the posterior portion of the antero-ventral CN. Parylene insulated tungsten electrodes were used to record action potentials. Wideband noise at 75 dB SPL was used to search for single units.

Upon isolation of a CN unit, its frequency tuning curve (Kiang et al. 1970) and spontaneous activity were measured. Then the response to tone bursts at CF with 30 ms duration were measured as a function of level. The CN unit type is determined following established classification methods (Bourk 1976, Blackburn and Sachs 1989, Young et al. 1988) using the post stimulus time histograms (PSTHs), the first-order interspike interval (ISI) histograms, regularity analysis (performed by taking the coefficient of variation (CV) of the mean ISI), and first spike latency.

Responses to Huffman sequences with different phase transitions (described in Chapter 2) at various stimulus levels were then studied. The transition frequency (FT) of the Huffman sequence was always set to the CF of the unit. As in Chapter 2, the

threshold level was found by listening for an audible response just above spontaneous activity or by measuring a quick rate-level function. A range of stimulus levels from near threshold to 40 dB above threshold in 10 or 15 dB increments was used. Level was interleaved randomly across trials to minimize any effect of the order of sound levels in the stimulus presentation.

Analysis

Coincidence detector model cells

In Chapter 2, we found that, with proper scaling in time and frequency, we can infer the spatio-temporal pattern of AN responses across CF to a Huffman Sequence with fixed FT from the responses of a single AN fiber (fixed CF) to varied FT stimuli. Here we use measured responses of real AN fibers to Huffman sequences with varying FT to form "virtual" spatio-temporal response patterns, which include the normal delays associated with the cochlear traveling wave. We use these virtual spatio-temporal response patterns as the inputs to a coincidence detector (CD) model to predict responses of a CD cell that would receive AN inputs across CF.

We use an analytical point process model of CD neural cells (Krips and Furst 2009) to compute the responses of neurons that receive several AN fibers as inputs. This model provides a convenient framework for studying responses of CD cells because, rather than using raw spike times (which can be computationally slow), the model analytically derives the instantaneous firing rate of the output cell from the instantaneous rates of its inputs. The model is described in more detail in the appendix.

We systematically varied parameters of the model, including the total number of input fibers (N = 4, 10, 16), the minimum number of inputs (L) required to fire within the

coincidence window (0.1 ms) to generate a spike in the output cell, and the range of their CFs (w in octaves). All inputs were assumed to be excitatory and from AN fibers.

To determine realistic values for L and w, frequency tuning curves were measured from CD cells receiving input from model AN fibers (Zilany and Bruce 2006) using the same automatic tracking algorithm (Kiang et al. 1970) used in the experiments. When the CD parameters were too strict (if L or w were too large), the CD cells were unresponsive to the 100 ms tone pips at any level. Table 1 shows the range of w for different values of N and L for which the model cells showed a reasonable tuning curve (thresholds to pure tones at CF below 50 dB SPL). These parameters were used in the CD model. The maximum frequency range of AN inputs was set to 1 octave because that is the range over which the FT parameter of the Huffman sequence was varied in Chapter 2, and therefore cochlear scaling invariance could be applied.

	N=4	*N=10*	*N=16*
L=2	1	1	1
L=3	0.25	1	1
L=4	X	0.5	1
L=5	X	0.25	0.5
L=6	X	X	0.1

Table 1. Maximum input range (w, in octaves) for CD cells that had reasonable tuning curves.

Interpolating responses across FT

To implement the CD cell, we would ideally have measured each AN fiber's response to fine steps of FT so as to sample the virtual spatio-temporal pattern with fine spatial resolution. In practice, due to limited holding times in the AN, the FT of the Huffman Sequences (see Chapter 2, Methods) was only varied in 1/6 octave steps across a range of one octave centered at the AN fiber's CF. Therefore, we used dynamic time warping to

interpolate between responses from two consecutive FTs to predict responses to a set of FTs at finer sampling (1/36 octave).

The algorithm for dynamic time warping begins by taking the distance (absolute difference) between each time point in a first signal and each point in a second signal, generating a two-dimensional matrix of distances. Then, starting at the bottom left corner (the first point in both signals) and ending at the top right (the end point in both signals), the optimal path between the two corners is traced by iteratively choosing the mapping with the lowest cumulative distance. The result is an efficient and nonlinear mapping in time between the two signals, which was used to derive responses to FTs in between those actually recorded.

To interpolate responses, we assumed that both the temporal map generated by

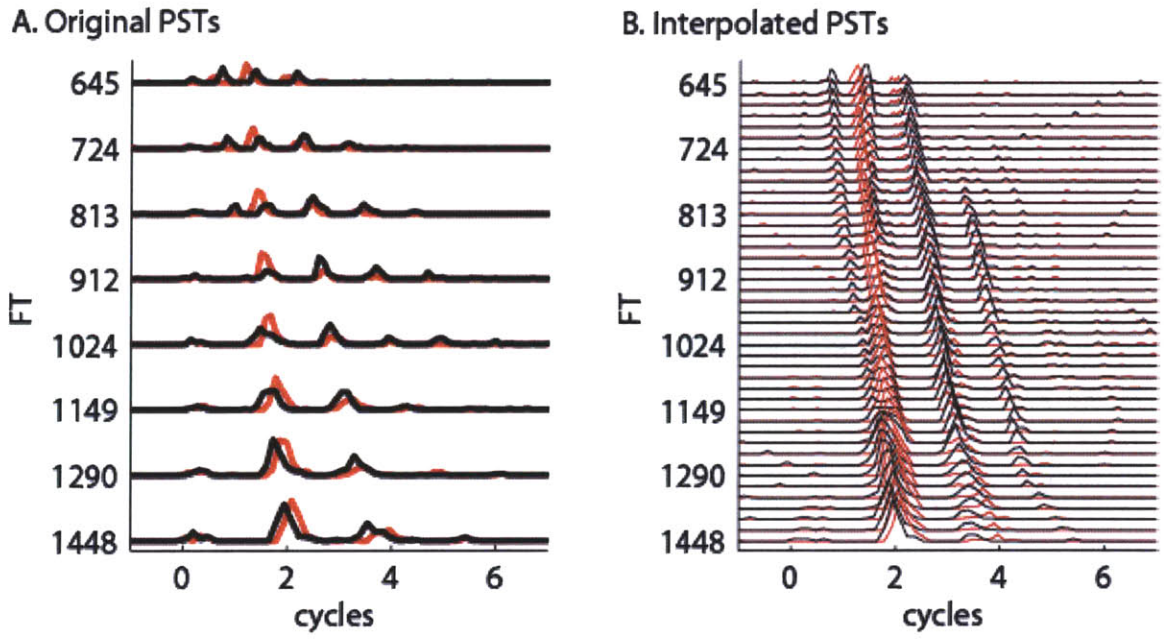


Figure 3.1: Responses to Huffman stimuli of a 972 Hz fiber at 100 dB SPL (65 dBreTh), with broad phase transitions are in red, and responses to sharp transition stimuli are in black. FT on the y-axis and time normalized by FT on the x-axis. **A.** Recorded PSTH responses for 6 values of FT. **B.** Estimated responses of the fiber for 35 values of FT, derived from dynamic time warping analysis.

the dynamic time warping algorithm as well as the amplitude in the response vary logarithmically with FT. An example of recorded responses of an AN fiber (CF 972 Hz) to stimuli with varying FT is shown in Figure 3.1A. Panel B repeats the original responses, and also plots the interpolated responses for FTs between those recorded. The interpolation algorithm generates response patterns that change smoothly with FT, such that as was done in Chapter 2, line segments could readily be fit to sets of peaks across FT with small residual error.

Quantifying sensitivity to phase transition and temporal characteristics

We use a normalized rate metric to quantify sensitivity to the phase manipulations in the Huffman stimulus. Carney (1990) used the rate difference (RD) in spikes/second between responses to the broad and sharp transition stimuli to quantify phase sensitivity. In order to normalize for overall differences in firing rates across units, we normalize the rate difference by the rate sum to get a dimensionless metric $RD = \frac{r_{broad} - r_{sharp}}{r_{broad} + r_{sharp}}$.

In Chapter 2, we showed that it was necessary to separate early and late portions of responses because the sharp transition stimulus was found to excite more coincidentally across CF than the broad transition stimulus only in the late portion of the response. Furthermore, the observation that the response to the sharp transition stimulus increased when FT was further from CF was only observed in the early portion. We defined the border between early and late windows to be halfway between peaks labeled 3 and 4 in AN responses. This border varies with CF and occurs around 1.5 to 2 cycles after the cochlear delay (see Figure 2.10B). Since the CD model adds no additional delay, the same early-late border was used for CD cells.

CN neurons have longer and more variable latencies than AN fibers, making it difficult to determine which CN response peaks corresponded to the early part of AN responses and which corresponded to the late response. For this reason, it was not ideal to use the same early-late border as was defined for AN fibers and CD cells. To remediate this problem, we found that on average, 25% of the spikes in the CD model occurred before the early-late border as described above. Thus, the early-late border in individual CN responses was defined to occur after 25% of the spikes had occurred.

While the RD quantifies overall preference for phase transition and thus sensitivity to the AN spatio-temporal pattern, there is still information in the responses not captured by this metric. Temporal characteristics of the responses also change as parameters of the CD model changes. At one extreme, a CD cell with strict requirements (if L=N, for example) can very roughly be approximated as performing a multiplication across its inputs, resulting in temporally sharpened responses relative to those of AN fibers. So, we expected responses of CD cells with strict parameters to have shorter overall durations and narrower peak widths than those of AN responses. At another extreme, if the requirements for the output cell to spike are loose ($L<<N$), the cell may behave as if it were simply summing its inputs, resulting in longer durations and wider peak widths if the inputs are not temporally aligned. Thus, in addition to using the RD metric, evaluating the temporal characteristics of CN responses should also provide insight into whether a CN unit performs an operation similar to coincidence detection, and the CD parameters that describe that operation.

PSTHs were computed with binwidth 1/25 cycles of FT. For each response pattern, a "threshold" was set at 5 standard deviations above baseline activity measured

during the last 10 ms of each inter-stimulus interval (long after both stimulus and response had decayed). Peaks in the response pattern were found by detecting all local maxima in the PSTH above this threshold. The overall normalized duration (ND, in cycles of FT) of each response was measured by taking the distance from the first time bin to the last time bin that the response exceeded threshold. The width (in cycles) of each response peak (PW) was measured by the time interval surrounding the peak that exceeded threshold. When the response pattern had more than one peak, the mean peak width across all peaks in the pattern was used.

In total, we used six metrics to quantify and compare the responses of AN fibers, CD model cells operating on AN activity, and CN neurons to Huffman stimuli. We used two RD metrics to quantify sensitivity to the changes in stimulus phase transition for the early and late response, and two metrics to quantify temporal characteristics for each transition width:

$$\boldsymbol{Y} = [\, RD_{Early}\,,\, RD_{Late}\,,\, ND_{Broad}\,,\, ND_{Sharp}\,,\, PW_{Broad}\,,\, PW_{Sharp}\,].$$

Maximum likelihood test

We used a maximum likelihood approach to quantitatively assess the similarity between responses of CN neurons and those of CD model cells and AN fibers. We used nonrandom parameter estimation from estimation theory, in which the model parameters are deterministic but unknown. Nonrandom parameter estimation is distinct from Bayesian estimation theory, which assumes random model parameters. We let X denote the set of CD parameters [N, L, w]. For each vector ($x \in X$) of CD parameters, we measured the distributions of each of the six metrics in $\boldsymbol{Y}$ for a set of AN inputs, and use

$p(y\,;x)$ to denote the joint distribution of metrics for each parameter set. Then for a set of metrics measured from CN responses y_{CN}, our goal is to find the CD parameters that are most likely to give rise to the observed CN response. We use maximum likelihood to estimate the most likely CD parameters $\hat{x}$:

$$\hat{x} = \arg\max_{x} \left(p(y_{CN}; x) \right)$$

We use the same approach to determine whether individual CN units are more similar to CD model cells or AN fibers.

We first characterized responses of CD cells with different model parameters (X = [N, L, w]) by measuring the distributions of RDs, NDs, and PWs for a set of AN inputs (125 responses from 44 fibers). Since the responses obviously varied with stimulus level, the responses were separated into low (<53 dB re. Th., or relative to the pure tone threshold) and high (>= 53 dB re Th) levels. For each vector of CD parameters x, six (one for each metric in $\boldsymbol{Y}$) histograms across all 125 input sets were computed. The distributions are plotted as images (with input range w on the y-axes, normalized by 125 so that each row sums to one) for [N, L]=[10, 2] in Figure 3.4. Beneath each image plot is the histogram computed from the input 125 AN responses when FT was equal to the CF. Histograms were smoothed with a three-point filter [¼ ½ ¼] in both dimensions (y and w). The ND was rounded to the nearest integer because peaks tended to be spaced one cycle apart. Finer sampling in ND resulted in multimodal distributions with integer spacing.

Our goal with the maximum likelihood test was to find the parameters of the CD model $\hat{x}$ that maximize the probability of obtaining the measured CN response $p(y_{CN}\,;x)$. We used the distributions measured from the set of CD responses to estimate the true

joint distribution p(y ; x). We assumed independence across the six metrics in $\boldsymbol{Y}$ as well as independence across recordings at different levels, such that the p(y ; x) may be approximated by the product of the distributions for each metric across stimulus levels (columns in Figure 3.4). We call this product $Q(y,x)$:

$$Q(y_{CN},x)=\prod_{i=1}^{\#levels}\prod_{j=1}^{6}p(y_{CN_{i,J}};x).$$

The resulting likelihood matrix $Q(y_{CN}, x)$ corresponds to the probability that the CN response pattern results from a CD cell with parameters x=[N,L,w]. We use $\hat{x}=\left[\hat{N},\hat{L},\hat{w}\right]$ to denote the model parameters that maximized this likelihood:

$$\hat{x}=[\hat{N},\hat{L},\hat{w}]=\arg\max_{x}\left(Q(y_{CN},x)\right).$$

Finally, to compare the CN responses to those of AN fibers, the six scalar probabilities corresponding to the AN distributions (bar plots in Figure 3.4) were multiplied together, and then multiplied across recordings at different stimulus levels. If the resulting probability was larger than $\max\left(Q(y_{CN},x)\right)$, then the CN unit was said to resemble AN fibers more than any CD cell.

Results

Responses of coincidence detector model cells

We limited our analysis to low CF AN fibers and CN units within the range of phase locking, such that the relative timing of discharges across CF in the AN could be manipulated in the fine structure. We recorded from 143 AN fibers (with CFs 107 – 2565 Hz) in 11 cats. In all fibers, we recorded the response to Huffman stimuli with FT set to the CF for two phase transition widths (r=0.85, r=0.95). In a subset of our fibers (44/143), we also varied the FT from a half octave below to a half octave above the fiber's CF in 6 steps per octave. The responses to varied-FT stimuli were interpolated to estimate responses to finer changes in FT (in 36 steps per octave). The resulting estimated "virtual" spatio-temporal patterns were input into a coincidence detector (CD) model. In this section, we compare the responses of an example AN fiber with those of CD cells operating on AN activity. We then describe how CD responses vary with changes in the CD parameters.

Coincidence detector cell example

Figure 3.2AC illustrate the interpolated responses of an AN fiber with CF 972 Hz responding to Huffman stimuli with different FT on the y-axis for two different levels: 50 dB re. Th. (relative pure tone threshold, panel A) and 65 dB re. Th. (panel C). Responses of model CD cells operating on this pattern are shown in the panels in Figure 3.2BD for the case when FT = CF. Different columns correspond to different values for the total number of inputs (N), and rows indicate different input frequency ranges (w, in octaves).

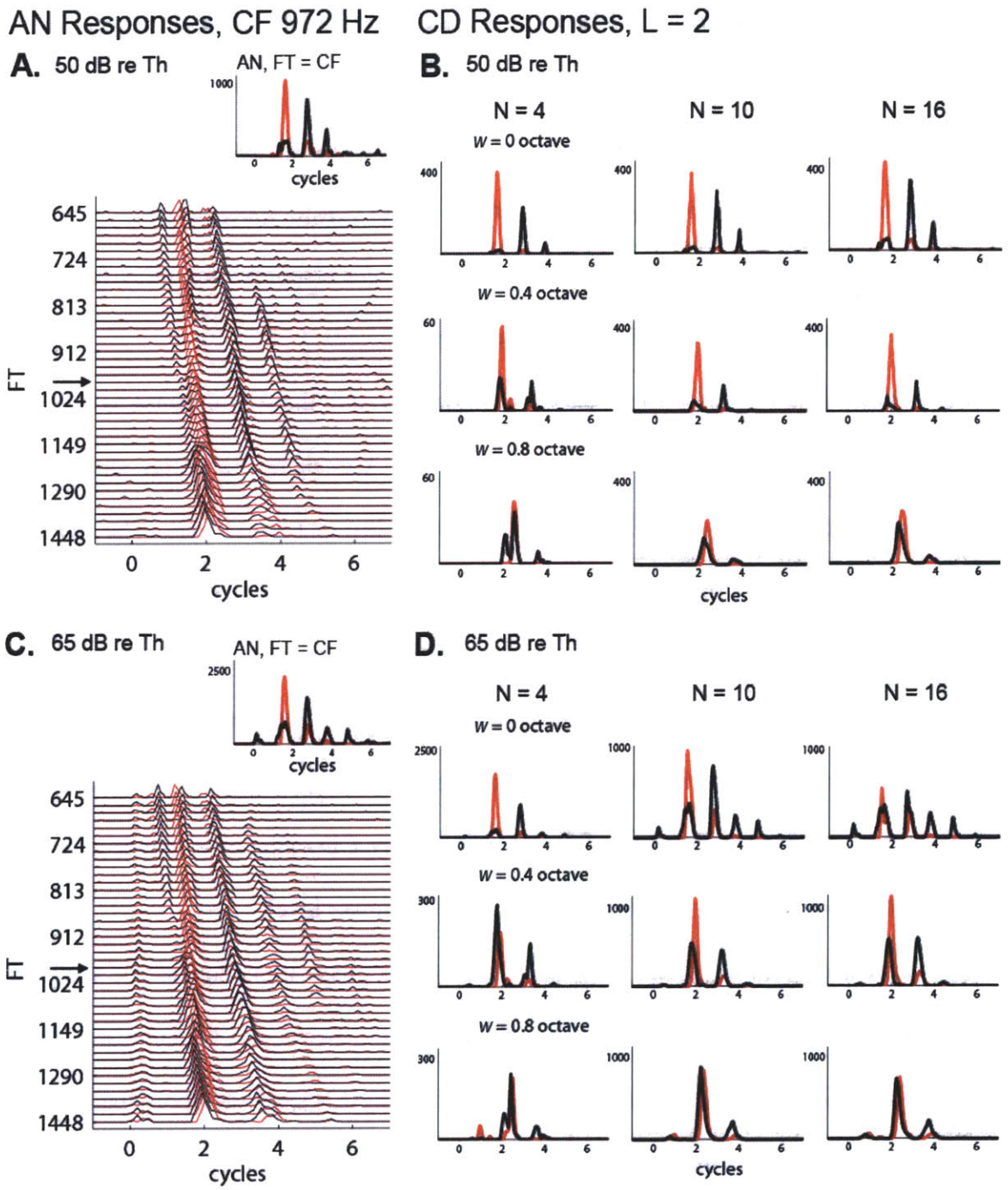


Figure 3.2: Responses of AN fiber and CD cells. **A.** Top panel shows responses (spikes/second) of the 972 Hz fiber when FT=CF. Bottom panel shows estimated responses with varying FT on the y-axis (arrow indicates which pattern is shown in the top panel). Level is 50 dB re. Th. (relative to pure tone threshold). **B.** Responses of CD cells with 4 (left), 10 (middle), and 16 (right) inputs taken from the AN pattern in panel A. The output cell produces a spike if at least 2 (L) of its inputs fire coincidentally (within 0.1 ms). Different ranges of input frequencies are shown: 0 (same-frequency CD, top), 0.4 (mid-range CD, middle), and 0.8 (wide range CD, bottom) octaves. Y-axes are in spikes/second. **C-D.** Same as in A-B but for 65 dB re. Th.

The AN inputs to the CD model all had FT ranges centered on the CF of the fiber.

All CD responses shown in Figure 3.2 have a fixed L of 2 (meaning at least 2 inputs have to fire within the coincidence interval 0.1 ms for the output cell to produce a spike). As N increases, the requirement for the output cell to spike becomes relatively less stringent. For large N, this is reflected in the longer responses for same-frequency CD cells and higher firing rates for cross-frequency CD cells.

When the FT is at CF (inset panels in Figure 3.2AC), the AN response to the early portion prefers the broad transition stimulus while the sharp transition stimulus is preferred in the late response. This is true for both the low and high levels. The CD cells that receive all inputs from fibers with the same CF, or same-frequency CD cells (w=0, top rows of Figure 3.2BD) enhance features present in the AN response pattern. So, no matter the number of inputs, the same-frequency CD cells all have an early preference for the broad transition stimulus and a late preference for the sharp transition stimulus. Furthermore, the N=10 and N=16 CD same-frequency CD cells have longer durations and slightly wider peaks at the higher level than at the lower level.

In Chapter 2, we found that regardless of the phase transition of the Huffman stimulus, the early AN response displayed greater coincidence across CF compared to the late portion. We further found that the late portion of the AN spatio-temporal pattern was excited more coincidentally across CF for the sharp transition stimulus compared to the broad transition. We expected that cross-frequency CD cells would convert the degree of AN cross-frequency coincidence into a rate response, and thus would have larger early responses compared to late responses, and larger late responses for the sharp transition stimulus compared to the broad.

The cross-frequency CD cells in Figure 3.2BD (w=0.4 and w=0.8 in middle and bottom rows respectively) for N=10 and 16 have responses closely matching our expectations. In each of these panels, the earlier peak is much larger than the later peak. In addition, only the sharp transition stimulus elicited a prominent late peak, while the broad transition stimulus had little or no late response.

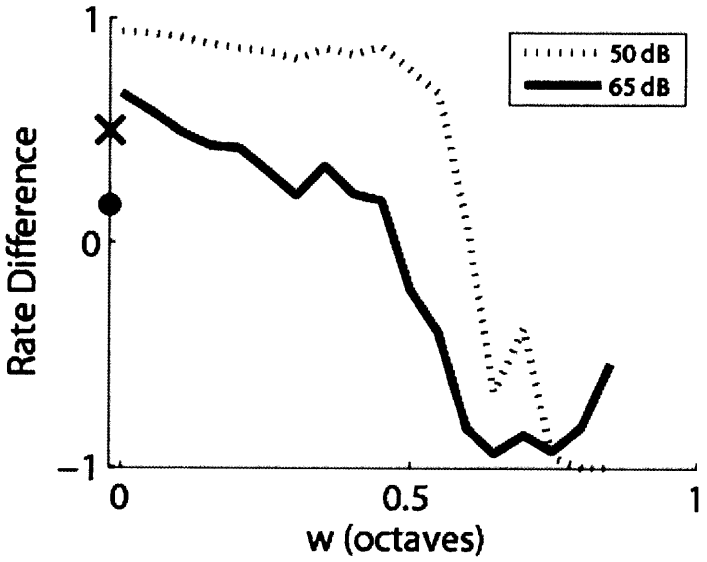


Figure 3. 3: Early rate differences for characterizing response patterns of AN fibers and CD cells shown in Figure 3.2, as a function of input range (w). AN data are plotted on the ordinate axis. N=10, L=2. 50 (CD dotted line, AN x) and 65 (CD solid line, AN solid dot) dB re Th.

Figure 3.3 plots the RD_{Early} as a function of input range w for the CD cell with N=10 shown in Figure 3.2. The RD_{Early} measured from the AN response (inset panels in Figure 3.2AC) are plotted on the ordinate axis. The dotted line (x for AN) indicates the lower level, and the solid line (solid dot for AN) indicates the higher level. The RD_{Early} of CD cells tends to decrease as w increases, becoming negative for w>0.4 octaves. This indicates a preference for the sharp transition stimulus for wide range CD cells. This occurs because the AN spatio-temporal pattern, when the CF is near FT, displays a small early response to the sharp transition stimulus relative to the broad transition stimulus. The response gets larger as CF is further from FT (Figure 3.2AC). The FT region for

which there is a small early AN response to the sharp transition stimulus is wider at the low level (Figure 3.2A) than at the high level (Figure 3.2C) and is illustrated with a schematic in Figure 3.11. This causes mid-range CD cells to have a larger early response to the sharp transition stimulus at the higher level (smaller RD_{Early}) compared to the lower level (Figure 3.3 and middle rows of Figure 3.2BD).

All the CD late peaks in the responses, regardless of input range w, are larger for the sharp transition stimulus compared to the stimulus with broad transition. Therefore, the CD late response alone does not distinguish between different values for w. However, the CD early response can provide information about the input frequency range because RD_{Early} systematically decreases as w increases.

CD population – Rate metrics

We used six metrics to characterize responses of AN fibers and CD cells to Huffman stimuli: RD_{Early}, RD_{Late}, ND_{Broad}, ND_{Sharp}, PW_{Broad}, and PW_{Sharp}. The image panels in Figure 3.4 plot the probability distributions for CD cell responses ([N, L]=[10, 2]) across 125 virtual spatio-temporal patterns measured from 44 AN fibers separately for low levels (<53 dB re Th., Figure 3.4A) and high levels (>53 dB re Th., Figure 3.4B). Distributions from CD cells with different input frequency ranges are plotted on the y-axes. Bar plots below each image show distributions of the metrics formed from the original 125 AN responses with FT set to the CF.

Consistent with the example responses shown in Figure 3.2, RD_{Late} is usually negative for both levels. For low levels, RD_{Late} is slightly larger for wider-range CD cells than for narrow range cells. Also at low levels, RD_{Early} tends to have large positive

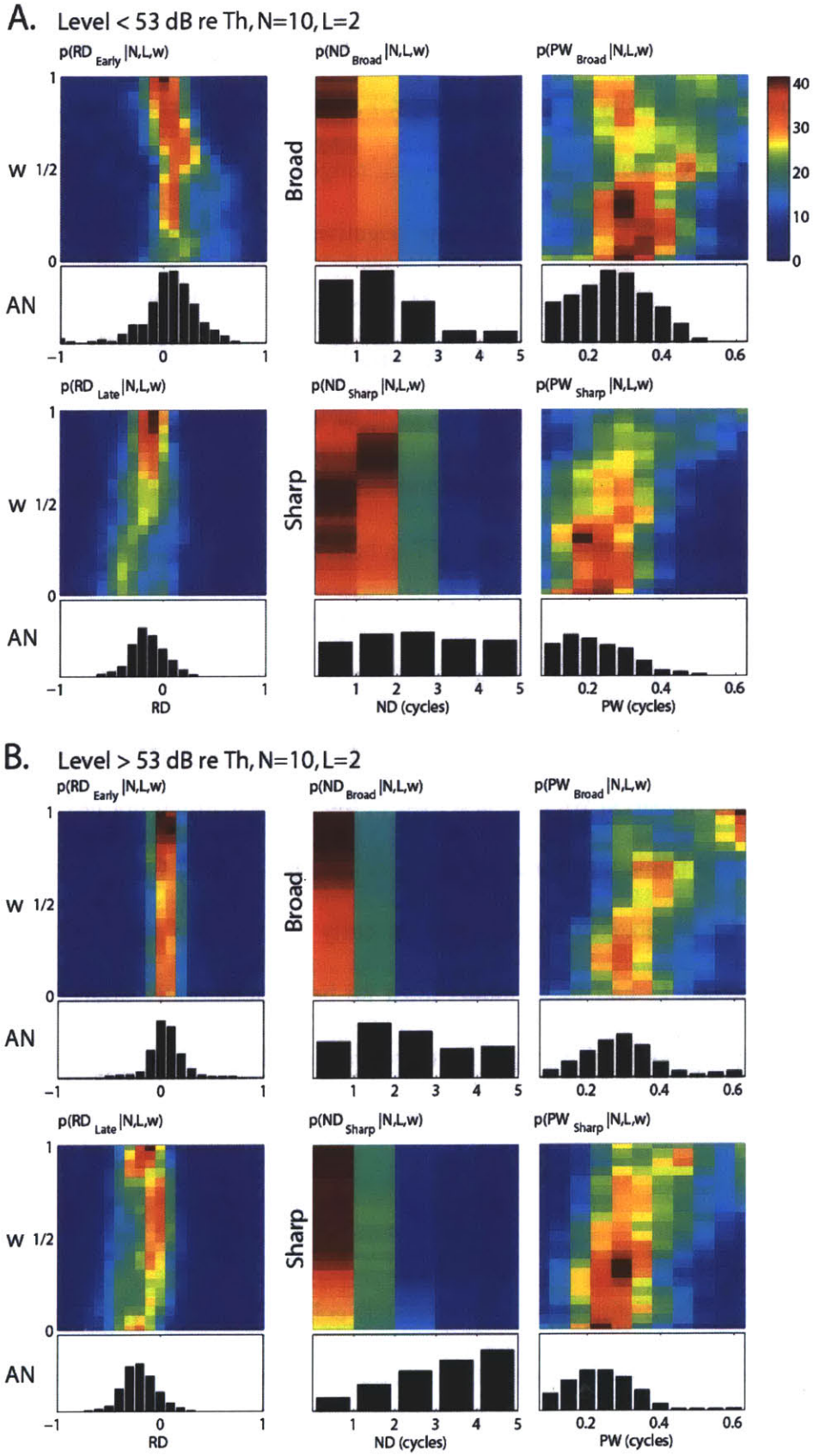


Figure 3.4: Histograms of the distributions of different metrics used to characterize responses of AN fibers and CD cells (N=10, L=2) to Huffman stimuli for low levels (**A**) and high levels (**B**). Images show histogram distributions for CD cells with input range w on the y-axes (pixel intensity denotes percentage of recordings). Bar plots below each image show distributions of AN fibers. Distributions of early rate differences (RD_{Early}) are shown in top left panels, and RD_{Late} is shown in bottom left panels. Temporal metrics normalized duration ND and mean peak width PW in cycles are shown in the middle and right columns, respectively. For ND and PW, broad transition responses are shown on top and sharp transition responses on the bottom.

values for narrow-range CD cells and decreases for wider-range CD cells. This happens because as input range increases, the early response of CD cells to the sharp transition increases due to increased AN activity when FT is off CF. At higher levels, RD_{Early} remains roughly unchanged for CD cells with different input ranges.

For some values of N and L, RD_{Early} became negative for wide range (>0.5 octave) CD cells with moderate thresholds ([N, L]=[4, 2], [10, 3], [16, 3], not shown). In general, RD_{Early} did not become negative for CD cells with low thresholds (L<<N) or high thresholds ([N, L]=[16, 4]). If the threshold is very low, the cross-frequency CD cell essentially behaves as a spatial integrator. Because the early portion of the AN inputs respond to the broad transition stimulus when FT is both near and away from CF while the response to the sharp transition stimulus only occurs when FT is away from CF, the low threshold CD cell is expected have an early preference for the broad transition stimulus. Conversely, if the threshold is too high such that the number of inputs giving an early response to the sharp transition width is less than L, the CD cell would not have an early response for the sharp transition stimulus. For a moderate threshold CD cell to have a negative RD_{Early}, the number of inputs with an early response to the sharp transition stimulus must exceed L, and these inputs must be better temporally aligned than for the broad transition stimulus. This appears to be the case in the AN example shown in Figure 3.2AC.

CD population – Temporal metrics

While the RD metric is useful for distinguishing between CD cells with narrow vs. wide input ranges, the normalized response duration ND is useful for distinguishing between

AN responses and CD responses. The ND histograms for CD cells and AN fibers are shown in the middle column of Figure 3.4. For both transition widths, CD cells tend to have shorter NDs than their AN inputs, often lasting just 1 or 2 cycles. For both AN and CD cells, the ND for the sharp transition stimulus tends to be longer on average than that for the broad transition stimulus, especially at the higher levels. This was not surprising due to the longer ringing in the stimulus waveform with the sharp transition.

At both levels, the peak width PW (right column of Figure 3.4) for both CD cells and AN fibers tends to be narrower for the stimulus with sharp transition compared to the broad transition stimulus. This was expected for cross-frequency CD cells because we had found in Chapter 2 that the sharp transition stimulus resulted in better temporal alignment of AN discharges across CF in the late response. The cross-frequency CD cells should convert this temporal alignment into a precisely timed large rate response. However, we also find narrower peaks for the sharp transition stimulus in the responses of AN fibers and same-frequency CD cells. This indicates that at least some of this effect observed in cross-frequency CD cells may be simply inherited from the AN. Responses to the sharp transition stimulus tend to have additional small late peaks compared to responses to the broad transition stimulus (Figure 3.2). These small peaks tend to be narrower than the larger early peaks and likely contribute to our finding of narrower peaks for the sharp transition stimulus.

We had expected for a CD cell with strict requirements (large L/N) to have shorter ND and PW than AN fibers. Indeed, as L was increased, CD cell responses had shortened NDs and narrower PWs regardless of input range (not shown). However, for the relatively relaxed set of parameters used in Figure 3.4 ([N, L]=[10, 2]), the PW of

cross-frequency CD cells sometimes exceeded those observed in AN fibers. This was more true for the responses to the broad transition stimulus than the responses to the stimulus with sharp transition. From Chapter 2, we observed that, compared to the sharp transition stimulus, the broad transition stimulus resulted in worse AN temporal alignment, in that responses of low-CF fibers had longer latencies than those of fibers with higher CF. In the response of a low threshold (L<<N) cross-frequency CD cell, the early portion of a peak may correspond to activity from its high-CF inputs, while the later portion of the same peak may correspond to its low-CF inputs. As the input range of the CD cell increases, more asynchronous activity is included, which effectively widens the PW of a low threshold CD cell. Because the broad transition stimulus excited more asynchronously than the sharp transition stimulus, this effect is expected to be even larger for the stimulus with broad transition. Therefore, PW can provide information about the CD input range as well as the threshold of the CD cell.

Responses of Cochlear Nucleus Units

We recorded from 110 single CN units (with CFs 151 – 2639 Hz) in 19 cats. In all units, we recorded the response to Huffman stimuli with FT set to the CF for two phase transition widths (r=0.85, r=0.95). Stimulus level was initially set to near threshold (see Methods) and increased in 10 or 15 dB steps. Table 2 summarizes the total number of units studied in each CN unit type, and the number of units whose responses most resembled AN fibers, narrow range CD cells, and wide range CD cells, based on the maximum likelihood approach described in Methods. As a control, the same test was performed on responses of 147 AN fibers, and 139 of the fibers returned correct predictions, while 8 fibers resembled CD cells, yielding an error rate of just 5%.

	AN (147)	Pri (40)	Pri-N (10)	PhL (31)	HiS (6)	On (3)
=AN	139	36	2	19	3	2
$\hat{w}$<1/3	5	3	0	9	2	0
$\hat{w}$>1/3	3	1	8	3	1	1

Table 2. Total number of units found to have responses to Huffman stimuli most like AN fibers, narrow range CD cells, and wide range CD cells.

Primary-like units

In response to short tone bursts at CF, the PSTHs of primary-like (Pri) units in the CN are characterized by a high onset response followed by a gradual decline to a steady discharge until the end of the stimulus. The top panel of Figure 3.5A shows the response of one phase locking Pri unit to 30 ms tone bursts at CF (1293 Hz) at 80 dB SPL. Pri units are associated with spherical bushy cells, which are known to receive only 1 or 2 excitatory inputs from the AN with large synapses called endbulbs of Held (Cant and Morest 1979, Sento and Ryugo 1989, Ryugo and Sento 1991). Due to the small number of convergent inputs and the secure synapses, the responses of these cells often resemble those of AN fibers and are unlikely to be sensitive to changes in the AN's spatio-temporal pattern.

Responses of the Pri unit to Huffman stimuli are shown in the bottom 3 panels of Figure 3.5A at different stimulus levels. The Pri responses resemble those of AN fibers (inset panels in Figure 3.2AC). The duration of the response lasts 3-4 cycles. Regardless of the stimulus level, the earliest peak has a preference for the broad phase transition, and the later peaks prefer the sharp transition. Using a maximum likelihood approach (see Methods), we found that the responses of this unit most closely resembles responses of an AN fiber rather than CD cell responses. These characteristics were common to almost all

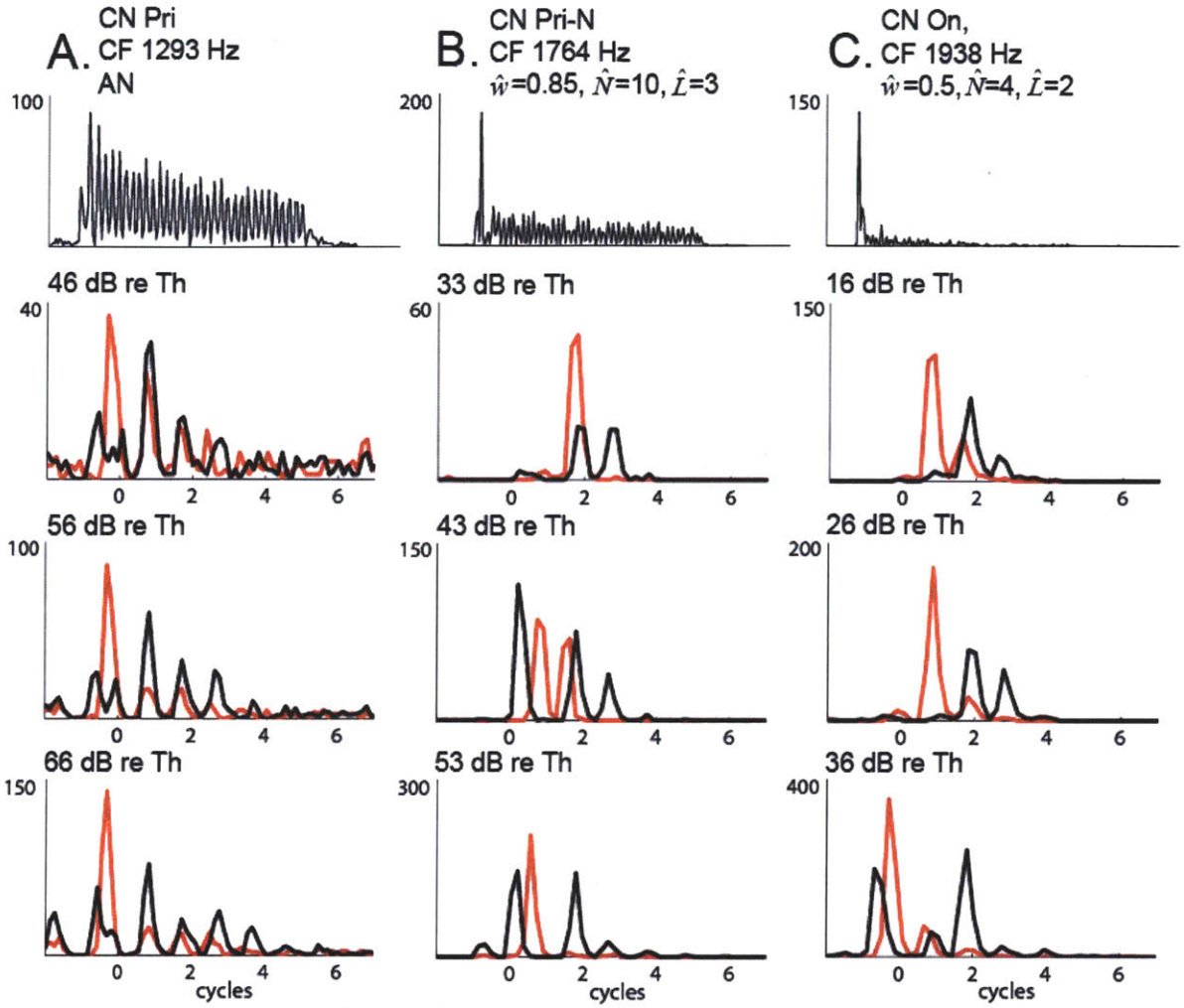


Figure 3.5: Responses of CN units. Top panels show responses to 30 ms tone bursts at CF at 80 (A), 70 (B), and 60 (C) dB SPL. 40 ms of the responses is shown. Bottom 3 rows show responses to Huffman stimuli with FT = CF. Broad transition is shown in red, sharp transition is shown in black. All y-axes are in spikes/second. **A.** Primary-like. **B.** Primary-like-with-notch. **C.** Onset.

of the Pri units in our population. Out of 40 Pri units studied, 36 had responses that matched most closely with AN responses, while only 4 units resembled CD cells. The panels in Figure 3.6A plot the distributions of RD_{Early} and RD_{Late} measured from the 40 Pri units. At low and high levels, both RD_{Early} and RD_{Late} are centered close to zero and are similar to those observed in the AN (Figure 3.4).

Out of the 4 Pri units that resembled CD cells, it is interesting to note that two of them had low CF (896 and 1078 Hz) and high synchronization indexes in response to short tone bursts at CF, meaning that they sharply phase lock. It is likely that these units form a continuum with the low-CF Hi-Sync units described later. The other two Pri units that resembled CD cells had fairly high CFs (2639 Hz, $[\hat{N}, \hat{L}, \hat{w}]$=[16, 5, 0.25], and 2750 Hz, $[\hat{N}, \hat{L}, \hat{w}]$=[16, 3, 0.25]). Interestingly, even though Pri units usually resemble AN fibers, the most likely N value for these two units was the largest value tested (N=16). The responses of one of these units (2639 Hz) displayed a short notch of decreased activity following the onset response to short tone bursts (not shown), but this decrease in activity was not sufficient to classify the unit as a primary-like-with-notch (Pri-N) unit

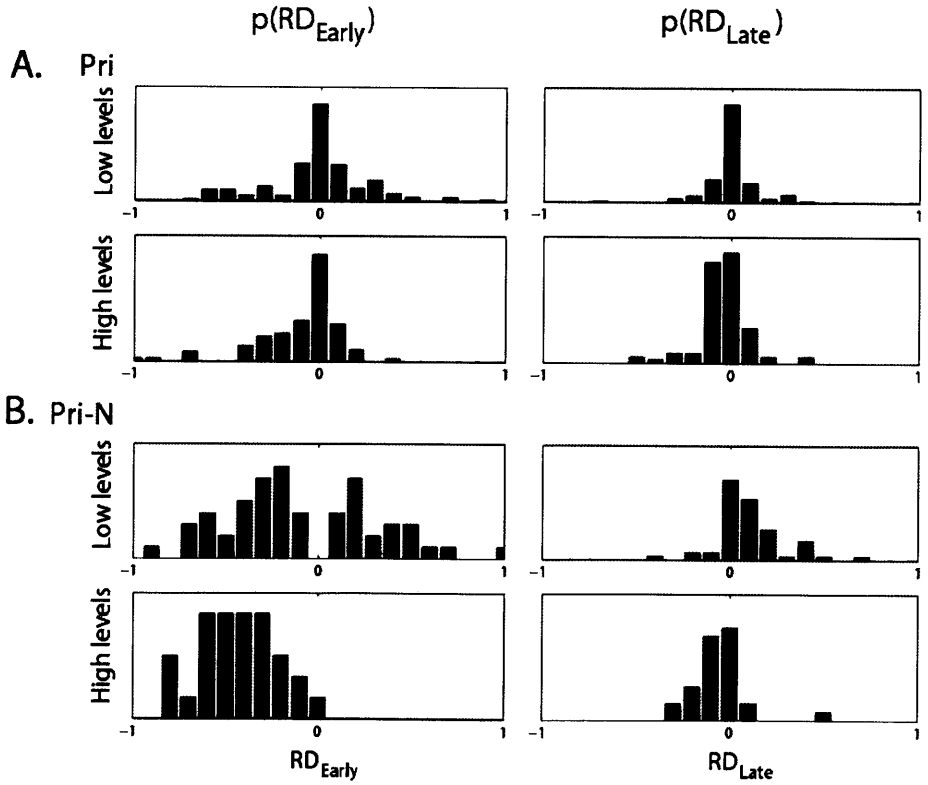


Figure 3.6: Distributions of early (left) and late (right) rate differences measured from responses from Pri (**A**) and Pri-N (**B**) units at low (top) and high (bottom) stimulus levels.

(see next section). These CD-like Pri units are examples showing how the unit types in the CN are not always discrete and frequently overlap. Rather, the continuum across different types gives rise to responses that may fit into several categories at the same time, or none at all. The maximum likelihood test may be a useful way to classify these units.

Primary-like-with-notch units

The top panel of Figure 3.5B shows the response of an example Pri-N unit to short tone bursts at CF (1764 Hz) at 70 dB SPL. Pri-N units are characterized by a high onset response, followed by a brief notch of inactivity (or decreased response) and recovery. These units are of particular interest because Pri-N response types are associated with globular bushy cells, which receive a large number of AN inputs (15-23 somatic inputs, Spirou et al. 2005, and more than 60 dendritic inputs, Liberman 1993) on modified endbulbs, which preserve the temporal firing pattern of AN inputs (Tolbert and Morest 1982ab, Oertel 1985). The notch of inactivity may arise from the refractory period of the cell's inputs after firing in synchrony to produce the high onset response at the output cell. This behavior is consistent with a CD mechanism that includes refractoriness (see Discussion).

Responses of the Pri-N cell to Huffman stimuli are shown in the lower 3 panels of Figure 3.5B. The normalized duration ND of the response is shorter (1-3 cycles) compared to that of AN and Pri responses (3-4 cycles). At the lowest level, the early peak is larger for the broad transition stimulus. At higher levels, there is a strong early response to the sharp transition stimulus. These observations are consistent with the

cross-frequency CD cells shown in Figure 3.2BD ([N, L, w]=[10 or 16, 2, 0.4]). The CD parameters that most closely matched this unit's responses were $[\hat{N}, \hat{L}, \hat{w}]$=[10, 3, 0.85].

The characteristics of the example Pri-N unit shown in Figure 3.5B were typical across our small population. The panels in Figure 3.6B show distributions of RD_{Early} (left) and RD_{Late} (right) measured from Pri-N units. At low levels, RD_{Early} is negative slightly more often than it is positive, and at high levels, RD_{Early} is almost always negative. As was described earlier, AN fibers and CD cells with low or high thresholds rarely exhibited negative RD_{Early}, but wide-range CD cells with moderate thresholds often had an early preference for the sharp transition stimulus. Using the maximum likelihood method, out of 10 Pri-N units studied, eight resembled CD cells with relatively wide input range ($\hat{w}$ >0.33 octaves). Out of these eight Pri-N units, two had $[\hat{N}, \hat{L}]$=[4, 2], four had $[\hat{N}, \hat{L}]$=[10, 2 or 3], and two had $[\hat{N}, \hat{L}]$=[16, 3]. Two units resembled AN fibers, and no Pri-N unit resembled a narrow range CD cell.

Onset units

In response to short tone bursts at CF, onset (On) units have a high onset response followed by almost no sustained activity. On units may be further subcategorized into On-I (no sustained activity), On-L (some sustained activity), or On-C (some regular chopping intervals at the response onset). Previous modeling work has shown that in responses like those observed in On units may be produced when CD is performed with a short time constant (~0.125 ms) and a large number (>100) of weak synaptic inputs (Kalluri and Delgutte 2003ab). On units are associated with several cell types, including

octopus, globular bushy, and stellate cells, which are all known to receive a larger number of AN inputs. Because of their wide tuning curves and very precise timing characteristics at the onset of their responses, On units are likely to receive input from a wide range of CFs.

The top panel of Figure 3.5C shows an example On-L unit with CF 1938 Hz at 70 dB SPL. In response to Huffman stimuli, the On-L unit displayed responses similar to the Pri-N example at even lower levels. At low levels (16 and 26 dB re Th), the early peak was larger for the broad transition stimulus, while the later peaks were larger for the sharp transition stimulus. At a higher level (36 dB re Th), an early response to the sharp transition stimulus appears, consistent with cross-frequency CD cells. This On unit was found to most closely match a CD cell with $\hat{w}$=0.5 octaves, $[\hat{N}, \hat{L}]$=[4, 2]. The other two On (both On-L) units were found to have responses to Huffman stimuli more similar to those of AN fibers than CD cells. However, the responses of these units to short tone bursts did not resemble the primary pattern observed in AN fibers.

To try to explain these apparent discrepancies, it could be the stimulus and methods used in this work are too restricted for describing the responses of these cells. The highest number of CD inputs examined in this work was 16, while cells giving rise to On-like patterns may receive a much larger number of inputs (>100). The CD model used here is very efficient for a relatively small number of inputs. However, the length of time to compute a CD response increases exponentially as N increases, so running our CD model with N>100 was not practical. Better models would likely involve taking into account the electrical properties of these cells and use a larger number of inputs (Kalluri and Delgutte 2003ab).

Furthermore, it is possible that the cells that give rise to On responses receive input from a wider range of CFs (Oertel et al. 2000) than considered in this study. Huffman stimuli are only able to manipulate the relative spatio-temporal pattern in the AN over a relatively narrow CF range. If a cell receives input from AN fibers outside of this range, its responses to Huffman stimuli would be more difficult to interpret. Thus the Huffman stimulus may not be the ideal stimulus for assessing the spatio-temporal sensitivity of On units.

Phase locker units

Low frequency (<1 kHz) CN units usually strongly phase lock to tone bursts at CF. This strong phase locking obscures the other features of the response pattern that are used to classify CN units. This makes it difficult to infer information about the cell's morphology based on its responses to tone bursts. One metric that might help to distinguish between cell populations at these low frequencies is the strength of phase locking. We classified most low-CF CN units into two groups depending on whether the synchrony index exceeds that observed in the AN or not (0.9). We call these high-synchrony (HiS, n=6) and phase lockers (PhL, n=31), respectively.

Figures 3.7AB show responses of two PhL units (CF 460 Hz in panel A, 871 in panel B). These examples demonstrate the variety of characteristics that units in the same group can exhibit. In response to short tone bursts at CF, the 460 Hz PhL unit shows a stronger onset response, which slowly decreases into high sustained activity. In contrast, the 871 Hz PhL unit in Figure 3.7B has a very strong onset response which rapidly decays to nearly zero over just a few cycles. Despite these differences in

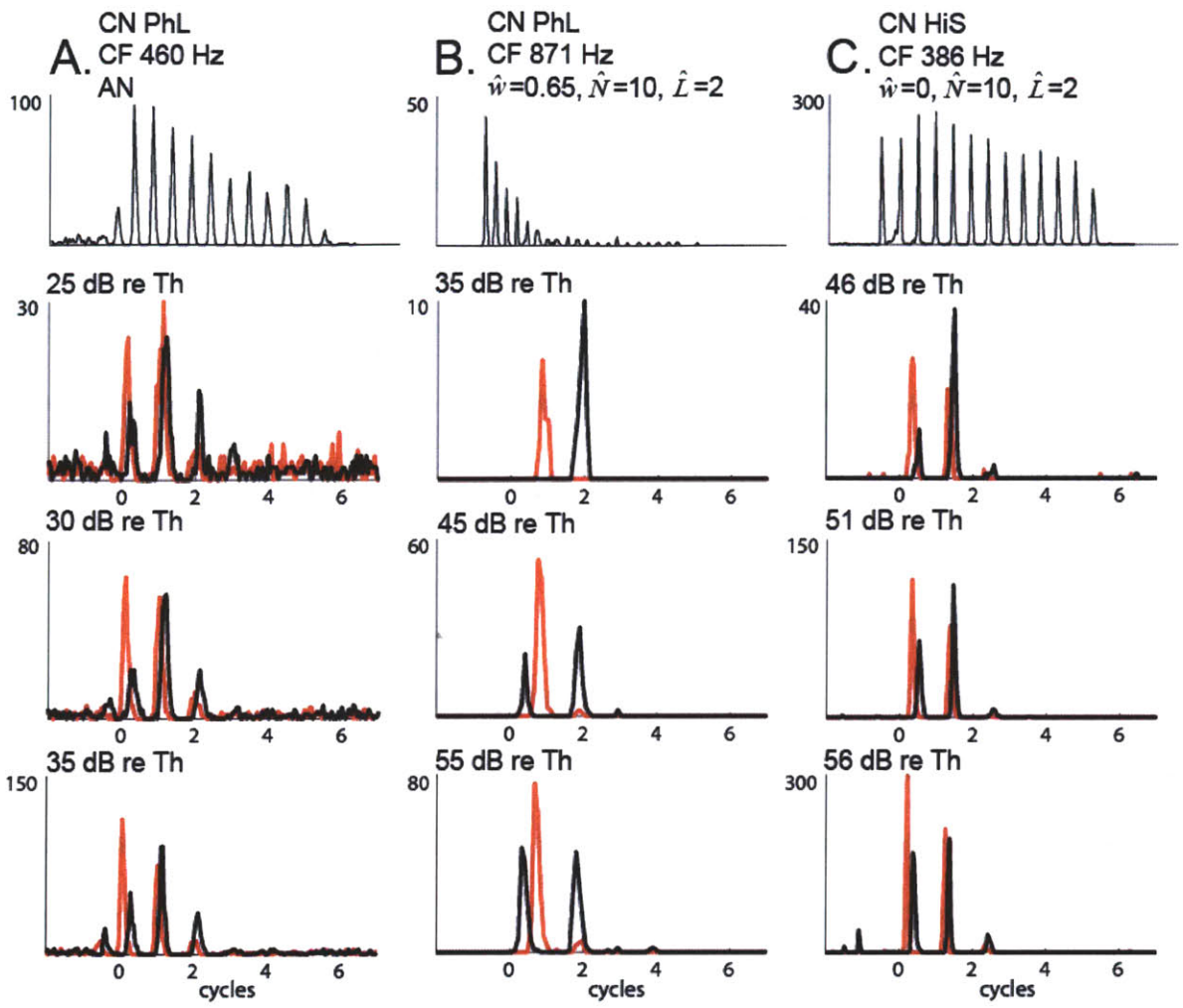


Figure 3.7: Responses of CN units, as in Figure 3.5. Top panels show responses to 30 ms tone bursts at CF at 65 (A), 80 (B), and 80 (C) dB SPL. **A.** Phase locker. **B.** Phase locker. **C.** Hi-Sync.

responses to short tone bursts, these units were both classified as PhL because they both exhibited phase locking within the range of AN, though the PhL unit in Figure 3.7B might also be classified as an On unit due to its large onset response.

The 460 Hz PhL unit in Figure 3.7A has responses to Huffman stimuli similar to those of AN fibers. The responses last for 3 cycles, and the early response has a strong preference for the broad transition stimulus at all 3 stimulus levels. In contrast, the 871 Hz PhL unit has a shorter duration, only displaying 1 peak at the lowest level and 1 or 2

peaks at the higher levels. Furthermore the early response to the sharp transition stimulus increases as level increases, consistent with a cross-frequency CD cell.

Out of our population of 31 PhL units, 19 had responses that most resembled AN fibers, 9 resembled narrow-range CD cells, and 3 resembled wide-range CD cells. Most (11) PhL units that resembled CD cells had $[\hat{N}, \hat{L}]$=[10 or 16, 2]. Our finding that some PhL units resemble AN while others resemble CD cells is not surprising considering that we cannot infer much information about these cells' anatomical features. The next section discusses an attempt to further subcategorize PhL units based on the ratio between early and late responses to tone bursts. However, we found no obvious correlation between this ratio and the best matching CD parameters for these PhL units.

Hi-sync units

Hi-sync (HiS) units comprise the other category of low frequency (<1 kHz) CN units. These units are characterized by their very strong phase locking to short tone bursts at CF (Figure 3.7C), stronger than that observed in AN fibers. HiS units have been associated with both spherical bushy cells and globular bushy cells (Smith et al. 1991, 1993, Joris et al. 1994a).

Figure 3.7C plots the responses of an example HiS unit with CF 386 Hz. Its sharpened temporal responses result in narrower peak widths for both tone bursts and the Huffman stimuli. At all levels, the responses of the HiS unit to Huffman stimuli displayed two narrow peaks, with a positive RD_{Early}, and RD_{Late} slightly negative or near zero. The unit had responses most similar to a same-frequency CD cell with $[\hat{N}, \hat{L}, \hat{w}]$=[10, 2, 0]. Three out of the six HiS units most resembled AN fibers, while two

resembled narrow-range CD cells ($[\hat{N}, \hat{L}, \hat{w}]$ =[10, 2, 0] and [16, 5, 0.25]), and one resembled a wide-range CD cell ($[\hat{N}, \hat{L}, \hat{w}]$ =[10, 2, 0.65]).

Since both PhL and HiS categories had some units that were characterized as resembling AN fibers as well as some units that resembled CD cells, using the synchronization index alone was not sufficient to predict like-AN or like-CD responses. Another possibility is to use the onset-to-sustained ratio in response to the short tone bursts at CF, which is the ratio between the firing rate to the first 10 ms of the response and the firing rate to the later 20 ms. For the six HiS units, there was a correlation between this ratio and the most likely CD parameters. The three HiS units that resembled AN fibers had the lowest ratios, while units with higher ratios resembled CD cells with increasing input range. It should also be noted that two low-CF Pri units with high synchronization indexes (0.9) had high onset-to-sustained ratios (~3 to 4) and resembled CD cells. However, the sample size is small, and other unit types, such as PhL, Pri, and Pri-N units displayed no obvious correlation between best input range and onset-to-sustained ratios.

Chopper units

Chopper units in the CN have a distinctive chopping pattern in response to short tone bursts at CF, where the spikes tend to occur at regular intervals unrelated to the stimulus frequency. As a result, these units have poor phase locking. Integrate-and-fire models (Laudanski et al. 2010, Hewitt et al. 1992) have been used to simulate responses of chopper units. In these models, the AN activity is summed over an integration window until threshold is reached. The neuron fires and the membrane potential resets to its

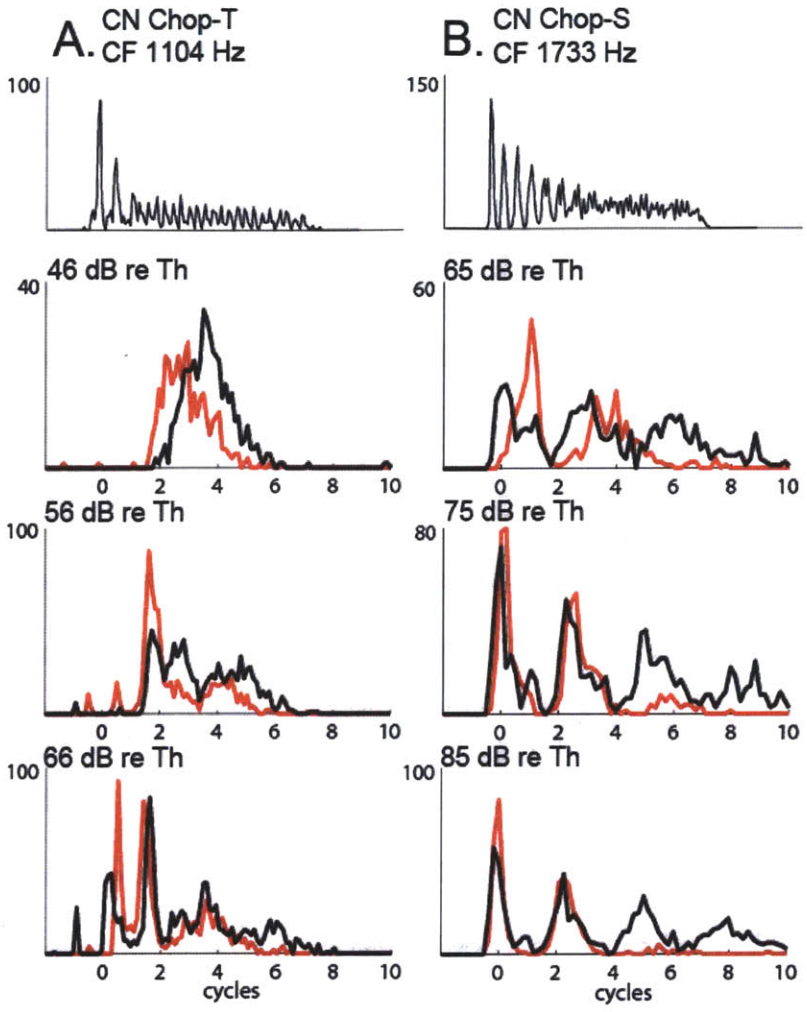


Figure 3.8: Responses of CN chopper units, as in Figure 3.5. Top panels show responses to 30 ms tone bursts at CF at 55 (A) and 60 (B) dB SPL. **A.** Transient Chopper. **B.** Sustained Chopper.

resting value (van Gisbergen et al. 1975). This summation has been shown to produce regular output patterns for irregular input patterns (Molnar and Pfeiffer 1968). These units are associated with stellate cells.

The coefficient of variation (CV) is a metric for irregularity in the unit's discharge pattern and is used to distinguish between two classes of choppers. Transient choppers (Ch-T) only display a chopping pattern during the early portion of the response, followed by irregular discharge (Figure 3.8A). Thus, the CV in Ch-T units will be low at the onset of the response, and will increase with time. In contrast, sustained choppers (Ch-S)

display regular chopping intervals throughout the 30 ms tone burst, and have lower CV values (Young et al. 1988).

Due to their tendency to produce spikes at intervals unrelated to the stimulus, it is unlikely that chopper units perform a coincidence detection mechanism as described here. However, stellate cells do receive multiple inputs from the AN in the form of small bouton endings (Wouterlood et al. 1984). In principle, there could be spatio-temporal sensitivity in the dendrites, even if phase locking is subsequently lost as the signals propagate to the cell body.

Responses of chopper units to Huffman stimuli tend to have very wide peaks, sometimes wider than a cycle of the CF. The Ch-T unit in Figure 3.8A has a long latency and a very wide uni-modal response at low levels. At the highest level, the unit's latency is shorter and displays some chopping for the first half of the response. The Ch-S unit in Figure 3.8B displays 2-4 wide peaks in response to Huffman stimuli, which are enhanced at higher stimulus levels. Unlike responses of other cell types, the spacing between peaks of Ch-T and Ch-S units is not related to the unit's CF (due to the poor phase locking in these units), and the average peak width is very wide, greater than 1 cycle.

Because CD cells strongly phase lock and do not produce regular chopping patterns, the maximum likelihood test was not performed on chopper units. It is unlikely for chopper units to have the very short time constants required for coincidence detection. However, the differences in the response patterns to the different transition widths indicate that at least some chopper units may be sensitive to the spatio-temporal pattern of AN activity.

Discussion

Inhibitory inputs

In this work, all inputs to the CD model cells were excitatory and from AN fibers. Many cells in the CN also receive inhibitory inputs that may come from the dorsal CN, stellate cells in the posteroventral and anteroventral cochlear nuclei (Oertel and Wickesberg 1993, Wickesberg and Oertel 1990), descending input from the superior olivary complex (Ostapoff et al. 1997), or from the contralateral CN (Needham and Paolini 2003, Babalian et al. 2002, Wenthold 1987).

While inhibition undoubtedly plays an important role in the responses of some if not all CN units (Gai and Carney 2008), we were interested in exploring a coincidence detection mechanism that operates directly on the spatio-temporal pattern of AN activity. All AN fibers provide direct excitatory input to cells in the CN. In contrast, the inhibitory inputs to the CN arrive indirectly from various locations, resulting in longer delays that are difficult to estimate. The responses of CN units to the click-like Huffman stimuli are short and decay after only a few cycles. It is not clear whether inhibitory inputs would play as important a role for responses to such transient stimuli as they do for longer stimuli. A sustained inhibitory input would likely raise the threshold of the output cell, effectively increasing the parameter L in the CD model used here.

Refractoriness in CD model

The CD model used in this work does not include refractory effects. Refractoriness refers to the period of time following an action potential, before a cell can fire again. For

Pri-N units that have very precisely timed onsets, the subsequent notch of inactivity may arise from the cell's refractory period.

Johnson and Swami (1983) described an iterative method for introducing refractory effects, which resulted in skewed peaks that have a sharper rise and a slower fall when applied to AN fibers. We implemented the method on CD responses to approximate the absolute refractoriness not included in the CD model. The changes produced by the model were mostly insignificant, especially for lower firing rates. For a small number of spikes, the effect of refractoriness is expected to be minimal.

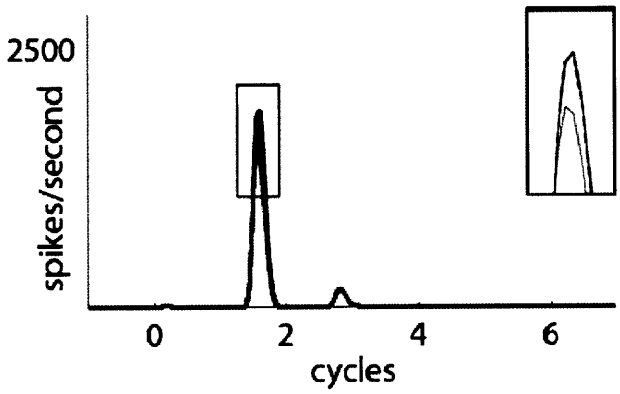


Figure 3.9: Effect of refractoriness in CD model. Black is the same as Figure 3.2D, N=4, L=2, w=0 for the broad transition stimulus. Gray plots the response after introducing refractory effects. Inset plots the zoomed in response at the largest peak.

In Figure 3.9, we show the effect of introducing refractoriness on an example CD response pattern (taken from Figure 3.2D, [N, L, w]=[4, 2, 0]). The original response is shown in black, and the response after taking into account refractory effects is shown in gray. Consistent with observations made by Johnson and Swami (1983), the effect of accounting for refractoriness results in peaks that are slightly skewed towards earlier times. As expected, the effect of refractoriness is largest for the higher peak, though the effect is still very small. Introducing refractoriness in all of the CD responses used in this study did not result in significant changes to the rate or temporal metrics used.

Furthermore, the characterization of CN units resembling CD cells or AN fibers based on their responses to Huffman stimuli remained unchanged after introducing refractory effects.

Comparison with Carney's (1990) study

Carney (1990) was first to use Huffman stimuli to evaluate spatio-temporal sensitivity in single CN units. She argued that the broad transition stimulus excited AN fibers more coincidentally across CF compared to the sharp transition stimulus. In Chapter 2, we observed that for both model fibers and real AN fibers, neither transition width consistently excited more coincidentally across CF in the early response, and that it is the sharp transition stimulus that excites more coincidentally across CF in the late response. Carney's criterion for spatio-temporal sensitivity in CN units was either a greater overall response to the broad transition stimulus, which she interpreted as consistent with a cross-frequency CD mechanism, or a larger qualitative difference between responses to the two transition widths than observed in AN fibers.

She found that Pri-N, Ch-T, and some HiS units preferred the broad transition stimulus over the sharp transition stimulus at low levels, and concluded that these observations were consistent with cross-frequency CD cells. Later (Carney 1992), she proposed a cross-frequency integrate-and-fire model with a short time constant of 0.05 ms to model coincidence detection. She used a Huffman stimulus with FT 1000 Hz to excite model AN fibers with different CFs (850-1150 Hz). Each input spike (from the model fibers) produced a temporary increase in postsynaptic voltage. When the membrane voltage reached a threshold of 1.25 times the amplitude of each input, the

output cell would fire and reset its potential to rest. The model cell responses resembled those of her CN units in that there was an overall preference for the broad phase transition stimulus compared to the stimulus with sharp transition, especially at low levels. The model she used is very similar to our [N,L,w]=[16,2,0.4] CD model, which also prefers the broad transition stimulus over the stimulus with sharp transition (Figure 3.2B) in the early response.

Carney used the overall rate difference to quantify sensitivity to the AN spatio-temporal pattern. We found it necessary to divide between early and late portions and also consider temporal characteristics in order to quantitatively determine whether the CN units resemble AN fibers or CD cells. To facilitate comparison with Carney's results, the images in Figure 3.10 show distributions of overall RD (RD_{Total}) for model CD cells (N=10, L=2) at low (panel A) and high (panel B) levels. Bar plots below each image show the distributions measured from AN fibers. Using only RD_{Total}, it is difficult to distinguish between AN fibers and CD cells, which both have slightly negative RD_{Total} at both levels. It is also difficult to distinguish between narrow range and wide range CD cells. In contrast, we showed that some of the six metrics used in this study RD_{Early}, RD_{Late}, and the temporal metrics systematically vary as the parameters of the CD model are varied. Using these metrics not only allowed the characterization of CN units as resembling CD cells or AN fibers, but also provided insight into the nature of the AN-to-CN convergence. For CN units that resembled CD cells, various CD parameters such as the input range and relative strictness of the CD model (relationship between L and N) were predicted.

We showed in Chapter 2 that, contrary to Carney's interpretation of the Huffman

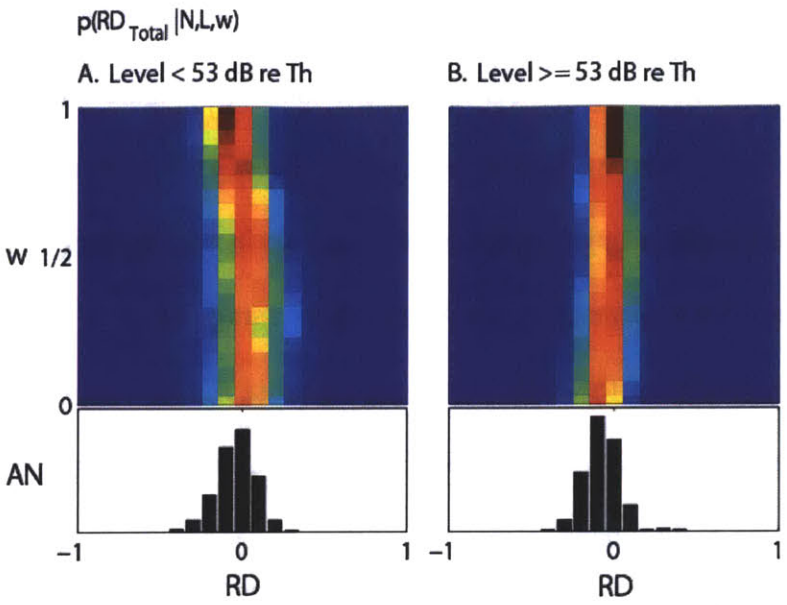


Figure 3.10: Distributions of RD_{Total} for CD cells (images) and AN fibers (bar plots) at low (**A**) and high (**B**) levels. As in Figures 3.2-4, N=10 and L=2.

stimuli, the late portion of the AN response is excited more coincidentally for the sharp transition stimulus than for the stimulus with broad transition. In several of her CN units (Pri-N, Ch-T, and HiS) that she labeled as spatio-temporally sensitive, she observed a preference for the broad transition stimulus at low levels. In contrast, we typically only observed an early (not overall) preference for the broad transition stimulus at low levels. This discrepancy might arise from a difference in the RD_{Total} measured from AN responses in the two studies. On average, at low levels, Carney's AN fibers tended to prefer the broad transition stimulus over the sharp transition stimulus by 14.4 spikes/second, while our AN population average RD_{Total} was just 4.8 spikes/second. Possible reasons for this discrepancy include differences in stimulus generation methods and have been explored in Chapter 2.

At high levels, Carney's spatio-temporally sensitive CN cells sometimes changed their preference, firing more for the sharp transition stimulus, though the responses at high levels were not used in determining spatio-temporal sensitivity. This increased

response to the sharp transition stimulus at higher levels is consistent with our observations in almost all of our Pri-N (Figure 3.5B and 3.6B), some On (Figure 3.5C), and some PhL (Figure 3.7B) units that we labeled as cross-frequency CD cells. A possible explanation for why cross-frequency CD cells would have an increased response to the sharp transition stimulus as level increases is given in the next paragraph.

Figure 3.11 shows a schematic of the early rate response of an AN fiber (whose CF is shown by the vertical dotted lines) as a function of FT. The top panel plots the rate response at a low level, while the bottom panel plots the response at a higher level. The early response to the broad transition stimulus (red line) generally does not change much with FT. In contrast, the early response to the sharp transition stimulus (black line) tends to decrease when FT is near CF, falling below the response to the broad transition stimulus (reflected in positive RD_{Early} AN values in Figure 3.4 when FT equals CF). The FT range over which the response to the sharp transition stimulus is smaller than the response to the broad transition stimulus is wider at low levels than at higher levels (Figure 3.2AC). We showed in Figure 3.2B that at low levels, cross-frequency CD cells tend to prefer the broad transition stimulus over the stimulus with sharp transition in their early response. At high levels (Figure 3.2D), the early response to the sharp transition stimulus was larger relative to the broad transition stimulus. This observation was in the same direction as the reversal of transition width preference at higher levels observed in many of Carney's spatio-temporally sensitive units. These observations suggest that the responses of CN units to Huffman stimuli at both low and high levels can be useful for determining their phase sensitivity.

So, despite the different analyses and a different interpretation of the way Huffman stimuli excite the spatio-temporal pattern in the AN, this study and Carney's study reach similar conclusions: that some CN units (particularly Pri-N) have responses consistent with those of cross-frequency CD cells.

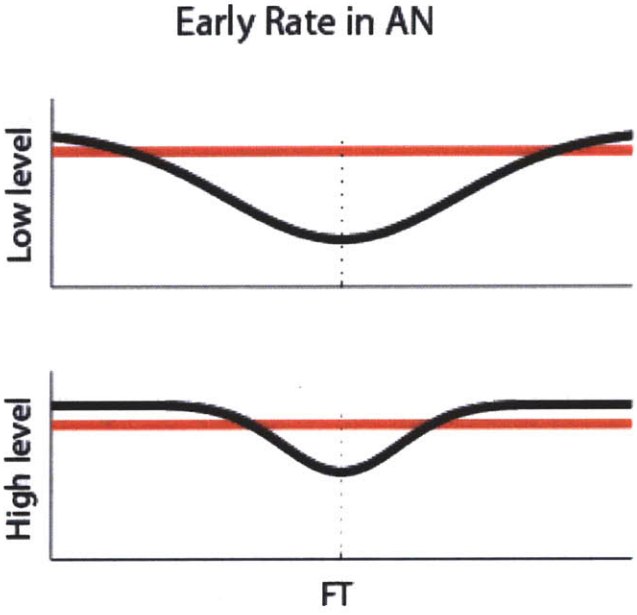


Figure 3.11: A schematic illustrating how the early firing rate in the AN varies with FT for a single fiber whose CF is shown by dotted line. Response to broad transition stimulus shown in red, and response to sharp transition stimulus shown in black. Top panel plots the rate at a low level, and bottom panel plots the rate at a higher level.

A different way to construct the virtual spatio-temporal pattern

The method for constructing virtual spatio-temporal patterns from the simple AN responses on which the CD cells operate were described in Chapter 2. Cochlear delays were accounted for by first removing the constant delay corresponding to the CF of the fiber, and then adding in a variable delay corresponding to the virtual CFs of the spatio-temporal response pattern. This was done so that the relative inputs to the CD model represent the true AN spatio-temporal pattern.

Our method contrasts with that of Carney (1992) who only subtracted the constant delay, and did not add the variable delay back into the pattern. Some CN units have large

responses to click stimuli, suggesting that these units may prefer to fire when AN responses across CF align with the cochlear delay. However, this method effectively assumes a compensation mechanism in which the AN-CN delay varies with CF so as to keep the overall delay constant across all CF. It assumes that high CF axons take longer than low CF axons to reach the CN, or that low CF inputs end closer to the cell body on the CN neuron, for which there is no evidence.

The analysis in this chapter was repeated using Carney's method of delay compensation, in which the AN spatio-temporal patterns had just the constant cochlear delay subtracted. We found that the relative pattern in the AN between the broad and sharp transition Huffman stimuli remained unchanged. The resulting probability distributions for the six metrics for the CD responses did not drastically change from what is shown in Figure 3.4. This indicates that the exact method of delay compensation matters less than the relative changes in the responses to the two phase transitions and the temporal characteristics imposed on the responses by the CD model.

Furthermore, most of the CN unit categorizations into AN vs. CD groups also did not change drastically. One exception is that using the Carney method of delay compensation, 4/10 Pri-N units resembled narrow-range CD cells (where there were none before, Table 2), 5/10 resembled wide-range CD cells, and one unit resembled the AN. All 4 narrow-range Pri-N units were most likely to receive inputs from either a 0.1 or 0.15 octave range. While all CD-like Pri-N units resembled wide-range CD cells using our delay compensation method, almost half of the Pri-N units resembled narrow-range CD cells with the Carney method. Several other CN unit types also had slightly increased numbers of units in the narrow-range CD category: 4 Pri, 13 PhL, 4 HS, 1 On.

Thus, the method for handling cochlear delay may change the relative proportion of wide-range vs. narrow-range CD cells, but does not change very much the classification as CD-like vs. AN-like.

Implications of cross-frequency coincidence detection

While most CN units do not have responses that resemble those of CD cells, the majority of Pri-N and some PhL and HiS units do have responses consistent with those of CD cells. Depending on the method of compensating for cochlear delay, these Pri-N units may resemble narrow-range or wide-range CD cells, but never same-frequency CD cells. This monaural cross-frequency CD mechanism results in a temporal sharpening across frequency channels that has functional implications for the processing of sound. One possible example lies in sound localization, in which binaural cues are used.

The cell types associated with these response patterns are globular bushy cells, whose axons provide strong inhibitory inputs (via the calyx of Held) to contralateral binaural neurons (Smith et al. 1991) in the medial superior olive (MSO) and lateral superior olive (LSO). MSO neurons are sensitive to interaural time differences and compare the timing of excitatory inputs (from CN spherical bushy cells) from both cochleas with some modulation by inhibitory inputs. The temporally precise inhibition provided by the globular bushy cells may arise from a monaural CD mechanism as described here, and appears to be necessary for the encoding of physiologically relevant ITDs (Brand et al. 2002).

LSO neurons are sensitive to interaural level differences through comparison between contralateral inhibitory inputs from globular bushy cells and ipsilateral

excitatory inputs from spherical bushy cells. At high sound levels, the cochlear filters widen, firing rates of AN fibers saturate, and the spatio-temporal pattern of activity across the AN gets more coincident across CF. Cross-frequency CD cells have higher thresholds and saturation levels than AN fibers and convert the degree of cross-CF coincidence in the AN into a temporally sharpened rate response. Thus, for LSO neurons comparing the sound levels at the two ears, if the monaural cross-CF CD cells are providing the inhibitory input, the temporal sharpening and conversion of spatio-temporal cues to rate cues likely play a role in the coding of interaural level differences. If the ipsilateral excitatory inputs saturate and the contralateral inhibitory inputs do not, the inhibition may become more potent at high levels. This may change the tuning of interaural level differences, possibly by increasing the dynamic range of these LSO neurons.

Summary

In this chapter, we computed responses of model cells performing coincidence detection on spatio-temporal patterns of AN activity. Based on our results from Chapter 2, we expected cross-frequency CD cells to have a larger late response to the sharp transition due to the better cross-CF synchrony elicited by this stimulus. However a late preference for the sharp transition was observed in both same- and cross-frequency CD model cells and in AN fibers, due to the longer ringing of the sharp transition stimulus compared to the stimulus with broad transition. We also observed that the range of FT for which the early AN response to the sharp transition stimulus is small compared to the response to the broad transition stimulus is wide at low levels, and narrower at higher levels. This

result was reflected in cross-frequency CD cells, which increasingly preferred the sharp transition stimulus at high levels in their early responses. The same was not observed for same-frequency CD cells or for AN fibers, which were less sensitive to changes in stimulus level. We then used a maximum likelihood framework to compare responses of CN units to CD model cells and identified the CD parameters that were most likely to give rise to the CN responses. We found that while most CN units had responses resembling those of AN fibers rather than those of CD cells, certain unit types (Pri-N and some PhL) had responses consistent with cross-frequency CD cells.

Appendix

The coincidence detection (CD) model (Krips and Furst 2009) used in this work analytically derived the instantaneous rate $\lambda(t)$ of a multiple input CD cell based on the instantaneous rates of its inputs. The CD model operates under several assumptions, including (1) the inputs may be modeled as non-homogeneous Poisson point processes and (2) the coincidence interval must be much less than the refractory period of the inputs. Both of these conditions are met: AN discharges are often modeled as Poisson process (Carney 1993), and we use a coincidence interval of 0.1 ms, which is much smaller than the 0.7 ms refractory period in the AN.

The model states that a CD cell with N independent inputs will fire if at least L of its inputs discharge within a coincidence interval Δ. We begin by presenting the instantaneous rate λ of a cell that fires when *exactly* l of its N inputs discharge within Δ. Let Ψ_{l_i} denote the set of l inputs that generate a spike within Δ, and the complementary set Ω_{l_i} describes the N-l inputs that do not spike. For each value of l ($L \le l \le N$), there exist $\binom{N}{l}$ possible combinations of inputs that can form Ψ_{l_i} and Ω_{l_i}, such that $1 \le i \le \binom{N}{l}$. The λ of the CD cell that spikes when *exactly* l of its N inputs spike within Δ is described by:

$$\lambda_{EE_{=l}^{N}(\Psi)} = \sum_{i=1}^{\binom{N}{l}} \lambda_{EE_l(\Psi_{l_i})} \lambda_{I(\Omega_{l_i})}$$

where $\lambda_{EE_l(\Psi_{l_i})}$ describes the instantaneous rate of a cell that spikes when all of its inputs Ψ_{l_i} spike within Δ, and is obtained by:

$$\lambda_{EE_L}(t) = \sum_{l=1}^{L} \lambda_{E_l}(t) \prod_{\substack{j=1 \\ j \neq l}}^{L} \int_{t-\Delta}^{t} \lambda_{E_j}(t')dt'$$

and $\lambda_{I(\Omega_{l_i})}$ describes the instantaneous rate of a cell that does not spike when any of its inputs Ω_{l_i} spike during the last Δ, and is obtained by:

$$\lambda_{I(\Omega_{l_i})} = \prod_{j=l+1}^{N} \left(1 - \int_{t-\Delta}^{t} \lambda_{E_j^{(i)}}(t')dt' \right).$$

The λ of the N input CD cell that fires when at least L of its inputs fire within Δ may be found by summing over all possibilities for l ($L \leq l \leq N$):

$$\lambda_{EE_L^N(\Psi)} = \sum_{l=L}^{N} \lambda_{EE_{=l}^N(\Psi)} .$$

CHAPTER 4

Rate representations of pitch in model coincidence detector and cochlear nucleus cells

Abstract

The pattern of activity across the auditory nerve (AN) contains phase cues to the pitch of harmonic complex tones. AN fibers tuned to frequencies near the stimulus components show rapid phase changes, and fibers tuned to frequencies in between stimulus components have more constant phases. In this work, we investigate whether a cross-frequency coincidence detection (CD) mechanism can extract these spatio-temporal cues by converting them into rate cues. We also investigated responses of cochlear nucleus (CN) units to harmonic complex tones, and interpreted these responses in light of whether they resembled cross-frequency CD model cells.

We found that, compared to the AN, model cross-frequency CD cells displayed enhanced rate cues to pitch at all levels. Responses of CN units to pitch stimuli differed according to unit type. Primary-like units tended to have degraded rate cues at high levels due to firing rate saturation, consistent with observations in AN fibers. Other CN units (primary-like-with-notch units and choppers) tended to have more robust rate cues with level, suggesting that these units extend the dynamic range of neural representations of pitch stimuli beyond that found in the AN.

However, there was no clear correlation between robustness in rate cues and similarity to CD cells as determined by responses to all-pass stimuli with phase manipulations in Chapter 3. It is likely that the CD model used is an oversimplification of the actual circuitry in the CN. However, our finding of several CN units that maintain robust rate representations of pitch at high levels is still relevant for understanding the neural mechanisms that encode pitch.

Introduction

In response to a harmonic complex tone, the pattern of activity of auditory nerve (AN) fibers across the tonotopic axis contains cues to the pitch of the stimulus. Fibers tuned to frequencies near the stimulus components show higher firing rates and rapid phase changes, whereas fibers tuned to frequencies in between stimulus components have smaller firing rates and relatively more constant phases. There are at least two possible mechanisms whereby central neurons can extract the phase cues in the AN. One possibility is a network of cross-frequency coincidence detectors (CD), in which neurons prefer to fire when their AN inputs discharge in synchrony. In Chapter 3, we have found some evidence for this type of processing in some units associated with globular bushy cells in the cochlear nucleus (CN). Our goal in this work is to investigate the implications that a cross-frequency coincidence detection mechanism may have on the representation of pitch of harmonic complex tones, and to assess the strength of pitch cues in the CN.

Another possible mechanism for extracting AN spatio-temporal cues is a lateral inhibitory network, which acts like a spatial derivative such that cells prefer to fire when their AN inputs across different characteristic frequencies (CF) fire *out of* synchrony (Shamma 1985b). Cedolin and Delgutte (2010) implemented the "mean absolute spatial derivative" (MASD) to model a lateral inhibitory network, which extracts these spatio-temporal cues by converting them into rate cues. To compute the MASD, they took the sum across time of the absolute difference between firing patterns of neighboring AN fibers. The resulting MASD profile had local maxima corresponding to rapid phase changes (integer values of CF/F0) and local minima when the phase is constant (halfway

between integer values of CF/F0). They found that a spatial derivative operation on AN activity resulted in a representation of pitch that was more consistent with psychophysical data than traditional representations. Specifically, they found that the spatio-temporal pitch cues were more robust with changes in level than rate cues, and predicted the upper frequency limit in the pitch perception of complex tones with missing fundamentals.

The lateral inhibition mechanism generates local maxima at locations corresponding to those of the stimulus components, consistent with the locations of AN fibers with higher firing rate. In contrast, the degree of cross-frequency coincidence in the AN is directly opposed to the firing rate of fibers at low levels. Because a cross-frequency CD cell depends on both the degree of cross-frequency coincidence as well as the firing rates of its inputs, it is unclear which characteristic plays the stronger role in determining the response of the CD cell. However, at high levels, when the firing rates of AN fibers are saturated so that there are minimal input rate cues, the degree of cross-frequency coincidence should dominate, and CD cells should generate local maxima in between locations tuned to those of stimulus components.

Previous studies of pitch in the CN have focused almost exclusively on temporal representations of pitch (Rhode 1995, Palmer and Winter 1992, 1993, Shofner 1991, 1999, Sayles and Winter 2007, 2008, Verhey and Winter 2006, Winter et al. 2001, 2003, Wiegrebe and Winter 2001), in which periodicity cues are reflected in neural phase locking. In contrast to the large body of work to understand temporal coding schemes in CN neurons, their rate responses to pitch stimuli have been less studied. Studies of vowel representation have demonstrated that some CN units have rate-place representations of

vowel spectra that are robust with increases in sound level (Blackburn and Sachs 1990, May et al. 1998).

In this chapter, we investigate the role of cross-frequency CD in the representation of pitch for harmonic complex tones and compare the responses of model CD cells to those of neurons in the CN. We find that model CD cells have enhanced rate cues to pitch at both low and high levels compared to those of AN fibers. Furthermore, the pitch estimates of cross-frequency CD cells remained stable with increases in level, unlike those of AN fibers. Responses of CN units to pitch stimuli differed by unit type. Those associated with globular bushy cells or stellate cells were more likely than other CN units to have robust rate cues and stable pitch estimates with increases with level. However, it is unclear whether cross-frequency CD plays a role in these observations, as there was no clear correlation between the characterization of these units as CD-like based on responses to Huffman stimuli and having robust rate cues and stable pitch estimates.

Methods

Neurophysiology and Stimuli

The methods of recording from single units in the auditory nerve (AN) and cochlear nucleus (CN) are well established (Bourk 1976, Pfeiffer 1966, Godfrey et al. 1975, Kiang et al. 1965b) and are the same as described in Chapters 2 and 3. Most of the CN units with lower CF (<2.6 kHz) studied in this chapter have been previously studied with Huffman stimuli in Chapter 3. The stimuli used in this study were the same as those described by Cedolin and Delgutte (2010). Harmonic complex tones consist of harmonics 2-20 with equal-amplitude and cosine phase. For each unit, the range of F0 was selected so that the neuron's CF would be near low-order harmonics (2, 3, and 4), which are likely to be resolved. Specifically, "neural harmonic number" (CF/F0) was varied between 1.5 and 4.5 in increments of 1/8. The F0 was stepped up and down over this range in order to minimize the effect of sensitivity to the order of notes. Each of the 50 (25 ascending and 25 descending) steps had a duration of 200 ms, including a 20 ms transition period where the waveform for the previous F0 step decayed and overlapped with the buildup of waveform for the next F0 step. Spikes occurring during this 20 ms transition were not used for analysis. Each 10 s stimulus was repeated 20 times with no interruption. Levels were typically initially set at 15-20 dB per component above pure tone threshold and then increased in 10 or 20 dB increments.

Analysis

Cochlear Scaling Invariance

To physiologically measure the AN spatio-temporal response pattern to harmonic complex tones across an array of real AN fibers would be very difficult, because it would require finely and uniformly sampling fibers with different CFs. We get around this challenge by applying the methods of Cedolin and Delgutte (2010) to use cochlear scaling invariance (Zweig 1976). In an invariant cochlea, scaling invariance states that the basilar membrane response $H(f,CF_0)$ at location CF_0 to a pure tone stimulus frequency f can be expressed as a function of the ratio $\frac{f}{CF_0}$. For another stimulus frequency f_1, if we let $\beta = \frac{f_1}{f}$, it follows that

$$H(f,CF_0) = H(\frac{f}{CF_0}) = H(\frac{f_1}{\beta \times CF_0}).$$

This means that the shape of the basilar response H at location CF_0 to frequency f is the same as the response at location $\beta \times CF_0$ to frequency f_1, and they are scaled (by β) versions of each other. Therefore, the basilar membrane response at a fixed location to stimuli with varying f_1 can be used to infer the response to a stimulus with fixed f at differing cochlear locations. This principle also holds true for every component of a harmonic complex tone, in which harmonicity is preserved in scaling by the same β.

The left panel of Figure 4.1 plots the spatio-temporal response pattern of an array of model AN fibers (Zilany and Bruce 2006) to a harmonic complex tone with F0 500 Hz at 60 dB SPL. CF varies from 750 to 2250 Hz and is shown on the right y-axis. The left y-axis plots the neural harmonic number $\frac{CF}{F0}$, and the x-axis is time normalized by F0. The right panel shows the responses of a single model fiber with CF 1500 Hz to complex

tones with varying F0 (333-1000 Hz) on the y-axis, so that the neural harmonic number is the same as the left panel. The nearly identical patterns between the left and right panels support the use of cochlear scaling invariance to infer the true AN spatio-temporal pattern from the responses of a single fiber to stimuli with varying F0.

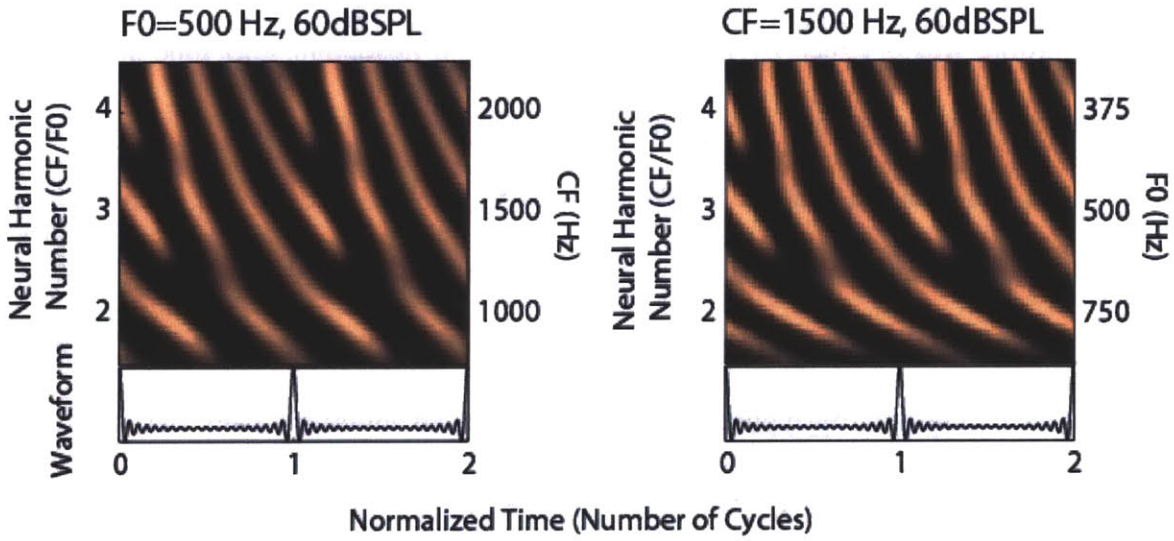


Figure 4.1: Responses of model AN fibers to complex stimuli. Stimulus level was 60 dB SPL. Left panel: responses of an array of fibers (CF 750-2250 Hz) to a stimulus with fixed F0 at 500 Hz. Right panel: responses of a single model fiber (CF 1500 Hz) to complex tone stimuli with F0 333-1000 Hz.

We use "virtual-F0" to refer to the F0 of the hypothetical fixed F0 stimulus, which corresponds to 1/3 of the fiber's true CF. We use "virtual-CF" to refer to the CFs of the hypothetical array of fibers, and set virtual-CF to be the neural harmonic number (CF/F0) times the virtual-F0.

Constructing "virtual" spatio-temporal patterns

Spike times in response to each 180 ms (200 ms minus the 20 ms overlap) duration complex tone were recorded and used for analysis. Spike times from the same F0 from

the ascending and descending parts of the sequence were combined because there were no systematic differences between the ascending and descending parts. Period histograms (with 50 bins per period) in response to each complex tone were computed as a function of time and F0 (expressed in the dimensionless units of number of cycles (t x F0) and "neural harmonic number" (CF / F0), respectively). The average discharge "rate-place" profile was obtained by integrating each period histogram over time at each F0 and plotting in spikes/second as a function of neural harmonic number.

As was done in Chapter 3, the non-scaling conduction delay for each cochlear location (Carney and Yin 1988) was subtracted from the time of each spike. This was done to allow the time scale to be normalized across F0s so that the principle of cochlear scaling invariance may be applied. In CN recordings, an additional 1 ms was subtracted from spike times to account for latency between the AN and CN. Then the cochlear delays corresponding to the virtual-CF locations were added back to each time axis to obtain virtual spatio-temporal patterns of AN activity.

Coincidence detector model cells

We use the virtual spatio-temporal response patterns of AN fibers as described in the previous section as the inputs to a coincidence detector (CD) model to predict responses of a CD cell receiving AN inputs across CF. As in Chapter 3, we use an analytical probabilistic model of CD neural cells (Krips and Furst 2009) to model the responses of neurons that receive AN fibers as excitatory inputs. We systematically varied several parameters of the model, including the total number of input fibers (N = 4, 10, 16), the

minimum number of inputs (L) that need to fire within the coincidence interval (0.1 ms) to generate a spike in the output cell, and the range of their CFs (w in octaves).

As in AN fibers, we expected the rate profiles of CD cells to show oscillations related to the presence of harmonics. Depending on the input range and level, we expected local maxima to occur in the rate profiles either on- or off-integer locations of the neural harmonic number. We use the strength of oscillation in the rate profile to characterize the strength of the rate cue to pitch. In order to get a reliable fit for the rate profile, and to reliably estimate the strength of oscillation, we required data over at least a range of two neural harmonic numbers. For a typical AN fiber in which responses to complex tones with neural harmonic number 1.5 to 4.5 were recorded, this meant that the widest input range w of a CD cell that we could measure would be dependent on edge effects. The typical maximum input range studied was just 0.5 octave, resulting in a range of neural harmonic number from 1.8 to 3.8 (for an input AN pattern with neural harmonic number 1.5 to 4.5).

Interpolating responses across F0

To implement the CD cell, we would ideally measure each AN fiber's responses in fine steps of F0 so as to sample the virtual spatio-temporal pattern with fine spatial resolution. In practice, due to limited holding times in the AN, the F0 was only varied in 1/8 "neural harmonic number" steps. Therefore, we used dynamic time warping to interpolate between responses from two consecutive F0s to predict responses to a set of F0s at finer sampling (1/40 neural harmonic number).

An example of measured responses of an AN fiber (CF 2340 Hz) to complex tones is shown in Figure 4.2A. Panel B shows the interpolated responses for F0s between those recorded along with the original responses. The interpolation algorithm generates response patterns that change smoothly with F0.

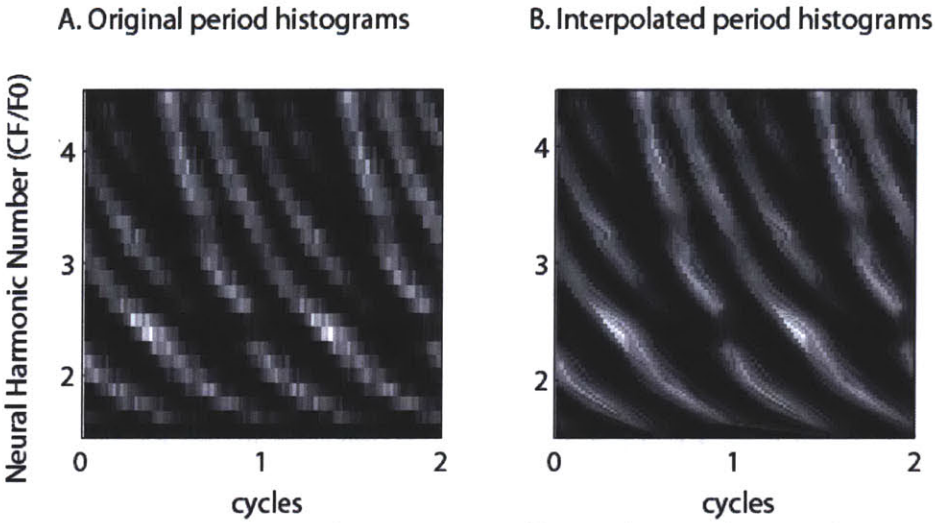


Figure 4.2: Responses of 2340 Hz AN fiber to harmonic complex tones at 42 dB re. Th. **A.** Recorded period histogram responses. **B.** Interpolated responses of the fiber for finer sampling of F0, derived from dynamic time warping analysis

Quantifying harmonic resolvability

The rate profiles of AN fibers, CD cells, and CN units were quantified following the methods of Cedolin and Delgutte (2010). To quantify harmonic resolvability, the following mathematical function was independently fit to each rate profile as a function of neural harmonic number using a least squares algorithm:

$$f(n) = A\cos(2\pi\phi n)e^{\frac{-n}{n_0}} + Be^{\frac{-n}{n_0}} + C$$

where the five parameters are: the amplitude (A), normalized frequency (ϕ), time constant of decay (n_0), amplitude of the baseline (B), and the DC component (C). This

function consists of an exponentially decaying sinusoidal component, an exponential term, and a constant term C. The parameter ϕ times the CF gives a rate-based estimate of the best frequency (BF), which characterizes the accuracy of the rate cue to pitch. We also fit each rate profile to a three-parameter model without an oscillatory component (A=0). The rate cue to pitch was only considered to be reliable if the five-parameter model provided a significantly (p<0.01) better fit than the non-oscillatory model based on an F-test for the ratio of residuals. To allow for either local maxima or minima to occur at integer locations of neural harmonic number, the amplitude A was allowed to be positive or negative.

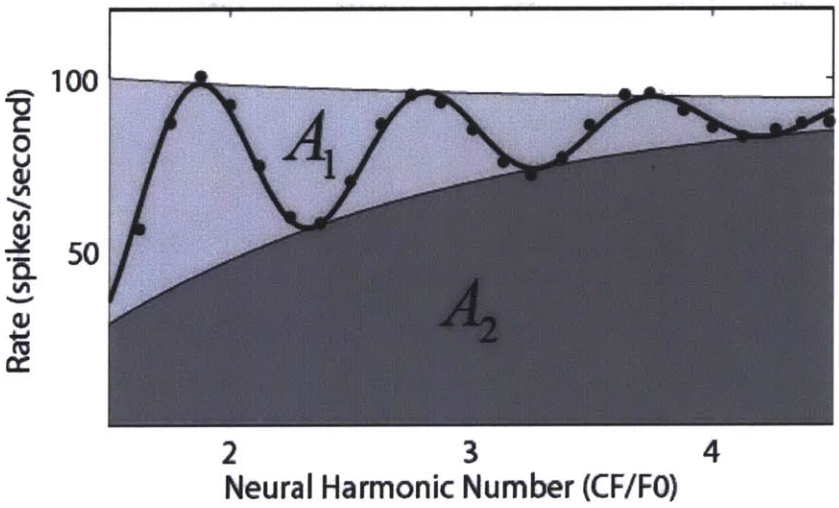


Figure 4.3: Curve fitting for rate profiles. Dots show data from an example rate profile (same as left panel in Figure 4.4). Solid line plots the fitted oscillating curve. Harmonic strength, or the strength of oscillation in the profile, was quantified by the area between the two exponential curves (A_1) normalized by the area under the top curve (A_1+A_2).

Figure 4.3 plots an example rate profile of an AN fiber (left panel of Figure 4.4) in spikes/second versus neural harmonic number n. The measured profile is shown in dots, and the solid line plots the fitted oscillating curve. Based on the values of the parameters of the best fitting curves, we derived a measure of the amplitude of the

harmonic oscillations in the spatial pattern. We call this the "harmonic strength" (HS_A) and define HS_A as the area between the two exponentials normalized by the area under the top exponential: $HS_A = \frac{A_1}{A_1 + A_2}$, which takes on values between zero and one. This metric is different from the one used by Cedolin and Delgutte (2010), who normalized the area between the two exponentials by the typical standard deviation in rate estimates measured from 100 bootstrap re-samplings. Because the CD model we used was based on an analytical derivation of the instantaneous firing rates, we did not work with raw spike times and thus could not use bootstrap re-sampling to estimate standard deviations.

Results

Responses of coincidence detector model cells

We have a database of 328 recordings from 171 AN fibers in 11 cats, most of which (279 recordings from 157 fibers in 9 cats) have been presented previously (Cedolin and Delgutte 2010). We additionally made 49 recordings from 14 fibers in 2 cats from which we also obtained responses to Huffman sequences. Among the entire set, 11 fibers had low spontaneous rates (SR<0.5 spikes/second), 38 had medium SR (0.5 to 18 spikes/second), and 122 had high SR (>18 spikes/second). The CFs of our AN fibers ranged from 301 Hz to 8881 Hz. The resulting estimated "virtual" spatio-temporal patterns were input into a coincidence detector (CD) model. In this section, we compare the responses of an example fiber with those of CD cells operating on AN activity. We then describe how CD responses vary with changes in the CD parameters.

Coincidence detector cell example

Figure 4.4 shows the responses of an AN fiber with CF 2340 Hz to harmonic complex tones with F0s 520-1560 Hz. Neural harmonic number n=CF/F0 is shown on the left y-axis of the left panel, F0 on the right y-axis of the right panel, and time normalized by 1/F0 is shown on the x-axes. The left image shows the response at a lower level (27 dB relative pure tone threshold at CF) than the right image (42 dB re. Th.). Rate profiles are shown in side panels to the right of each image, and the strength of oscillation, or harmonic strength (HS_A), is shown above each profile. At both levels, the fiber's temporal discharge pattern shows two peaks per cycle at neural harmonic number 2 and

three peaks per cycle near harmonic number 3, etc., indicating phase locking to the fine structure in the resolved harmonics. Furthermore, the relative phase changes rapidly near integer neural harmonic number (but slightly below), and is more constant halfway in between integer locations. These observations are consistent with the model predictions shown in Figure 4.1. At the lower level, the rate profile has moderately strong oscillations (HS_A=0.31) that decay as *n* increases. At the higher level, the fiber's firing rate saturates, resulting in decreased oscillations and a poorer rate cue to pitch (HS_A=0.12).

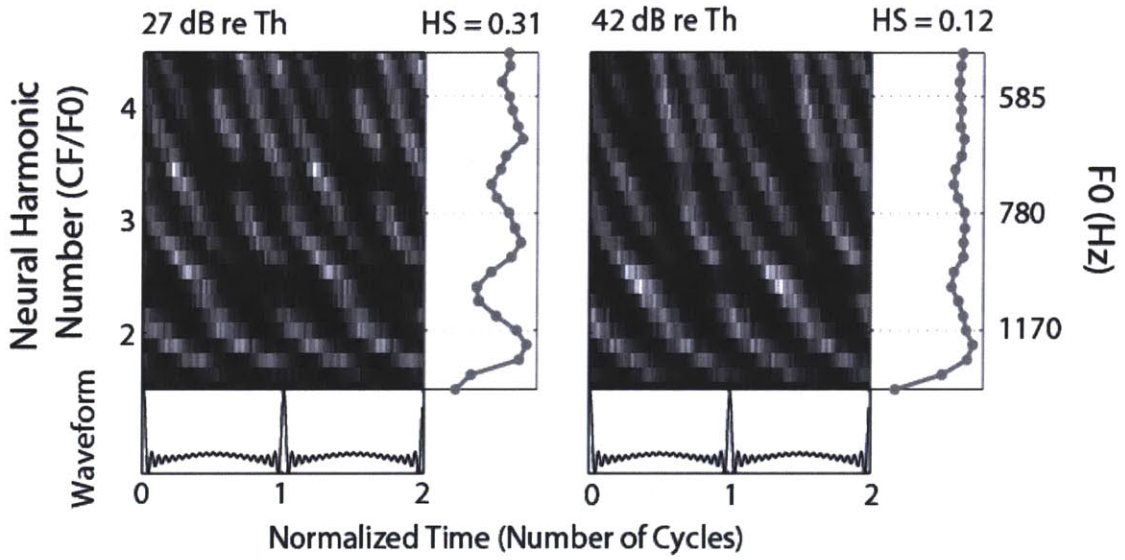


Figure 4.4: Responses to complex tones of a 2340 Hz AN fiber at two levels (27 and 42 dB re pure tone threshold). Right image is the same as the left image in Figure 4.2. F0 on the y-axis and time normalized by F0 on the x-axis. Right panels show the profile of average discharge rate.

This fiber's virtual spatio-temporal response pattern was used as input to a network of model CD cells with different virtual CFs. Figure 4.5 shows the spatio-temporal responses of CD cells with N=10, L=2, and w=0 (top), 0.25 (middle), and 0.5 (bottom) octaves. The neural harmonic number axis for the cross-frequency CD cells

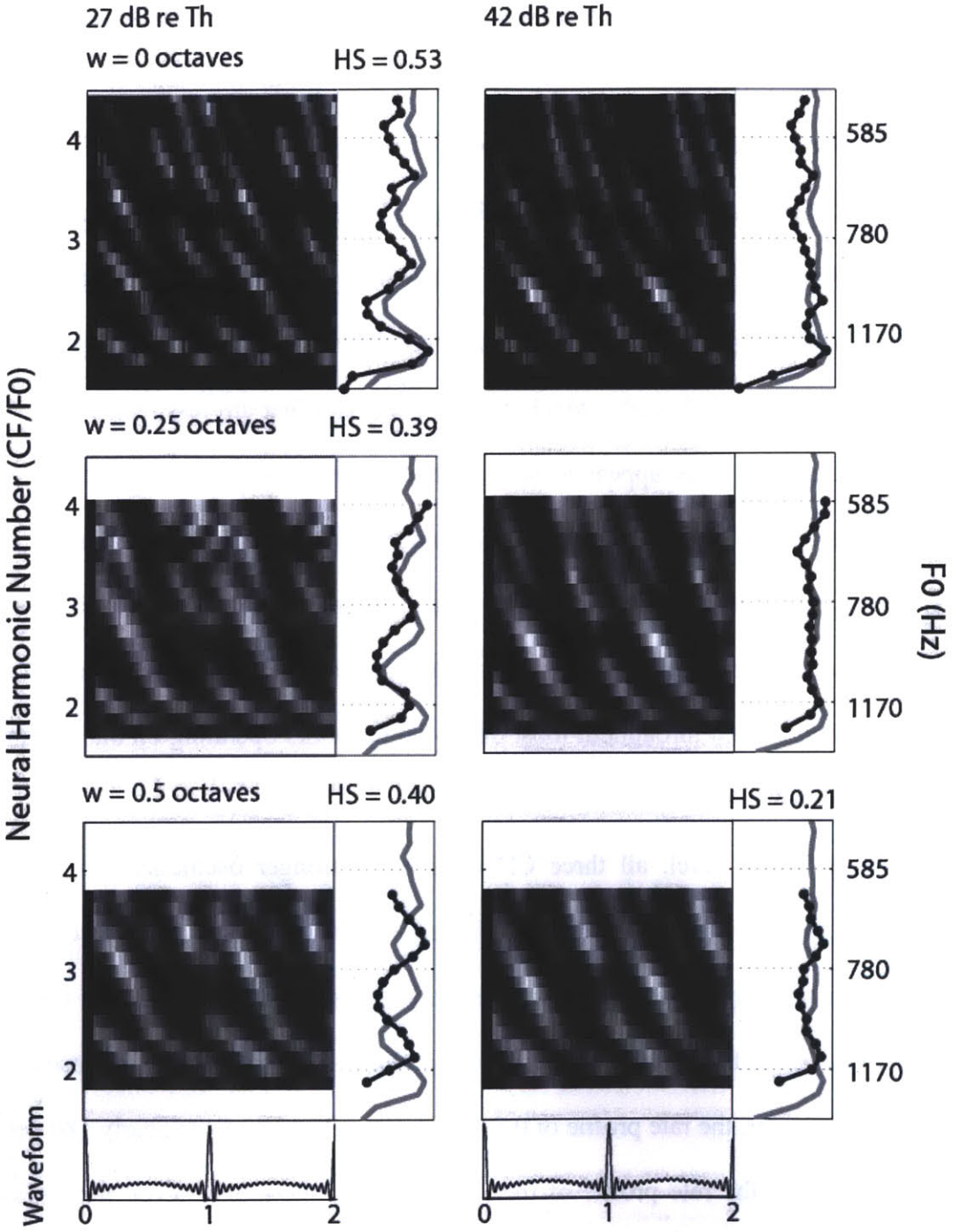


Figure 4.5: Responses of model coincidence detector cells operating on input from the AN fiber in Figure 4.4 at two levels (in dB re. pure tone threshold, columns) for three input ranges (top: 0 octaves, middle: 0.25 octaves, bottom: 0.5 octaves). Rate profiles are shown in the right panels (CD cells in black, AN fiber in gray).

does not extend over the full range of the AN y-axis because of edge effects. Left panels show the responses at the lower level, while right panels show the responses at the higher level.

Similar to the spatio-temporal pattern of its inputs, the responses of the same-frequency CD cell also display phase locking to the fine structure; however it shows temporal sharpening compared to the AN (as indicated by sharper peaks in the temporal patterns). In addition, the phase cues are preserved from the AN. Phase locking is poorer in the cross-frequency CD cells, which display responses that are less temporally sharp. Moreover, the phase cues appear to shift with changes in w. In the AN fiber, phase changed most rapidly when neural harmonic number *n* was slightly below integer values. In CD cells, the most rapid phase change occurs at integer locations for w=0.25, and slightly above integer locations for w=0.5. This systematic shift in phase cues with input range w was consistent throughout most other CD networks operating on our population of 171 AN fibers.

At the lower level, all three CD cells have stronger oscillations in their rate profiles than the AN fiber in Figure 4.4. Furthermore, while the same-frequency CD cell had a rate profile with maxima at the same locations as those of the AN fiber, rate profiles of the cross-frequency CD cells shows a phase shift with respect to those of their inputs. In particular, the rate profile of 0.5 octave CD cell seems to be nearly 180 degrees out of phase with the rate profile of the AN fiber. At the higher level, for the same-frequency and narrow range CD cells (w=0 and 0.25 octaves respectively), the oscillatory model was inadequate to fit the data (the oscillatory model did not provide a significantly better fit than the non-oscillatory model). In contrast, the rate profile of the w=0.5 octave

CD cell showed significant oscillations at the high level, and showed an enhanced rate cue to pitch (HS_A=0.21) compared to the AN fiber. Similar to observations at the lower level, the rate profile of the 0.5 octave CD cell was out of phase with that of the AN fiber. However, a difficulty with this analysis is that for the w=0.5 octave CD cells, when rate profiles could only be measured over a range of two neural harmonic numbers, it is hard to distinguish a change in phase from a change in best frequency (BF), which determines the frequency of oscillation in the rate profile.

CD population – Strength of rate cue

The panels in Figure 4.6 plot HS_A against CF (on the x-axes) for three ranges of level (in dB re. pure tone threshold, columns). Gray and black dots show AN and CD data (respectively) for which the parameter *A* was positive in the best fitting oscillating model, meaning that the best fitting curve had cosine phase. Pink and red dots show AN and CD data (respectively) for which the best fitting model had a negative *A*, meaning the best fitting curve was 180 degrees out of phase with cosine, or in anti-cosine phase. The CD parameters are N=10 and L=2, and w=0, 0.25, and 0.5 octaves (rows). The number of best fitting curves in anti-cosine phase is small in the AN, and for CD cells increases with increasing input range. In both AN fibers and CD cells, the harmonic strength improves as CF increases (Figure 4.6). This is consistent with better relative frequency selectivity, or relatively sharper cochlear filters, at high CFs (Kiang et al. 1965a, Liberman 1978). CD cells generally tend to have stronger oscillation in their rate profiles compared to AN fibers. This is especially true for same-frequency CD cells at low levels (top left panel) and cross-frequency CD cells at high levels.

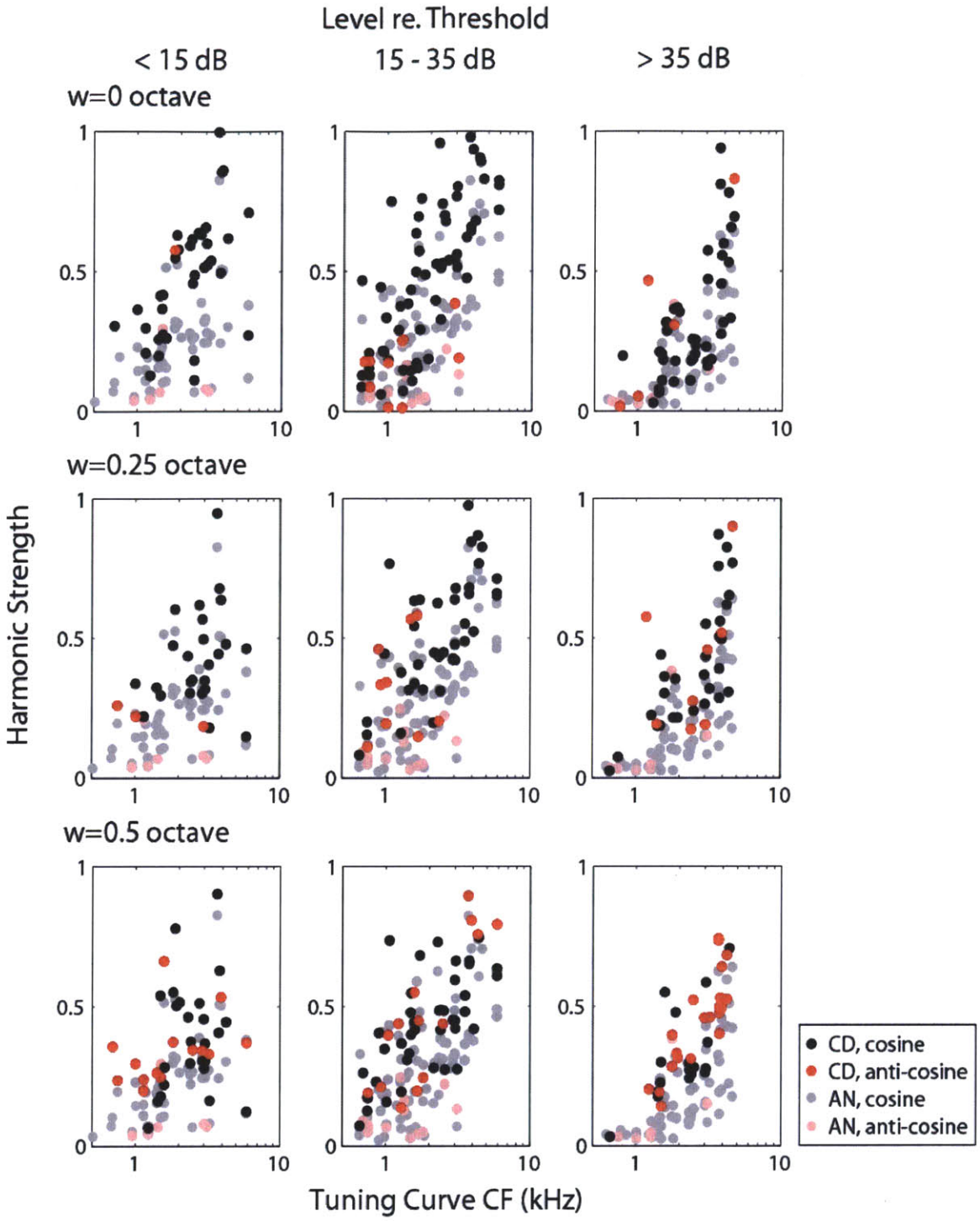


Figure 4.6: Harmonic strength of AN fibers (best fit with cosine phase: gray, best fit with anti-cosine phase: pink) and CD cells (cosine phase: black, anti-cosine phase: red) versus tuning curve CF (x-axes) and level (in dB re pure tone Threshold, columns). Top: w=0, middle: w=0.25, bottom: w=0.5 octaves.

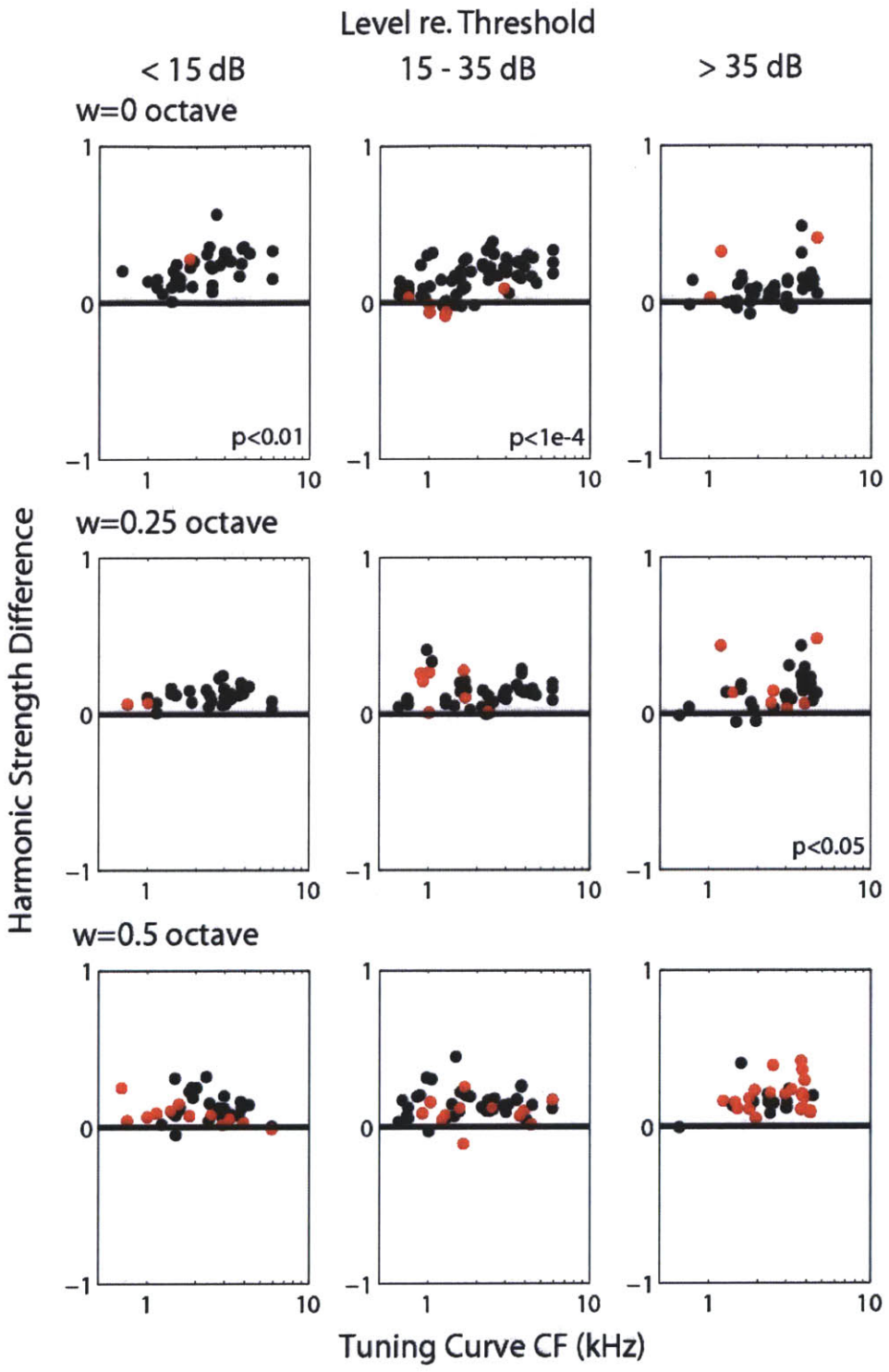


Figure 4.7: Strength differences between CD and AN responses as a function of the tuning curve CF (x-axes), level (columns), and input range (rows). Data for which the AN rate profile was in cosine phase and the CD profile was in anti-cosine phase are in red. All other data are shown in black. Positive values indicate that the rate profiles of CD cells have stronger oscillations than those of AN fibers.

To quantify the difference in oscillation strength between CD cells and their AN inputs, the harmonic strength difference $HS_{CD} - HS_{AN}$ is plotted in Figure 4.7 versus tuning curve CF (x-axes), level (in dB re. pure tone threshold, columns), and input range (rows). The strength difference tends to be positive for all CD cells, independent of level, CF, and input range, indicating that CD cells tend to have stronger oscillations in their rate profiles than their input AN fibers. Furthermore, the strength difference significantly decreases with increasing level for the same-frequency CD cell ($p<0.05$) and significantly *increases* with level for the w=0.5 octave CD cell ($p<0.05$). For the most part, CF does not have a strong effect on the strength difference, though there is a significant positive correlation between strength difference and CF for same-frequency CD cells at the lower and medium levels, and the w=0.25 CD cells at high levels (see p values in the bottom right of corresponding panels). These results suggest that while same-frequency CD cells display enhanced rate cues to pitch compared to the AN at low levels, this enhancement decreases at high levels. Moreover, the w=0.5 octave CD cells show rate cue enhancement over the AN at all levels, and enhancement improves at high levels.

CD population – Accuracy of rate cue

In addition to using HS_A to characterize the strength of the rate cue to pitch, we also characterize the accuracy of the rate cue by using the rate best frequency (BF), or the best fitted normalized frequency ϕ times the CF. The panels in Figure 4.8 show the relationship between the rate BF measured from the rate profiles and the tuning curve CF, plotted as a percentage of the CF. Data from AN fibers are plotted in the top row, while data from CD cells are plotted in the bottom three rows. The symbols are the same as in

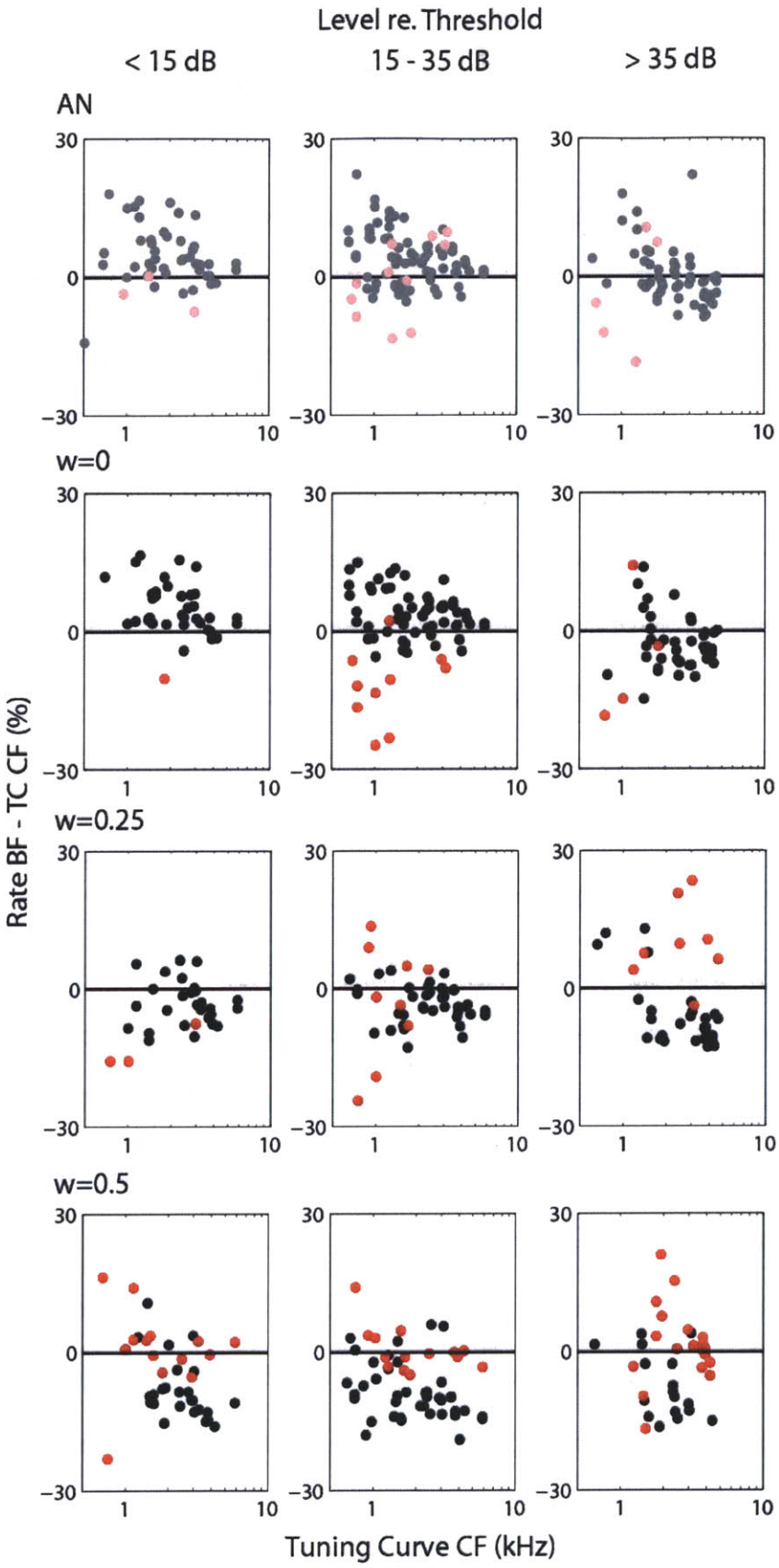


Figure 4.8: Relationship between rate-estimated BF and tuning curve CF for AN (top row) and CD cells (bottom three rows). Gray and black dots represent data for which the best fitted curve is in cosine phase. Pink and red dots represent data for which best fitted curve is in anti-cosine phase.

Figure 4.6. When the BF is equal to the CF for a gray or black data point (in cosine phase), the peaks of the rate profile are located at exactly integer values of the neural harmonic number *n*. In contrast, when BF is equal to CF for a pink or red dot (in anti-cosine phase), the peaks of the rate profile are located exactly halfway between integer values of *n*.

Few AN fibers (16/171) had best fitting curves in anti-cosine phase, as expected. For CD cells, the rate BF systematically varied with input range and whether the fitted curve was in cosine or anti-cosine phase. Same-frequency CD cells tended to have anti-cosine-phase fitted curves when rate BF was much smaller than CF, which often placed the peak locations near integer values of neural harmonic number (if the rate BF was 15% less than the tuning curve CF and the best fitting curve was in anti-cosine phase, the local maxima were located at 2.1, 3.0, and 3.8 neural harmonic numbers). This trend reversed directions for the w=0.5 octave CD cells, which tended to have anti-cosine-phase when rate BF was very close to or above CF. The w=0.25 octave CD cell tended to have more mixed results, with anti-cosine-phase fitted curves for small rate BF at low levels and large rate BF at high levels. As expected, at high levels (when the AN rate cues are weak), cross-frequency CD cells were more likely to have peaks in their rate profiles in between integer neural harmonic numbers, resulting in more anti-cosine-phase fitted curves than AN fibers or same-frequency CD cells.

At low levels, AN fibers and same frequency CD cells tend to have higher rate BFs than CFs. This trend reverses direction at higher levels, where the rate BFs tended to be smaller than CF. This result is consistent with AN rate BF estimates from Cedolin and Delgutte (2010). In contrast, the relationship between rate BF and CF for cross-

frequency CD cells (bottom two rows) was more level invariant: the rate BF tended to be less than CF at all levels, especially when the oscillations were in cosine phase. An analysis of covariance showed significant effects of CF ($p<1e-8$) and level ($p<1e-6$), but no interaction for the AN data. There was also a significant effect of level ($p<0.01$) for the same-frequency CD cell, but not for the cross-frequency CD cells. There was no significant effect of CF for any of the CD cells. CD cells with N=4 or N=16 yielded very similar results overall.

Responses of Cochlear Nucleus Units

We made 394 recordings from 132 CN units in 19 cats (with CFs 100-8580 Hz). These included 43 Pri, 8 Pri-N, 37 PhL, 1 HiS, 1 On, 20 Ch-S, and 22 Ch-T units. In this section, we compare the responses of CN units to those of AN fibers using the same quantitative metrics used for comparison between model CD cells and AN fibers: harmonic strength and rate best frequency. In Chapter 2, we used Huffman stimuli to assess whether each CN unit more likely resembles an AN fiber or a model CD cell. Here, we show results from CN units most resembling AN fibers, narrow-range CD cells, or wide-range CD cells separately to investigate whether a unit's responses to Huffman stimuli can be used to predict characteristics of the same unit's responses to complex tones.

Primary-like units

CN primary-like (Pri) units are associated with spherical bushy cells, which receive only 1 or 2 excitatory inputs from the AN with large synapses called endbulbs of Held (Cant and Morest 1979, Sento and Ryugo 1989, Ryugo and Sento 1991). Due to the small

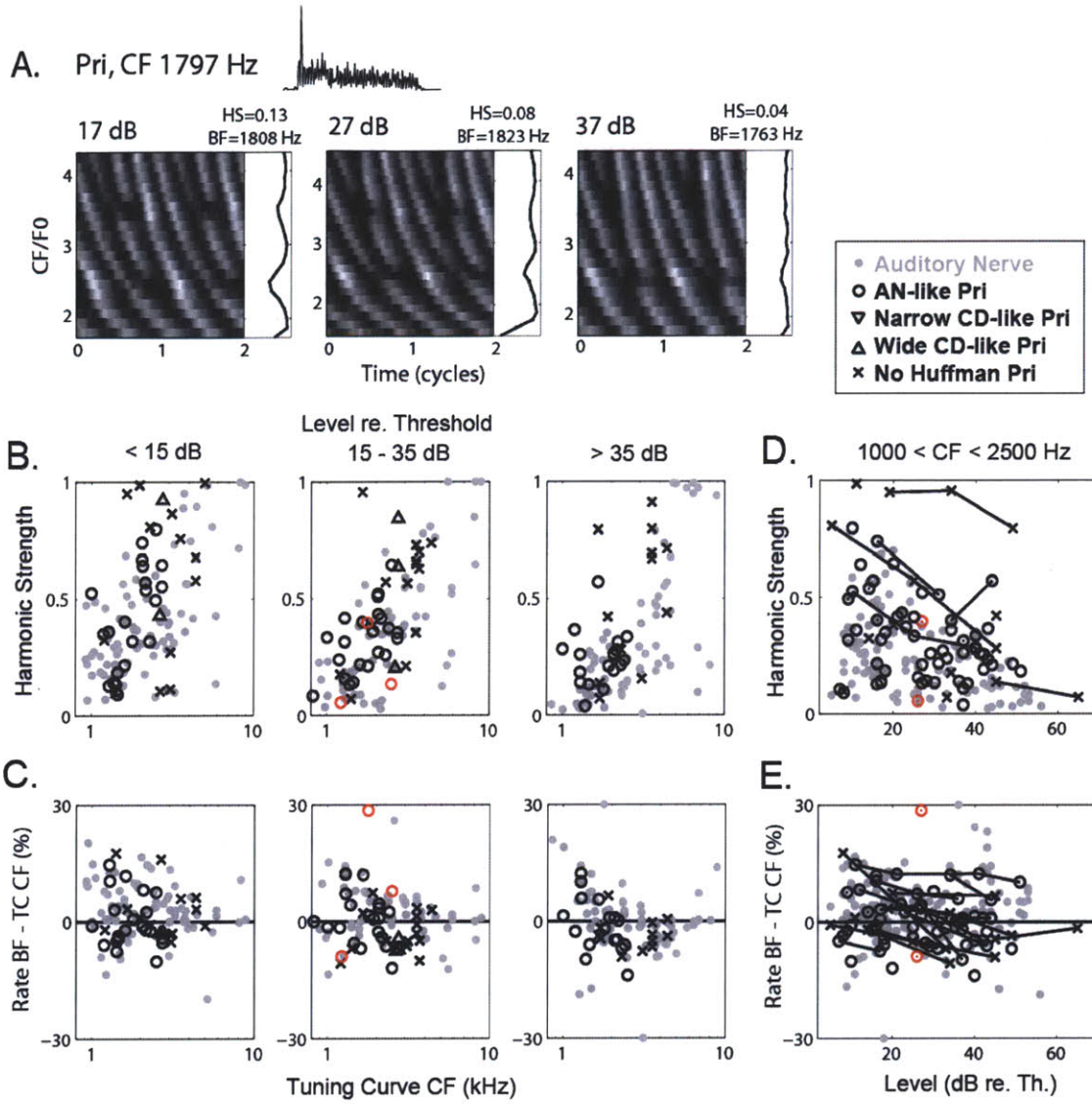


Figure 4.9: A. Example response of Pri unit at different levels (in dB re pure tone threshold). Inset shows the response of the unit to 30-ms tone bursts at CF (1797 Hz) used to characterize the unit. **B-E.** Most Pri units were further subcategorized into categories based on their responses to Huffman stimuli in Chapter 3: AN-like (open circles), narrow-range CD-like (w<0.25 octaves, downward facing triangles), or wide-range CD like (w>0.25 octaves, upward facing triangles). Data from Pri units for which no subcategory was assigned are shown in x's. Data from Pri units for which the best fitting curve was in anti-cosine phase are shown in red. All AN fiber data are shown as gray dots. The same legend applies to figures 4.10-4.11. **B.** Harmonic strength versus CF (x-axes) and level re. pure tone threshold (different panels). **C.** Rate BF – tuning curve CF relationship versus CF and level. **D.** Harmonic strength versus level for CF 1000 to 2500 Hz. **E.** Rate BF – tuning curve CF relationship versus level for CF 1000 to 2500 Hz.

number of convergent inputs, the responses of these cells often resemble those of AN fibers and we expected this to hold for harmonic complex tones as well. Indeed, in Chapter 3, we found that of the 40 Pri units examined, 36 had responses to Huffman stimuli that were more similar to AN fibers than CD cells.

Figure 4.9A plots the responses of a typical Pri unit to complex tones. The inset shows the unit's responses to 30-ms tone bursts at CF (1797 Hz) that were used to characterize the unit. In Chapter 3, this unit was determined to have responses to Huffman stimuli more similar to AN fibers than CD cells. As for AN fibers, the firing rate of the unit saturated at high levels, decreasing the strength of oscillation in the rate profile and degrading the rate cue. Moreover, as was observed in AN fibers and same-frequency CD cells, this unit's rate BF was smallest at the highest level (1763 Hz).

Panels B and C plot the harmonic strength (HS_A, B) and rate BF-tuning curve CF relationship (C) measured from the rate profiles of AN fibers (gray) and Pri units (red and black) as a function of CF (x-axes) for three different ranges of levels (in dB re. pure tone threshold). Panels D and E plot HS_A and the rate BF-tuning curve CF relationship, respectively, as a function of level for a narrow CF range (1000 to 2500 Hz). Pri units most resembling AN fibers (as determined by their responses to Huffman sequences in Chapter 3) are plotted as open circles, narrow-range CD-like units are plotted as downward facing triangles, and wide-range CD-like units are plotted as upward facing triangles. Pri units whose responses to Huffman sequences were not measured are plotted as x's. Data for which the best fitting curves were in cosine phase are shown in black, and anti-cosine phase are shown in red. Population responses of other CN unit types are plotted in the same way in Figures 4.10-4.11.

Not surprisingly, HS_A for Pri units fell within the HS_A range of AN fibers and showed similar trends with CF and level. Pri units tended to have increased rate cues at higher CF ($p<1e-6$) and degraded rate cues to pitch at high levels ($p=0.015$). Very few rate profiles measured from Pri units were in anti-cosine phase. Furthermore, the rate BF for Pri units tended to decrease with increasing level, also consistent with observations in the AN, though this trend was not significant ($p<0.1$). The one Pri unit that resembled a wide range CD cell had slightly stronger rate oscillations than AN fibers with similar CF at low and medium levels.

Phase locker units

Low frequency (<1 kHz) CN units usually strongly phase lock to tone bursts at CF. This strong phase locking obscures the other features of the response pattern that are used to classify CN units. Out of the 38 low-frequency CN units from which we recorded responses to complex tones, just one had stronger phase locking than observed in AN fibers, and 37 had phase locking within the range of low frequency AN fibers. We characterized these 37 CN units as phase lockers (PhL).

Figure 4.10A shows the responses of a PhL unit (CF 870 Hz) to harmonic complex tones. The responses of this fiber show poor phase locking to the fine structure in the stimulus, possibly due to poor cochlear frequency selectivity at low CFs. This is especially true at higher neural harmonic numbers, where the responses are phase locked to the stimulus envelope. The oscillation in the rate profile at the lowest level suggests that the first two harmonics in the stimulus may be resolved. Similar to the responses of AN fibers and Pri units, the rate profiles of this PhL unit at the higher levels had less

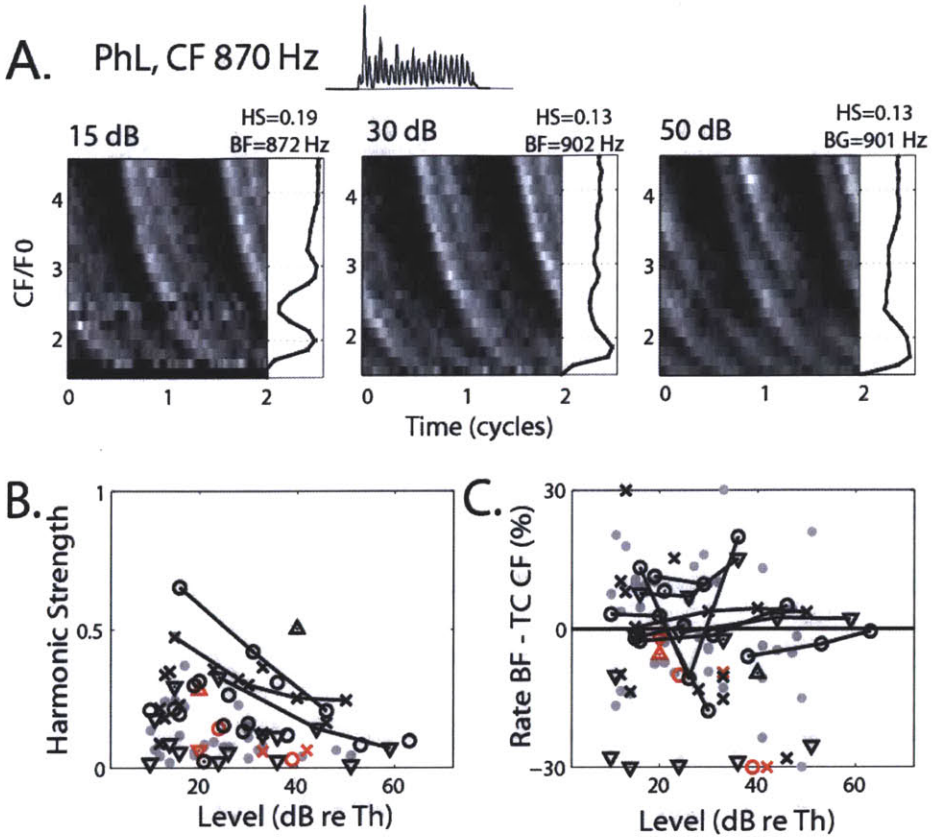


Figure 4.10: A. Example response of PhL unit at different levels. Inset shows the response of the unit to 30-ms tone bursts at CF (870 Hz) used to characterize the unit. **B.** Harmonic strength for CN PhL units and AN fibers versus level. **C.** Rate BF – tuning curve CF relationship versus level. Legend is the same as in Figure 4.9. Lines connect data points measured from the same unit at different levels. Data in panels B-C are shown for CF<1000 Hz.

pronounced oscillations than at the lower level. However, unlike AN fibers or Pri units, the degradation of this unit's rate cues at high levels does not appear to be due to a saturation of firing rates because there is still a clear global maximum near the second harmonic (CF/F0=2).

Figure 4.10B shows HS_A measured from our population of PhL responses. The symbols are the same as in Figure 4.9. Some PhL units have higher HS_A than low-CF AN fibers, regardless of their resemblance to AN fibers or CD cells based on responses to

Huffman sequences. PhL units were more likely to have a significant oscillatory component in their rate profiles than AN fibers. For levels above 35 dB re. pure tone threshold, only 3 out of 11 low-CF AN fibers had rate profiles that were significantly better fit with an oscillatory model compared to a non-oscillatory model. In contrast, 13 out of 19 PhL units had strongly oscillating rate profiles at these high levels. However, as for the example unit in panel A, a significant oscillatory component could be found when the rate profile shows just one peak. Effects of CF ($p<0.001$) and level ($p<0.01$) on HS_A were significant for PhL units.

Figure 4.10C shows the rate BF-tuning curve CF relationship for PhL units compared to AN fibers. These low-CF units tend to have larger differences between the rate BF and tuning curve CF compared to higher CF units (Figures 4.9, 4.11, 4.12), likely due to poorer relative frequency selectivity at these low CFs. Furthermore, for the example unit in Figure 4.10A at the two higher levels, even though the oscillatory model provided a significantly better fit than the non-oscillatory model, only one peak is clearly present in the rate profile. Without at least two peaks, it is very difficult to precisely estimate the rate BF. The relationship between the rate BF and tuning curve CF did not appear to depend on whether the PhL unit was more AN-like or CD-like. As with Pri units, few profiles measured from PhL units were in anti-cosine phase.

Primary-like-with-notch units

Primary-like-with-notch (Pri-N) units are characterized by a high onset response, followed by a brief notch of inactivity (or decreased response) and recovery. These units are of particular interest to us because the notch of inactivity may arise from a recovery

period after the cell's inputs fire in synchrony to produce the high onset response in the output cell. This pattern is consistent with a coincidence detection mechanism. From our results in Chapter 3, eight out of ten Pri-N units were determined to resemble wide-range CD cells, while the other two resembled AN fibers.

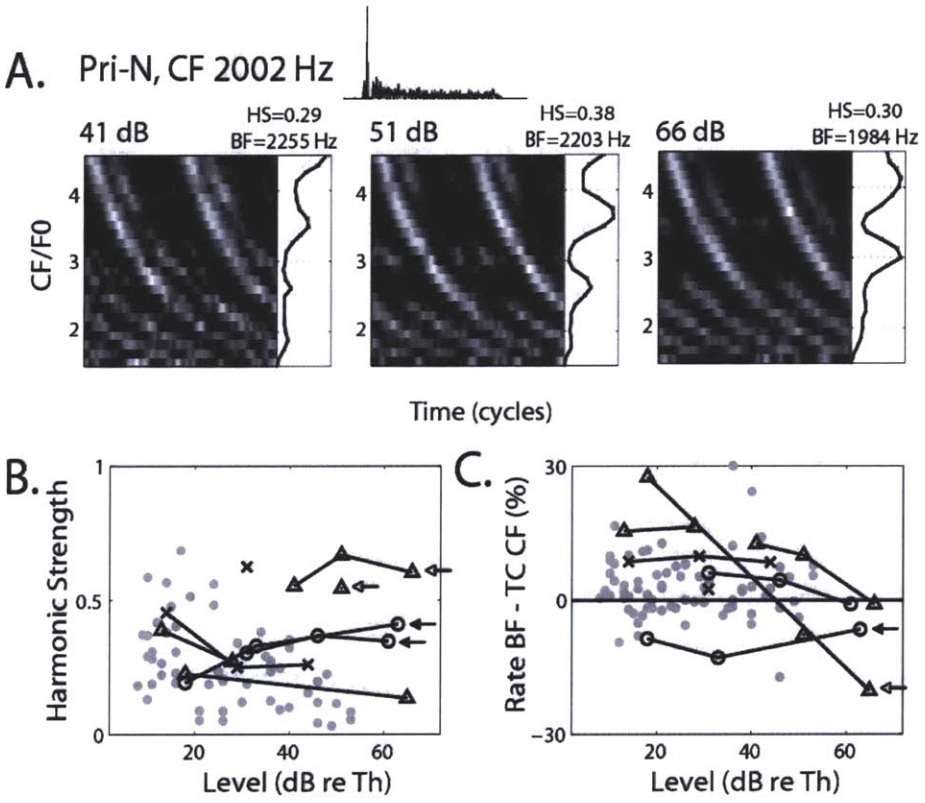


Figure 4.11: A. Example response of CN Pri-N unit at different levels. Inset shows the response of the unit to 30-ms tone bursts at CF (2002 Hz) used to characterize the unit. **B.** Harmonic strength versus level. **C.** Rate BF – tuning curve CF relationship versus level. Legend is the same as in Figure 4.9. Lines connect data points measured from the same unit at different levels. Data in panels B-C are shown for CFs 1400-2100 Hz.

Figure 4.11A shows the response patterns of an example Pri-N unit with CF 2002 Hz. From our analysis in Chapter 3, this unit most resembled a cross-frequency CD cell with input range 0.55 octaves. This unit phase locked to the fine structure for low neural harmonic numbers, and twice per period for high neural harmonic number. Furthermore,

the response phase of this unit does not obviously change near integer n vs. halfway between integer n, suggesting that unlike Pri units, this Pri-N unit poorly preserves the phase cues in the AN. In this unit, the strength of oscillation in the rate profile is largest at the medium level 51 dB re. Th., but does not greatly decrease at the highest level 61 dB re. Th. Furthermore, the peaks of the rate profile occur between integer values of the neural harmonic number at the lower two levels (41 and 51 dB re. Th.), and occur near integer values at the highest level (61 dB re. Th.). This led to rate BFs 13% and 10% above CF at the two lower levels 41 and 51 dB re. Th. respectively. Meanwhile, the rate BF at the highest level 61 dB re. Th. was just 1% below CF. Despite the occurrence of local maxima in between integer neural harmonic number in these profiles at the two lower levels, the best fitting curve was in cosine phase because the BF estimates were significantly different from CF.

All 8 Pri-N units had CFs between 1400 and 2100 Hz. This CF range was small enough that there was no significant change in HS_A or rate BF offset with CF, so these quantities are plotted against level in Figure 4.11BC. Due to the saturation of firing rates in the AN at high levels, the AN rate cue (and therefore HS_A) decreases as level increases. While most Pri-N units also displayed degraded rate cues at high levels, 4/9 units had stronger rate cues at high levels than observed in the AN. Surprisingly, two Pri-N units that were determined to be more AN-like in Chapter 3 had HS_A that increased with level, resulting in much stronger rate cues than observed in the AN at high levels (see black arrows in panel B). Out of the four Pri-N units that resembled CD cells from our results in Chapter 3, two had HS_A that fell within the range of the AN, and two displayed much higher HS_A than observed in the AN (see white arrows in panel B).

Out of the six Pri-N units for which responses to complex tones were measured at more than one level, five units displayed little relative change in rate BF with level. The rate BF of the remaining Pri-N unit was very sensitive to level (white arrow in panel C), with a rate BF 28% above CF at 18 dB and falling to -20% at 65 dB. From Figure 4.8, we observed that wide-range (w=0.5 octave) model CD cells have smaller rate BFs than their tuning curve CFs and were more likely to have profiles in anti-cosine phase than same-frequency model CD cells or AN fibers. However, in Figure 4.11C, we find that Pri-N units most resembling wide-range CD cells according to their responses to Huffman sequences actually have larger rate BFs than CF. One AN-like Pri-N unit had rate BFs that were consistently lower than CF (black arrow in panel C). Furthermore, none of the Pri-N units had profiles in anti-cosine phase.

Chopper units

Chopper units in the CN tend to have an intrinsic preferred chopping interval. This preferred interval is often unrelated to the stimulus or the unit's CF, and as a result, these units poorly phase lock. It is therefore unlikely that chopper units have the very short time constants required for coincidence detection, and in Chapter 3, the most likely CD parameters to give rise to chopper-like responses were not assessed. Chopper units are further divided into two categories: sustained choppers (Ch-S) display regular chopper intervals throughout their responses, while transient choppers (Ch-T) tend to only display a chopping pattern during the early response.

Figure 4.12A shows an example Ch-T unit's (CF 1760 Hz) responses to complex tones. This unit displays poor phase locking (at the middle and higher levels), and shows

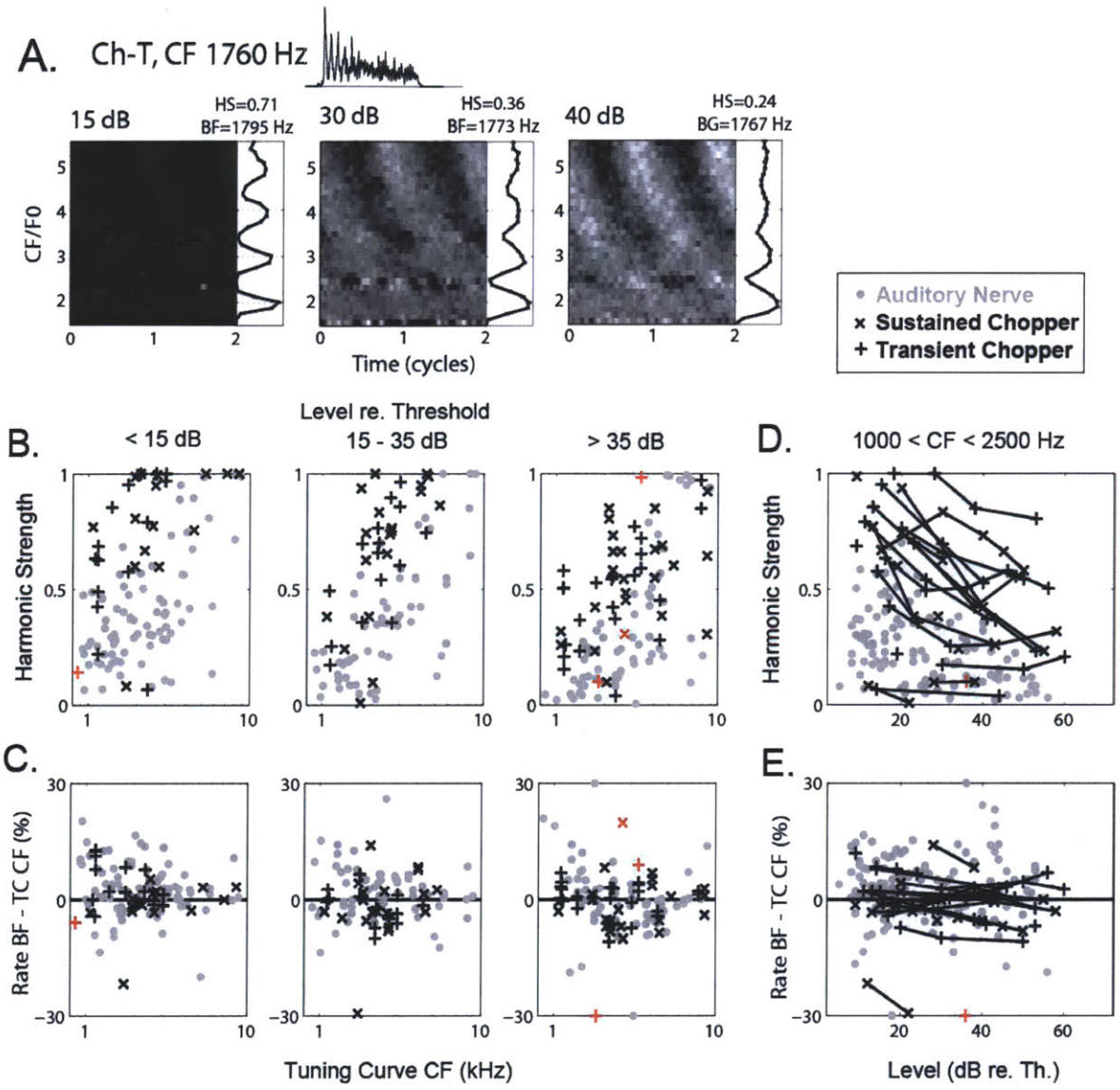


Figure 4.12: A. Example response of Ch-T unit at different levels. Inset shows the response of the unit to 30-ms tone bursts at CF (1760 Hz) used to characterize the unit. **B.** Harmonic strength versus CF (x-axes) and level (panels). **C.** Rate BF – tuning curve CF relationship versus CF and level. **D.** Harmonic strength versus level for CF 1000 to 2500 Hz. **E.** Rate BF – tuning curve CF relationship versus level for CF 1000 to 2500 Hz. Lines connect data points measured from the same unit at different levels. In panels B-E, data from Ch-S units are shown as x, and data from Ch-T units are shown as +. Data plotted in black means that the best fitting curve was in cosine phase, and red means anti-cosine phase.

very strong oscillations in its rate profiles. Even though the oscillation strength decreases as level increases, the oscillation is still quite strong at the highest level of 40 dB re. Th.

The panels in Figure 4.12B-E show the HS_A and rate BF-tuning curve CF relationship for Ch-T, Ch-S, and AN fibers. Panels D and E show data measured from choppers with CFs 1000 to 2500 Hz. At low levels, most choppers have stronger rate cues than those observed in the AN. At medium and high levels, the rate cues of choppers were degraded, though in most units the degradation was not as much as in AN fibers. The effects of level and CF were significant for both Ch-T and Ch-S units. No significant differences were observed in the rate cue strength between Ch-S and Ch-T units.

Almost all chopper units had rate BFs near their tuning curve CFs and within the range of those observed in the AN. Both Ch-S and Ch-T units had rate BF estimates that were largely level invariant.

Discussion

Comparison with the Cedolin and Delgutte (2010) study

Cedolin and Delgutte (2010) implemented a network of hypothetical lateral inhibitory (LI) neurons, which prefer to fire when their AN inputs across different CFs discharge asynchronously. For harmonic complex tones, these neurons essentially converted the AN spatio-temporal cues into rate cues by having local maxima at integer locations of neural harmonic number, where rapid phase changes occur in the AN. They found that this type of processing produced pitch cues that were more consistent with psychophysical observations by being more robust with level than rate cues in AN fibers. Furthermore, by relying on phase locking, the spatio-temporal representation predicted an upper limit to the pitch of the missing fundamental consistent with psychophysics (Moore 1973a), though a recent study questions the existence of this upper limit (Oxenham et al. 2011).

In this work, we examined a different type of neural mechanism: cross-frequency coincidence detection (CD), in which neurons prefer to fire when their inputs are temporally aligned because in Chapter 3 we found some evidence for this type of processing in some neurons in the cochlear nucleus (CN). In the spatio-temporal responses to complex tones of AN fibers, cross-frequency coincidence is stronger in between integer values for neural harmonic number, while the firing rates are higher at integer values. Thus, the coincidence cues and rate cues are opposed to each other, making it difficult to predict which has a stronger effect at low levels. However, at high levels, AN firing rates saturate, so that coincidence cues, which are relatively more robust with level, may dominate.

Responses of LI and CD neurons were similar in some respects in that they both produced rate cues relatively more robust with level than those of AN fibers, consistent with psychophysical observations. Furthermore, while rate-based best frequency (BF) estimates of AN fibers and same-frequency CD cells significantly decrease with level, both LI and CD cells had level invariant rate-BF estimates.

Despite these similarities, there were some important differences between the two types of processing. LI profiles rely on phase locking and showed degraded rate cues at high CFs. In contrast, while the CD operation enhances phase locking, the CD cells did not appear to require phase locking to have strong oscillations in their rate profiles. Thus the CD rate cues remained strong for CFs up to 5 kHz (Figure 4.7), indicating that the rate cues of CD cells are dominated by the rate cues in the AN at high CF rather than the degradation of temporal cues. The observations from LI cells are consistent with the common notion of an upper frequency limit to the pitch of complex tones with missing fundamentals (Schouten et al. 1962, Moore 1973b). However a recent study has found pitch perception even when all harmonics are above the currently accepted pitch region and limits of phase locking (Oxenham et al. 2011), suggesting that our CD results may not contradict psychophysical observations.

As in Chapter 3, we accounted for cochlear delays in the CD model by first removing the non-scaling delay corresponding to the CF of the fiber, and then adding in a variable delay corresponding to the virtual CFs of the spatio-temporal response pattern. This was done so that the inputs to the CD model best approximate the true AN spatio-temporal pattern and assumes no cochlear delay compensatory mechanism in which low-CF AN input activity would arrive earlier than activity from inputs with higher CF. This

method contrasts with that of Cedolin and Delgutte (2010) who only subtracted the constant delay, and did not add the variable delay back into the pattern when computing the response of the LI network. By performing computations on such patterns, they essentially assume a compensation mechanism in which high CF axons take longer than low CF axons to reach the LI network, or that the position of the endings on the dendrites of low CF inputs are closer to the cell body, for which there is no evidence in the CN.

To include a variable delay to these patterns corresponding to "virtual-CF" locations, we added longer delays for lower neural harmonic numbers, and shorter delays for higher neural harmonic numbers. This caused the phase changes to occur even more rapidly at integer neural harmonic number locations, especially for the lower harmonic numbers. This would likely have enhanced the spatio-temporal cues for the LI network described in Cedolin and Delgutte 2010. The CD analysis in this chapter was repeated without adding the variable "virtual" delay, and the trends illustrated in Figures 4.6-4.8 were mostly unchanged. One minor change was that the rate-BF estimate for narrow-range CD cells (w=0.25 octaves) acquired a significant dependence on level (p=0.047), with the rate-estimated-BF more likely to be above CF at low levels and below CF and high levels, causing these cells to behave more like same-frequency CD cells.

Scaling invariance

We applied the principle of cochlear scaling invariance for inferring the spatio-temporal responses of an array of AN fibers from the responses of a single fiber to complex tones with varying F0. In a completely invariant cochlea, the response at location CF_0 to a stimulus frequency f depends on only the ratio $\frac{f}{CF_0}$. One issue with the assumption of

scaling invariance is that the principle only holds locally, meaning it does not hold when CF_0 is far from f (Shera and Guinan 2003, van der Heijden and Joris 2006, Versteegh et al. 2011). In particular, the Q-factor (ratio of CF to tuning curve bandwidth) is known to increase with increasing CF (Kiang et al. 1965a, Liberman 1978), while scaling invariance assumes a constant Q across CF. A power law with exponent 0.37 fits AN data well in the cat measured with complex tones (Cedolin and Delgutte 2005) or pure tones (Shera and Guinan 2003, Shera et al. 2002).

In this study, F0 was typically varied over a narrow range: neural harmonic number 1.5 to 4.5 meant a range of 1.6 octaves around the "virtual" F0, or CF/3, meaning that the response at neural harmonic number 1.5 has a Q of $(1.5/3)^{0.37}$. This means that the sharpness in tuning was overestimated by 22%. Similarly, the sharpness in tuning at 4.5 was underestimated by 16% (Cedolin and Delgute 2010, Larsen et al. 2008). Therefore, the true rate-place profiles in the AN are likely to have broader peaks at low neural harmonic numbers and sharper peaks at high neural harmonic numbers than those estimated here.

These effects are likely to also be reflected in the responses of CD cells operating on the AN spatio-temporal pattern. In the responses of some CN units to complex tone stimuli, we observed smaller responses and poorer harmonic resolvability at lower harmonic numbers than at higher harmonics (Figure 4.11A). These observations are consistent with the overestimation of sharpness at lower harmonic numbers and underestimation of sharpness at higher harmonic numbers associated with scaling invariance described above. However, the analysis of CN responses also assume scaling

invariance, so it is unclear whether these observations are a reflection of the discrepancies of scaling invariance or a processing mechanism that emphasizes these higher harmonics.

If we assume scaling invariance in the CN such that we interpret the individual CN response patterns shown in Figures 4.9-12 as virtual spatio-temporal patterns of CN neurons across CF, we effectively assume a uniform array of CN cells across CF performing the same operation on their AN inputs. The operation performed by the CN cells would likely differ by cell type. While these assumptions are not inconceivable, it is unclear whether the CN circuitry is set up in this way. Regardless of the assumptions involved, the responses of CN units to varying F0 stimuli likely provide at least some insight for the characteristics of the true CN rate profiles.

Implications of cross-frequency coincidence detection for pitch stimuli

Unlike other CN unit types and AN fibers, some CN Pri-N and chopper units maintained robust rate cues to pitch and stable pitch estimates even at high levels (up to 80 dB SPL). Pri-N units are associated with globular bushy cells which provide temporally precise inhibitory inputs to contralateral neurons (Smith et al. 1991) that process binaural information. Pri-N units and cross-frequency CD cells sometimes produced local minima at integer locations of neural harmonic number, and local maxima in between integers (left and middle panels in figure 4.11, bottom left panel in Figure 4.5). If these cells provide inhibitory inputs to higher centers, then a possible function for these cells is to suppress neural activity away from integer harmonic locations at higher levels, and to enhance activity at integer locations.

Neurons in the lateral superior olive (LSO) compare their ipsilateral excitatory inputs from CN spherical bushy cells (whose Pri responses mostly resemble those of AN fibers) and contralateral inhibitory inputs from CN globular bushy cells. Similar to the rate responses of AN fibers, those of individual Pri units saturated at high levels. In contrast, some globular bushy cells have responses consistent with a cross-frequency CD mechanism, which maintains robust rate cues at high levels. Though pitch perception does not require binaural inputs, it is possible that the robust cues at high levels found in some Pri-N units serve to effectively increase the dynamic range of LSO neurons to pitch stimuli beyond that of AN fibers and CN Pri units.

CN stellate cells give rise to chopper response patterns (Cant 1981, Tolbert and Morest 1982ab, Rhode et al. 1983b), which show a wide dynamic range that might arise from inhibitory inputs from CN interneurons and the contralateral AN (Ferragamo et al. 1998, Babalian et al. 2002). In contrast to projections of bushy cells, those of stellate cells are more varied. These cells provide input to local circuits in the DCN (Palmer et al. 2003, Doucet and Ryugo 1997) as well as direct and indirect ascending pathways that converge in the inferior colliculus (Osen 1972, Adams 1979, Oliver 1987, Palmer et al. 2003, Smith and Rhode 1989).

Particular attention has been paid to temporal representations of pitch in chopper units (Verhey and Winter 2006, Winter et al. 2003). When the period of the stimulus is near the intrinsic chopping interval of these units, the unit's response patterns become more temporally regular (Winter et al. 2001, Wiegrebe and Winter 2001). Thus, the interspike-interval statistics of choppers may be important for temporal pitch processing,

and it has been hypothesized that these units may convert the AN temporal pitch cues into a temporal place code (Wiegrebe and Meddis 2004).

In addition, chopper units have robust rate representations of vowels at high levels (Blackburn and Sachs 1990, May et al. 1998), which is consistent with our finding that some of these units maintain robust rate cues to pitch at high levels. While these units are unlikely to perform CD, another mechanism called level-dependent "selective processing" has been proposed to explain the wide dynamic ranges observed in these cells (Lai et al. 1994). In a selective processing mechanism, responses of a chopper unit are determined by its AN inputs with high spontaneous rates and low thresholds at low stimulus levels and by its AN inputs with low spontaneous rates and high thresholds at high stimulus levels. Regardless of the exact mechanism that gives rise to these responses with wide dynamic range, it is likely that an important role of these neurons is to extend the dynamic range of periodicity encoding from that found in the AN.

Summary

In this chapter, we recorded responses of CN neurons to pitch stimuli and compared these responses to those of model cross-frequency coincidence detector (CD) cells operating on AN responses to harmonic complex tones. At high levels, the firing rates of AN fibers saturate, effectively degrading the rate cue to pitch. However, Cedolin and Delgutte (2010) found that a spatio-temporal representation of pitch provided more robust cues than rate cues at high levels. This representation was manifested in rapid relative AN phase changes near integer neural harmonic numbers, and more constant phase in between integer values. In theory, a cross-frequency CD model cell can extract these AN

spatio-temporal cues at high levels by producing local maxima in firing rate in between integer locations of the neural harmonic number. The intervals between the local maxima serve as the rate cue in these CD cells to the pitch of the stimulus.

We found that CD model cells displayed stronger rate cues than those found in the AN at high levels, but this was also true at low levels. We further found that unlike AN fibers and same-frequency CD cells, cross-frequency CD cells had rate-BF estimates that were invariant with level. Not surprisingly, responses of CN neurons to pitch stimuli differed by unit type. Similar to those of AN fibers, rate cues of Pri units degraded at high levels, and the rate-estimated best frequency (BF) decreased as level increased.

Low-CF CN units (PhL) also tended to have degraded rate cues at high levels, though this degradation was not always a result of saturation in firing rate. While the degradation of rate cues in AN fibers tended to be due to firing rate saturation, several fibers and some PhL units sometimes displayed a peak near the second harmonic, even when other harmonics were unresolved. Several of these PhL units were found to resemble CD cells in Chapter 3. However, it is difficult to interpret our finding that the rate responses of some low-frequency CN units were sensitive to the phase manipulations in the context of harmonic resolvability because the cochlea has poor frequency selectivity in cat and thus does not have strong spatio-temporal cues at these low frequencies.

We found that some Pri-N and most chopper units displayed relatively more robust rate cues at high levels compared to AN fibers, but it is unclear whether this result arises from a network of CD cells as described in this work. CN chopper units exhibit poor phase locking and were not expected to extract the spatio-temporal pitch cues in the

AN as described by the models used here because they lack the short time constants required for a CD mechanism. Moreover, the Pri-N units that displayed robust rate cues to pitch were not always those that most resembled cross-frequency CD cells as evaluated with Huffman sequences in Chapter 3. It is possible that Huffman sequences and harmonic complex tone stimuli are too different to use the responses to one stimulus to try to predict responses to the other. The nonlinearity of CN cells makes it difficult to use insight obtained from one stimulus about cellular mechanisms to interpret responses to the other stimulus.

The CD model makes a number of simplifying assumptions compared to actual CN neurons. One such oversimplification arises from assuming all inputs are from the AN and excitatory, when in fact the CN circuitry is known to be much more complicated. CN neurons receive inhibitory inputs (Ferragamo et al. 1998, Babalian et al. 2002), efferent inputs from descending pathways (Fujino and Oertel 2001), as well as inputs from interneurons in the CN (Bahmer and Langner 2009, Arnott et al. 2004). Moreover, the inputs based on assuming scaling invariance are too simple in that all inputs to the CD model have identical response properties (including spontaneous rates) and do not include the substantial variability that would be expected if inputs came from different AN fibers. It is possible that the wider dynamic range observed in some CN units may be a result of receiving inputs from AN fibers with differing spontaneous rates (Lai et al. 1994). A Hodgkin-Huxley (1952) model including more realistic indirect inputs, membrane channels, and spiking mechanism (Rothman and Manis 2003, Rothman et al. 1993) would likely provide more insight into the neural mechanisms that give rise to the robust rate cues observed in these units. Furthermore, cross-frequency CD is not the only

mechanism that can extract spatio-temporal pitch cues. It is possible that some CN units are making use of the spatio-temporal information across the AN via some other neural mechanisms. One possibility is lateral inhibition, in which a cell preferentially fires when its inputs discharge out of synchrony, would maintain the local maxima of rate profiles for harmonic complex tones at integer values of neural harmonic number (Cedolin and Delgutte 2010). Another possible mechanism is spatial and temporal integration over spatially distributed inputs, which has been shown to give rise to responses similar to chopper units (Oertel 1985). Further investigation of spatial and temporal summation mechanisms may help to explain the spatio-temporal sensitivity observed in some of these units in Chapter 3 or the robust rate cues at high levels shown in this chapter.

CHAPTER 5

General Conclusions and Discussion

The goal of this thesis was to assess the sensitivity of central neurons to the spatio-temporal pattern of discharges in the auditory periphery and to evaluate the implications of this processing for the coding of pitch. Specifically, we tested the hypothesis that a cross-frequency coincidence detection (CD) mechanism would convert the spatio-temporal pitch cues in the auditory periphery into a rate cue that is robust with level. Overall, we found that a small fraction of neurons in the cochlear nucleus (CN) had responses similar to those of cross-frequency CD model cells. Furthermore, a subset of these CD-like CN units had robust rate cues to pitch at high levels.

In Chapter 2, we demonstrated that the relative timing of discharges across auditory nerve (AN) fibers innervating different cochlear locations can be manipulated without greatly changing their discharge rates. We used complex transient stimuli called Huffman sequences, which have all-pass magnitude spectra with a single phase transition (Carney 1990). To facilitate experiments, we applied the principle of cochlear scaling invariance (Zweig 1976) to infer the spatio-temporal response pattern of AN fibers to a fixed Huffman stimulus from the responses of a fixed AN fiber to multiple Huffman stimuli with varying phase transition frequencies. Scaling invariance was first verified in model AN fibers (Zilany and Bruce 2006) and then applied to responses of physiological AN fibers.

We showed that changing the slope of the phase transition resulted in systematic changes in the spatio-temporal pattern of activity across the AN in both model and real fibers. We found that at all levels, the early responses of AN fibers were more synchronous than the later portions. In addition, at low levels, when the transition frequency FT was near the fiber's characteristic frequency (CF), the early AN response to the Huffman stimulus with sharper phase transition was smaller than that for the broad transition. When FT was further from CF, the early AN response to the sharp transition stimulus increased. In contrast, the early AN response to the broad transition stimulus was more constant across CF/FT. This difference in early response to the two transition widths was smaller at high levels. Furthermore, we found that for the later response peaks in the AN, the Huffman stimulus with a sharper phase transition excited more coincidentally across CF than the stimulus with the broad phase transition. These results were true for different stimulus levels and AN fibers tuned to different CFs, and was observed in both real fibers and model predictions. As expected, across the population of fibers studied, AN fibers showed little overall preference in their rate responses for either phase transition bandwidth. Overall, we verified that these stimuli can be used to systematically manipulate the spatio-temporal pattern in the AN without greatly changing their discharge rates.

The AN data from Chapter 2 serve as a baseline for comparison to responses of cross-frequency CD model cells and CN neurons measured in Chapter 3. We first examined responses of model CD cells to Huffman stimuli for extraction of the spatio-temporal cues in the AN, manifested by a conversion of the degree of cross-frequency coincidence in the AN pattern into a rate response. Indeed, model CD cells did extract

AN spatio-temporal cues by preferentially firing when their inputs tuned to different CFs discharged in synchrony. Cross-frequency CD model cells had larger earlier responses compared to their late responses, and also had larger late responses to the stimulus with the sharp transition. As the input range (or the range of CFs of AN fibers provided as input to the model CD cell) increased, the CD cell increasingly had larger early responses compared to the late responses, and within the early response, increasingly preferred the Huffman stimulus with sharp phase transition.

To quantitatively determine the similarity between responses of CN units and those of model CD cells, we used a maximum likelihood method to find the CD parameters most likely to give rise to the measured CN responses. We found that most CN units (particularly primary-like, or Pri) had responses that resembled AN fibers more than CD cells. However, a subset of CN units (primary-like-with-notch (Pri-N) and some phase lockers) had responses consistent with cross-frequency CD cells.

In Chapter 4, we examined the implications of a cross-frequency CD mechanism on the neural representation of pitch. One way in which AN fibers represent the pitch of harmonic complex tones is in their rate-place profiles, in which fibers tuned to frequencies near stimulus components fire more than fibers tuned to frequencies in between components. At high levels, the firing rates of AN fibers saturate, thus degrading the rate cue to pitch. This is in contrast to psychophysical observations that show that pitch perception remains robust even at high levels, suggesting there must be other neural mechanisms for pitch representation beyond the rate representation in the AN.

Cedolin and Delgutte (2010) examined a spatio-temporal representation of pitch, represented by rapid relative AN phase changes when the CF coincides with a resolved harmonic, and more constant phase in between harmonics. They found that these spatio-temporal cues were more robust than rate cues at high levels in the AN. In theory, in the absence of rate cues at high levels, a model cross-frequency CD cell could extract these AN spatio-temporal cues by producing local firing rate maxima when the CF falls halfway between two harmonics. The intervals between the local maxima serve as the rate cue in these CD cells to the pitch of the stimulus.

Responses of model cross-frequency CD cells operating on AN patterns of activity to harmonic complex tones displayed enhanced rate cues (stronger oscillations in their rate profiles) to pitch at all levels compared to the AN. As expected, at high levels, the rate profiles of these model cells were more likely than AN fibers to have local maxima when the CF was halfway between two harmonics.

Not surprisingly, CN Pri units had responses to pitch stimuli similar to those of AN fibers, with degraded rate cues at high levels. In contrast, some of the other CN unit types (in particular Pri-N and chopper units) displayed robust rate cues even at high levels. However, it is not clear whether this robustness arises from a cross-frequency CD mechanism, because there was no obvious correlation between rate cue robustness and similarity to CD cells as assessed by responses to Huffman stimuli in Chapter 3. Two out of the four Pri-N units found to have robust rate cues were determined to resemble cross-frequency CD cells, but the other two more resembled AN fibers. Furthermore, due to the strong regularity in the response patterns of chopper units, it is unlikely that these units perform a CD operation with the short time constants assumed in our CD model.

Thus the responses of chopper units to Huffman stimuli were *a priori* not assessed for similarity to CD cells.

It is possible that the Huffman sequence is not the ideal stimulus for making judgments about the cellular mechanisms of a CN cell, if the goal is to interpret responses to harmonic complex tones. Moreover, the CD model used in this work is likely an oversimplification of the actual CN neural circuitry. Our CD model assumes all inputs arise directly from AN fibers and are excitatory, while it is known that many CN cells receive both excitatory and inhibitory inputs. Furthermore, even though scaling invariance is a very powerful tool for facilitating our experiments, the inputs to the CD model (that were necessarily based on the scaling invariance assumption) did not include the variability that would arise if inputs came from different AN fibers. A more realistic model would include indirect and inhibitory inputs and more plausible membrane channel properties.

Cat vocalizations typically have fundamental frequencies between 400 and 1200 Hz (Nicastro and Owren 2003, Shipley et al. 1991, Brown et al. 1978). Within this range, rate representations of pitch in AN fibers, CD model cells, and CN units were strong. However, rate representations at lower fundamental frequencies relevant for the human voice (<300 Hz) were poorer, which likely reflects the poorer cochlear frequency selectivity at these lower frequencies. There is some evidence that the human cochlear filters may be roughly three times sharper than those in the cat (Shera et al. 2002), suggesting that a rate representation of pitch in humans might be robust even at these low frequencies. Cochlear scaling invariance has been shown to break down near the apex (van der Heijden and Joris 2006, Versteegh et al. 2011, Dhar et al. 2011), suggesting that

direct comparison between low-CF CN units and low-CF model CD cells is not appropriate. Furthermore, the responses of CN units at these low frequencies are particularly difficult to interpret because they likely arise from cells with a variety of anatomical characteristics.

In spite of our mixed results of whether the cross-frequency CD mechanism found in the CN gives rise to robust rate representations of pitch, our finding that at least some CN neurons effectively extend the dynamic range of the rate representation of pitch beyond that observed in AN fibers is interesting in itself. Pri-N units are associated with globular bushy cells in the CN, which provide inhibitory input to binaural neurons in the lateral superior olive (LSO). LSO neurons are sensitive to interaural level differences by comparing their excitatory ipsilateral inputs (from CN spherical bushy cells, which typically give rise to Pri patterns) with inhibitory contralateral inputs. A possible functional role of the globular bushy cells is to extend the dynamic range of these binaural neurons by providing strong inhibition at high levels when the excitatory inputs are saturated.

If the rate representation of pitch is actually used in the neural pathways that process pitch, an obvious question is how are the rate cues extracted? It has generally been assumed that a harmonic template matching mechanism is sensitive to the frequencies of individual components in a stimulus, and selects a "best-fitting" F0 from internally stored harmonic templates (Goldstein 1973, Wightman 1973, Terhardt 1974). These harmonic templates may arise from frequent exposure to harmonically rich sounds such as speech (Terhardt 1974). Even exposure to non-harmonic broadband sounds may give rise to harmonic templates via cochlear filtering and accumulating coincidences

across all CFs. This sort of processing has been shown to produce high degrees of coincidences spaced at harmonic intervals (Shamma and Klein 2000).

This thesis addressed fundamental questions regarding sensitivity of central neurons in the CN to the spatio-temporal pattern of the auditory periphery. We further examined the implications of spatio-temporal sensitivity on the coding of pitch by CN neurons. This study focused on the hypothesis that CN units could represent the spatio-temporal pitch cues in the AN in their rate responses. It is also possible that instead, CN neurons enhance or at least preserve these spatio-temporal pitch cues which may be extracted at later stages. A study examining the available temporal or spatio-temporal pitch cues in the CN may suggest where to look next for the pathway that processes pitch.

References

Adams, J. C. (1979). Ascending projections to the inferior colliculus. *J. Comp. Neurol.* 183:519-538.

Anderson M. J., Young E. D. (2004). Isoflurane/N2O anesthesia suppresses narrow-band but not wide-band inhibition in dorsal cochlear nucleus. *Hear. Res.* 188:29-41.

Arnott R. H., Wallace M. N., Shackleton T. M., Palmer A. R. (2004). Onset neurons in the anteroventral cochlear nucleus project to the dorsal cochlear nucleus. *J. Assoc. Res. Otolaryngol.* 5:153-170.

Assmann, P. F., Summerfield Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *J. Acoust. Soc. Am.* 88:680–697.

Babalian A. L., Jacomme A. V., Doucet, J. R., Ryugo D. K., Rouiller E. M. (2002). Commissural glycinergic inhibition of bushy and stellate cells in the anteroventral cochlear nucleus. *NeuroReport*, 13:555-558.

Bahmer A., Langner G. (2009). A simulation of chopper neurons in the cochlear nucleus with wideband input from onset neurons. *Biol. Cybern.* 100:21-33.

Batra R., Kuwada S., Fitzpatrick D. C. (1997). Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *J Neurophysiol.* 78:1237-1247.

Bendor D., Wang X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*. 436:1161-1165.

Bendor D., Wang X. (2010). Neural coding of periodicity in marmoset auditory cortex. *J. Neurophysiol.* 103:1809-1822.

Bernstein J. G. (2006). Pitch perception and harmonic resolvability in normal-hearing and hearing-impaired listeners. Ph.D. Thesis, MIT, Cambridge, MA.

Bernstein J. G., Oxenham A. J. (2003). Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number? *J Acoust Soc Am.* 113:3323-3334.

Bernstein J. G. W., Oxenham A. J. (2005). An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination. *J. Acoust. Soc. Am.* 117: 3816-3831.

Blackburn C. C., Sachs M. B. (1989) Classification of unit types in the anteroventral cochlear nucleus: post-stimulus time histograms and regularity analysis. *J. Neurophysiol.* 62:1303-1329.

Blackburn C. C., & Sachs, M. B. (1990). The representations of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. *J. Neurophysiol.* 6:1191-1212.

Bourk T. R. (1976). Electrical Responses of Neural Units in the Anteroventral Cochlear Nucleus of the Cat. Ph.D. Thesis. MIT, Cambridge, MA.

Brand A., Behrend O., Marquardt T., McAlpine D., Grothe B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature.* 417:543–547.

Brawer J. R., Morest D. K. (1975). Relations between auditory nerve endings and cell types in the cat's anteroventral cochlear nucleus seen with the Golgi method and Nomarski optics. *J. Comp. Neurol.* 160:491-506.

Brown K. A., Buchwald J. S., Johnson J. R., Mikolich D. J. (1978). Vocalization in the cat and kitten. *Dev. Psychobiol.* 11:559-570.

Bruce I. C., Sachs M. B., Young E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *J. Acoust. Soc. Am.* 113:369-388.

Cant N. B. (1981). The fine structure of two types of stellate cells in the anterior division of the anteroventral cochlear nucleus of the cat. *Neuroscience.* 6:2643-2655.

Cant N. B., Morest D. K. (1979). Organization of the neurons in the anterior division of the anteroventral cochlear nucleus of the cat. Light-microscopic observations. *J. Neurosci.* 4:1909-1923.

Cao X., Oertel D. (2010). Auditory nerve fibers excite targets through synapses that vary in convergence, strength and short-term plasticity. *J. Neurophysiol.* 104:2308-2320.

Carlyon R. P., Shamma S. (2003). An account of monaural phase sensitivity. *J. Acoust. Soc. Am.* 114:333-348.

Carlyon R. P. (1998). Comments on "A unitary model of pitch perception." [*J. Acoust. Soc. Am.* 102:1811-1820 (1997)]. *J. Acoust. Soc. Am.* 104:1118-1121.

Carney L. H. (1990). Sensitivities of cells in anteroventral cochlear nucleus of cat to spatiotemporal discharge patterns across primary afferents. *J. Neurophysiol.* 64:437-456.

Carney L. H. (1992). Modelling the sensitivity of cells in the anteroventral cochlear nucleus to spatiotemporal discharge patterns. *Phil. Trans. R. Soc. Lond. B.* 336:405-406.

Carney L. H. (1993). A model for the responses of low-frequency auditory-nerve fibers in cat. *J. Acoust. Soc. Am.* 93:401-417.

Carney L. H. (1994). Spatio-temporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hear. Res.* 76:31-44.

Carney L. H., Yin T. C. T. (1988). Temporal coding of resonances by low-frequency auditory nerve fibers: single-fiber responses and a population model. *J. Neurophysiol.* 60:1653-1677.

Cedolin L., Delgutte B. (2005). Pitch of complex tones: rate-place and interspike-interval representations in the auditory nerve. *J. Neurophysiol.* 94:347-362.

Cedolin L., Delgutte B. (2010). Spatio-temporal representation of the pitch of harmonic complex tones in the auditory nerve. *J. Neurosci.* 30:12712-12724.

Colburn H. S. (1996) Computational models of binaural processing. In: *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Spring-Verlag, New York), pp. 332-400.

Cutler A., Chen H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Perc. & Psychophys.* 59:165-179.

Dhar S., Rogers A., Abdala C. (2011). Breaking away: violation of distortion emission phase-frequency invariance at low frequencies. *J. Acoust. Soc. Am.* 129:3115-3122.

Doucet J. R., Ryugo D. K. (1997). Projections from the ventral cochlear nucleus tot eh dorsal cochlear nucleus in rats. *J. Comp. Neurol.* 385:245-264.

Evans E. F. (1981). The dynamic range problem: Place and time coding at the level of cochlear nerve and nucleus. In: *Neuronal Mechanisms of Hearing*, edited by J. Syka and L. Aitkin (Plenum, New York), pp. 69-85.

Evans E. F., Palmer A. R. (1980). Relationship between the dynamic range of cochlear nerve fibers and their spontaneous activity. *Exp. Brain Res.* 40: 115-118.

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

Florentine M., Buus S., and Mason S. R. (1987). Level discrimination as a function of level for tones from 0.25 to 16 kHz. *J. Acoust. Soc. Am.* 81: 1528-1541.

Fujino K., Oertel D. (2001). Cholinergic modulation of stellate cells in the mammalian ventral cochlear nucleus. *J. Neurosci.* 21:7372-7383.

Gai Y., Carney L. H. (2008). Influence of inhibitory inputs on rate and timing of responses in the anteroventral cochlear nucleus. *J. Neurophysiol.* 99:1077-1095.

Godfrey D.A., Kiang N.Y.S., Norris B. E. (1975). Single unit activity in the dorsal cochlear nucleus of the cat. *J Comp Neurol.* 162:269-284.

Goldberg J. M., Brown P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol.* 32: 613–636.

Goldstein J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *J. Acoust. Soc. Am.* (54):1496-1516.

Greenwood D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *J. Acoust. Soc. Am.* 33:1344-1356.

Greenwood D. D. (1990). A cochlear frequency-position function for several species: 29 years later. *J. Acoust. Soc. Am.* 87:2592-2605.

Heffner H., Whitfield I. C. (1976). Perception of the missing fundamental by cats. *J. Acoust. Soc. Am.* 59:915-919.

Heinz M. G., Colburn H. S., Carney L. H. (2001). Rate and timing cues associated with the cochlear amplifier: Level discrimination based on monaural cross-frequency coincidence detection. *J. Acoust. Soc. Am.* 110:2065-2084.

Hewitt M. J., Meddis R., Shacketon T. M. (1992). A compute model of a cochlear-nucleus stellate cell: responses to amplitude-modulated and pure-tone stimuli. *J. Acoust. Soc. Am.* 91:2096-2109.

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

Huffman D. A. (1962). The generation of impulse-equivalent pulse trains. *IRE Trans IT.* 8:S10-S16.

Jeffress L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41:35-39.

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

Johnson D. H., Swami A. (1983). The transmission of signals by auditory-nerve fiber discharge patterns. *J. Acoust. Soc. Am.* 74:493-501.

Joris P. X., Carney L. H., Smith P. H., Yin T. C. T. (1994a). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I: Responses to tones at the characteristic frequency. *J. Neurophysiol.* 71:1022-1036.

Joris P. X., Carney L. H., Smith P. H., Yin T. C. T. (1994b). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II: Responses in the tuning curve tail. *J. Neurophysiol.* 71:1037-1051.

Joris P. X., Smith P. H. (2008). The volley theory and the spherical cell puzzle. *Neuroscience*. 154:65-76.

Kalluri S., Delgutte B. (2003a). Mathematical models of cochlear nucleus onset neurons: I. Point neuron with many weak synaptic inputs. *J. Comp. Neuro.* 14:71-90.

Kalluri S., Delgutte B. (2003b). Mathematical models of cochlear nucleus onset neurons: II. Model with dynamic spike-blocking state. *J. Comp. Neuro.* 14:91-110.

Kane E. C. (1973). Octopus cells in the cochlear nucleus of the cat: heterotypic synapses upon homeotypic neurons. *Int. J. Neurosci*. 5:251-279.

Kiang N.Y.S., Moxon E.C., Levine R.A. (1970). Auditory-nerve activity in cats with normal and abnormal cochleas. In: *Sensorineural hearing loss,* edited by G. E. W. Wolstenholme, J. Knight, pp. 241-267.

Kiang N. Y. S. (1990). Curious oddments of auditory-nerve studies. *Hear. Res*. 1-16.

Kiang N. Y. S., Watanabe T., Thomas E. C., Clark L. F. (1965a). Discharge Patterns of Single Fibers in the Cat's Auditory Nerve. *The MIT Press*.

Kiang N. Y. S., Pfeiffer R. R., Warr W. B., Backus A. S. N. (1965b). Stimulus coding in the cochlear nucleus. *Annals of Otology, Rhinology and Laryngology*. 74:463-485.

Krips R., Furst M. (2009). Stochastic Properties of Coincidence-Detector Neural Cells. *Neural Comp*. 21:2524-2553.

Lai Y. C., Winslow R. L., Sachs M. B. (1994). A model of selective processing of auditory-nerve inputs by stellate cells of the antero-ventral cochlear nucleus. *J. Comp. Neurosci.* 1:167-194.

Larsen E., Cedolin L., Delgutte B. (2008). Pitch representations in the auditory nerve: two concurrent complex tones. *J. Neurophysiol.* 100:1301-1319.

Laudanski J., Coombes S., Palmer A. R., Sumner C. J. (2010). Mode-locked spike trains in responses of ventral cochlear nucleus chopper and onset neurons to periodic stimuli. *J. Neurophysiol.* 103:1226-1237.

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

Liberman M. C. (1991). Central projections of auditory-nerve fibers of differing spontaneous rate. I. Anteroventral cochlear nucleus. *J. Comp. Neurol.* 313: 240-258.

Liberman M. C. (1993). Central projections of auditory-nerve fibers of differing spontaneous rate. II. Posterventral and dorsal cochlear nuclei. *J. Comp. Neurol.* 327: 17-36.

Lyon R., Shamma S. (1996). Auditory representation of timbre and pitch. In: *Auditory Computations,* edited by H. Hawkins, E. T. McMullen, A. Popper, and R. Fay (Berlin), pp. 221-270.

May B.J., Glenn S.L.P., Sachs M.B. (1998). Vowel representations in the ventral cochlear nucleus of the cat: effects of level, background noise, and behavioral state. *J. Neurophysiol.* 79:1755-1767.

Molnar C. E., Pfeiffer R. R. (1968). Interpretation of spontaneous spike discharge patterns of neurons in the cochlear nucleus. *Proc. IEEE.* 56:993-1004.

Moore B.C. J. (1973a). Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am.* 54:610-619.

Moore B.C. J. (1973b). Some experiments relating to the perception of complex tones. *Q. J. Exp. Psychol.* 25:451-475.

Moore B. C. J., Carlyon R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In: *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. R. Fay, A. N. Popper (Springer: New York), 24: 234–277.

Moushegian G., Rupert A. I., Gidda J. S. (1975). Functional characteristics of superior olivary neurons to binaural stimuli. *J. Neurophysiol.* 38:1037-1048.

Needham K., Paolini A. G. (2003). Fast inhibition underlies the transmission of auditory information between cochlear nuclei. *J. Neurosci.* 23:6357-6361.

Nicastro N., Owren M. J. (2003). Classification of domestic cat (Felis catus) vocalizations by naive and experienced human listeners. *J. Comp. Psychol.* 117: 44-52.

Nolan F. (2002). Intonation in Speaker Identification: An Experiment on Pitch alignment Features. *Int. J. Speech. Lang. Law.* 9:1-21.

Oertel D. (1985). Use of brain slices in the study of the auditory system: Spatial and temporal summation of synaptic inputs in cells in the anteroventral cochlear nucleus of the Mouse. *J. Acoust. Soc. Am.* 78:328-333.

Oertel D., Bal R., Gardner S. M., Smith P. H., Joris P. X. (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proc. Nat. Acad. Sci.* 97:11773-11779.

Oertel D., Wickesberg R. E. (1993). Glycinergic inhibition in the cochlear nuclei: evidence for tuberculoventral neurons being glycinergic. In: *The Mammalizan Cochlear Nuclei: Organization and Function*, edited by M. A. Merchan, J. J. Juiz, D. A. Godfrey, E. Mugnaini (Plenum: New York), pp. 225-238.

Oertel D., Young E. D. (2004). What's a cerebellar circuit doing in the auditory system? *Trends Neurosci.* 27:104-110.

Oliver D. L. (1987). Projections to the inferior colliculus from the anteroventral cochlear nucleus in the cat: possible substrates for binaural interaction. *J. Comp. Neurol.* 264:24-46.

Osen K. K. (1969). The intrinsic organization of the cochlear nuclei in the cat. *Acta. Otolaryngol.* 67:352-359.

Osen K. K. (1970). Course and terminations of the primary afferents in the cochlear nuclei of the cat. An experimental anatomical study. *Arch. Ital. Biol.* 108:21-51.

Osen K. K. (1972). Projection of the cochlear nuclei on the inferior colliculus in the cat. *J. Comp. Neurol.* 144:355-372.

Ostapoff E. M., Benson C. G., Saint Marie R. L. (1998). GABA- and glycine-immunoreactive projections from the superior olivary complex to the cochlear nucleus in guinea pig. *J. Comp. Neurol.* 381:500-512.

Oxenham A. J., Micheyl C., Keebler M. V., Loper A., Santurette S. (2011). Pitch perception beyond the traditional existence region of pitch. *Proc. Nat. Acad. Sci.* 108:7629-7634.

Palmer A. R., Russell I. J. (1986). Phase-locking in the cochlear nerve of the guinea pig and its relation to the receptor potential of inner hair cells. *Hearing Res.* 24:1-15.

Palmer A. R., Wallace M. N., Arnott R. H., Shackleton T. M. (2003). Morphology of physiological characterized ventral cochlear nucleus stellate cells. *Exp. Brain. Res.* 153:418-426.

Palmer A. R., Winter I. M. (1992). Cochlear nerve and cochlear nucleus responses to the fundamental frequency of voiced speech sounds and harmonic complex tones. In: *Auditory Physiology and Perception*, edited by K. Horner (Pergamon: Oxford), pp. 231-240.

Palmer A. R., Winter I. M. (1993). Coding of the fundamental frequency of voiced speech sounds and harmonic complex tones in the ventral cochlear nucleus. In: *Mammalian Cochlear Nuclei: Organization and Function*, edited by E. Mugnaini (Plenum: New York), pp. 373-384.

Patterson J. H., Green D. M. (1970). Discrimination of transient signals having identical energy spectra. *J. Acoust. Soc. Am.* 48:894-905.

Patterson J. H., Ronken D. A., Green D. M. (1969). Phase perception of transient signals having identical power spectra. *J. Acoust. Soc. Am.* 46:121-121.

Pfeiffer R. R. (1966). Classification of response patterns of spike discharges for units in the cochlear nucleus: tone-burst stimulation. *Exp. Brain. Res.* 1:220-235.

Pfeiffer R. R., Kim D. O. (1975). Cochlear nerve fiber responses: Distribution along the cochlear partition. *J. Acoust. Soc. Am.* 58:867-965.

Pfeiffer R. R., Molnar (1970). Cochlear nerve fiber discharge patterns: Relationship to the cochlear microphonic. *Science.* 167:1614-1616.

Pickles J. O. (1988). *An Introduction to the Physiology of Hearing.* Academic Press: London.

Plomp R. (1964). The ear as a frequency analyzer. *J. Acoust. Soc. Am.* 36:1628-1636.

Retzius G. (1884). Das Gehororgan der Wirbeltiere: II. Das Gehororgan der Amnioten. (Samson und Wallin: Stockholm), pp. 345.

Rhode W. S. (1995). Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. *J. Acoust. Soc. Am.* 97:2413-2429.

Rhode W. S., Oertel D., Smith P. H. (1983a) Physiological response properties of cells labeled intracellularly with horseradish peroxidase in cat dorsal cochlear nucleus. *J. Comp. Neurol.* 213:426-447.

Rhode W. S., Oertel D., Smith P. H. (1983b) Physiological response properties of cells labeled intracellularly with horseradish peroxidase in cat ventral cochlear nucleus. *J. Comp. Neurol.* 213:448-463.

Rothman J. S., Manis P. B. (2003). The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. *J. Neurophysiol.* 89:3097-3113.

Rothman J. S., Young E. D., Manis P. B. (1993). Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model. *J. Neurophysiol.* 70:2562-2583.

Rose J. E., Gross N. B., Geisler C. D., Hind J. E. (1966). Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. *J. Neurophysiol.* 29:288-314.

Ruggero M. A., Robles L., Rich N. C. (1992). Two-tone suppression in the basilar membrane of the cochlear: mechanical basis of auditory-nerve rate suppression. *J. Neurophysiol.* 68:1087-1099.

Ryugo D. K., Sento S. (1991). Synaptic connections of the auditory nerve in cats: Relationship between endbulbs of Held and spherical bushy cells. *J. Comp. Neurol.* 305:35-48.

Sayles M., Winter I. M. (2007). The temporal representation of the delay of dynamic iterated rippled noise with positive and negative gain by single units in the ventral cochlear nucleus. *Brain Res.* 1171:52-66.

Sayles M., Winter I. M. (2008). Ambiguous pitch and the temporal representation of inharmonic iterated rippled noise in the ventral cochlear nucleus. *J. Neurosci.* 28:11925-11938.

Scheffers, M. T. (1979). The role of pitch in perceptual separation of simultaneous vowels. *Inst. Perc. Res.* 14:51-54.

Schouten, J. F. (1940). The residue and the mechanism of hearing. *Proc. Kon. Akad. Wetenschap.* 43:991-999.

Schouten J. F., Ritsma R. J., Cardozo B. L. (1962). Pitch of the residue. *J. Acoust. Soc. Am.* 34:1418-1424.

Sento S., Ryugo D. K. (1989). Endbulbs of Held and spherical bushy cells in cats: morphological correlates with physiological properties. *J. Comp. Neurol.* 280:553-562.

Shamma S. (1985a). Speech processing in the auditory system: I. representation of speech sounds in the responses of the auditory nerve. *J. Acoust. Soc. Am.* 78:1612-1621.

Shamma S. (1985b). Speech processing in the auditory system: II: lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *J. Acoust. Soc. Am.* 78:1622:1632.

Shamma S., Klein D. (2000). The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. *J. Acoust. Soc. Am.* 107:2631-2644.

Shera C. A., Guinan J. J. Jr., Oxenham A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc. Nat. Acad. Sci.* 99:3318-3323.

Shera C. A., Guinan J. J. Jr. (2003). Stimulus-frequency-emission delay: a test of coherent reflection filtering and a window on cochlear tuning. *J. Acoust. Soc. Am.* 113: 2762–2772.

Shipley C., Carterette E. C., and Buchwald J. S. (1991). The effect of articulation on the acoustical structure of feline vocalizations. *J. Acoust. Soc. Am.* 89: 902-909.

Shofner W. P. (1991). Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla. *J. Acoust. Soc. Am.* 90:2450-2466.

Shofner W. P. (1999). Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: Analysis of neural autocorrelograms. *J. Neurophysiol.* 81:2662-2674.

Smith P. H., Joris P. X., Carney L. H., Yin T. C. T. (1991). Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *J. Comp. Neurol.* 304:387-407.

Smith P. H., Joris P. X., Yin T. C. T. (1993). Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J. Comp. Neurol.* 331:245-260.

Smith P. H., Rhode W. S. (1987). Characterization of HRP-labelled globular bushy cells in the cat anteroventral cochlear nucleus. *J. Comp. Neurol.* 266:360-375.

Smith P. H., Rhode W. S. (1989). Structural and functional properties distinguish two types of multipolar cells in the ventral cochlear nucleus. *J. Comp. Neurol.* 282:595-616.

Spirou G. A., Rager J., Manis P. B. (2005). Convergence of auditory-nerve fiber projections onto globular bushy cells. *Neuroscience*. 136:843-863.

Spitzer M. W., Semple M. N. (1995). Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *J. Neurosphyiol.* 73:1668-1690.

Summerfield, A. Q., Assmann, P. F. (1991). Perception of concurrent vowels: effects of pitch-pulse asynchrony and harmonic misalignment. *J. Acoust. Soc. Am.* 89:1364–1377.

Tan Q., Carney L. H. (2003). A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. *J. Acoust. Soc. Am.* 114:2007-2020.

Terhardt E. (1974). Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55:1061-1068.

Tolbert L. P., Morest D. K. (1982a). The neuronal architecture of the anteroventral cochlear nucleus of the cat in the region of the cochlear nerve root: Electron microscopy. *Neuroscience*. 7:3053-3068.

Tolbert L. P., Morest D. K. (1982b). The neuronal architecture of the anteroventral cochlear nucleus of the cat in the region of the cochlear nerve root: Golgi and nissl methods. *Neuroscience*. 7:3013-3030.

van der Heijden M., Joris P. X. (2006). Panoramic measurements of the apex of the cochlea. *J Neurosci.* 26:11462-11473.

van Gisbergen J. A. M, Grashuis J. L., Johannesma P. I. M, Vendrick A. J. H. (1975). Statistical analysis and interpretation of the initial response of cochlear nucleus neurons to tone bursts. *Exp. Brain Res*. 23:407-423.

Verhey J. L., Winter I. M. (2006). The temporal representation of the delay of iterated rippled noise with positive and negative gain by chopper units in the cochlear nucleus. *Hear. Res*. 216-217:43-51.

Versteegh C. P. C., Meenderink S. W. F., van der Heijden M. (2011). Response characteristics in the apex of the gerbil cochlea studied through auditory nerve recordings. *J. Assoc. Res. Otolaryngol.* 12:301-316.

Viemeister N. F. (1988a) Pschophysical aspects of auditory intensity coding. In: *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley: New York), pp. 213-241.

Viemeister N. F. (1988b) Intensity coding and the dynamic range problem. *Hear Res.* 34:267-274.

von Bekesy G. (1960). *Experiments in hearing*. McGraw Hill: New York.

Wenthold R. J. (1987). Evidence for a glycinergic pathway connecting the two cochlear nuclei: an immunocytochemical and retrograde transport study. *Brain Res.* 415:183-187.

Westerman L. A., Smith R. L. (1988). A diffusion model for the transient response of the cochlear inner hair cell synapse. *J. Acoust. Soc. Am.* 83:2266-2276.

Wickesberg R. E., Oertel D. (1990). Delayed, frequency-specific inhibition in the cochlear nuclei of mice: a mechanism for monaural echo suppression. *J. Neurosci.* 10:1762-1768.

Wiegrebe L., Meddis R. (2004). The representation of periodic sounds in simulated sustained chopper units in the ventral cochlear nucleus. *J. Acoust. Soc. Am.* 115:1207-1217.

Wiegrebe L., Winter I. M. (2001). Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus. *J. Neurophysiol.* 85:1206-1219.

Wightman F. L. (1973). The pattern-transformation model of pitch. *J. Acoust. Soc. Am.* (54):407-416.

Winter I. M., Palmer A. R., Wiegrebe L., Patterson R. D. (2003). Temporal coding of the pitch of complex sounds by presumed multipolar cells in the ventral cochlear nucleus. *Speech Comm.* 41:135-149.

Winter I. M., Wiegrebe L., Patterson R. D. (2001). The temporal representation fo the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *J. Physiol.* 537:553-566.

Wouterlood F. G., Mugnani E., Osen K. K., Dahl A. L. (1984) Stellate neurons in rat dorsal cochlear nucleus with combined Golgi impregnation and electron microscopy: synaptic connections and mutual coupling by gap junctions. *J. Neurocytol.* 4:639-664.

Wu S. H., Oertel D. (1984). Intracellular injection with horseradish peroxidase of physiologically characterized stellate and bushy cells in slices of mouse anteroventral cochlear nucleus. *J. Neurosci.* 4:1577-1588.

Yin T.C.T., Chan J. C. K. (1990). Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol.* 64:465-488.

Young E. D. (1984). Response characteristics of neurons of the cochlear nuclei. In: *Hearing Science*, edited by C. I. Berlin (College-Hill: San Diego), pp. 423-460.

Young E. D., Robert J. M., Shofner W. P. (1988). Regularity and latency of units in the ventral cochlear nucleus: Implications for unit classification and generation of response properties. *J. Neurophysiol.* 60:1-29.

Zhang X., Heinz M. G., Bruce I. C., Carney L. H. (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.* 109:648-670.

Zhang S., Oertel D. (1993). Cartwheel and superficial stellate cells of the dorsal cochlear nucleus of mice: intracellular recordings in slices. *J. Neurophysiol.* 69:1384-1397.

Zilany M. S. A., Bruce I. C. (2006). Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. *J. Acoust. Soc. Am.* 120:1446-1466.

Zilany M. S. A., Bruce I. C. (2007). Representation of the vowel /ɛ/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats. *J. Acoust. Soc. Am.* 122:402-417.

Zilany M. S. A., Bruce I. C. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics. *J. Acoust. Soc. Am.* 126:2390-2412.

Zitova B., Flusser J. (2003). Image registration methods: a survey. *Image and Vis. Comp.* 21:977-1000.

Zweig G. (1976). Basilar membrane motion. *Cold Spring Harb. Symp. Quant. Biol.* 40:619-633.