Spatio-Temporal Representation of the Pitch of Complex Tones in the Auditory Nerve and Cochlear Nucleus

by

Grace I. Wang

B.S. Electrical Engineering
University of Illinois, Urbana-Champaign 2005

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

Master of Science in Electrical Engineering and Computer Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2007

Signature of author: ....

....
ineering and Computer Science
August 30, 2007

Certified by: ....

....
Bertrand Delgutte
Research Scientist
Thesis Supervisor

Accepted by: ....(

....
Arthur C. Smith
Chairman, Department Committee for Graduate Students

1

Spatio-temporal representation of the pitch of complex tones in the auditory nerve and cochlear nucleus

by

Grace I. Wang

ABSTRACT

Traditional models for pitch processing have relied either on a purely spatial representation based on the frequency selectivity and frequency-to-place mapping in the cochlea, or on a purely temporal representation dependent on neural phase locking to the stimulus period. However, rate-place cues to the pitch of harmonic complex tones in the auditory nerve (AN) saturate at high levels, and temporal cues fail to predict the strong pitch of resolved harmonics. Thus, neither pitch representation accounts for all key psychophysical observations.

Recently, it has been shown that the AN also contains spatio-temporal cues to the individual harmonics of a complex tone that might be used in pitch extraction and are more consistent with psychophysical data (Cedolin 2006) This thesis aims to evaluate whether these cues are extracted in cochlear nucleus (CN) neurons, which receive inputs from AN fibers.

We used transient complex stimuli ("Huffman sequences") designed to manipulate the relative timing between adjacent AN fibers to evaluate sensitivity of CN units to the spatio-temporal pattern of their inputs. A small minority of units were sensitive to the stimulus manipulations, with only a few sensitive in the direction consistent with cross-frequency coincidence detection. We also measured the strength of the rate representation for the pitch of harmonic complex tones in CN units and found a few units that maintained salient pitch cues at high stimulus levels. However there was no obvious correlation between spatio-temporal sensitivity and robust pitch cues. Instead of a conversion of spatio-temporal cues in the AN into rate cues, our results indicate that temporal sharpening occurs at the level of the CN.

Thesis Supervisor: Bertrand Delgutte
Title: Senior Research Scientist

Motivation

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. 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 role in speech, such as for speaker identification (Nolan 2001), portrayal of emotions, and tonal languages (Cutler et al 1997). Differences in pitch can serve as cues for sound source segregation (Scheffers 1982, Summerfield and Assmann 1990). Pitch perception degrades in both hearing impaired listeners (Bernstein 2005) 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. In this thesis, we compared neural responses to harmonic complex tones in the auditory nerve (AN) and cochlear nucleus (CN) to determine whether spatio-temporal pitch cues available in the AN are extracted by CN neurons. We also used complex transient stimuli, "Huffman sequences," to assess the sensitivity of neurons to the spatio-temporal patterns of activity in the AN.

Background

Peripheral Auditory Processing

As sound travels through the inner ear, the cochlea performs a frequency analysis, breaking 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 may be thought of as a tonotopically mapped array, or a spatial representation of frequencies, of overlapping band-pass filters (also referred to as "auditory filters") (Fletcher 1940). 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 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 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.

In response to a low-frequency pure tone (<5 kHz), AN fibers tend to fire action potentials at a specific phase of the stimulus, a phenomenon called phase locking. Even though it is unlikely for a fiber to respond to each cycle, a population of fibers responding to the same low frequency sound may be able to give a complete temporal representation of the stimulus waveform (Javel et al 1988). Due to limitations presented by the mechanics of the inner hair cells and their synapses, phase locking breaks down for high frequency tones above 4 to 5 kHz (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 also be found in the fiber firing rates. Fibers with high spontaneous activity have low thresholds and small dynamic range, 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 each fiber 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.

Cochlear Nucleus

AN fibers convey the electrical signals carrying coded information about the acoustic stimulus to neurons in the CN, which is the first site of computation along the ascending central auditory pathway. The three subdivisions of the CN (anteroventral CN (AVCN), posteroventral CN (PVCN), and dorsal CN (DCN)) contain distinct cell types, which have differing morphologies, input patterns, 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.

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 (STBCFs) (Kiang et al 1965b, Pfeiffer 1966a). The common unit types are primary-like (Pri), primary-like-with-notch (Pri-N), onset (On), sustained and transient chopper (Chop-S and Chop-T respectively), pauser, and buildup.

Spherical bushy cells are located in the anterior part of the AVCN (Osen 1969, Cant and Morest 1979a). Their AN inputs end in large endbulb synapses, which result in nearly one-to-one synaptic transmission with short response latencies. Because they receive only a few AN inputs, spherical bushy cells respond in a nearly one-to-one fashion to AN action potentials, behaving as relays for the temporal information contained in the AN firing patterns (Molnar and Pfeiffer 1968, Pfeiffer 1966, Rhode and Smith 1986, Young 1984, Young et al 1988). In response to STBCFs, spherical bushy cells are usually associated with a Pri response pattern, which is characterized by a high discharge rate at the stimulus onset, followed by a smooth decrease to a steady rate during the rest of the tone burst. In addition, Pri units tend to strongly phase lock to low frequency acoustic stimuli.

Globular bushy cells predominantly appear in the posterior part of the AVCN and anterior part of the PVCN (Osen 1969a, Tolbert and Morest 1982a). On average, 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). The firing pattern of these cells in response to STBCFs is associated with a Pri-N response 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, the high onset response could indicate a high probability that many of the cells' inputs are firing at the stimulus onset. The refractory period property prohibits action potentials from firing within around 0.6 ms of one another and could explain the notch of no activity following the

5

onset. If this is the case, then globular bushy cells make use of precise timing information of their inputs.

Octopus cells are found in the caudal region of the PVCN. Like bushy cells, they have short response latencies. However instead of receiving a few large terminals from a few AN inputs, octopus cells receive many small bouton endings from many AN fibers (Kane 1973). These cells are usually associated with an onset (On) pattern in response to STBCFs, characterized by a sharply timed response at the beginning of the stimulus followed by little to no activity throughout the remainder of the tone burst. On-I units (associated with octopus cells) respond only at the onset, while On-L units (associated with octopus and bushy cells) respond with an initial burst of activity followed by some late activity. It is likely that there is a continuum between On-Ls and Pri-Ns. Several ideas have been proposed to explain the onset responses of octopus cells and others. Coincidence detection is most common but has not been completely validated with physiological data.

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 1983, Smith and Rhode 1989). Stellates receive many AN inputs in the form of small bouton endings terminated at some distance from the soma and respond to STBCFs with a chopper pattern (Cant 1981, Tolbert and Morest 1982ab, Rhode et al 1983b). A striking feature of chopper units is the high discharge regularity compared to AN firing patterns, which are irregular. Choppers can be further categorized as transient choppers (Chop-T), which chop only at the beginning of the tone and become more irregular later during the tone, and sustained choppers (Chop-S), which chop throughout the stimulus duration. Another unit type that is associated with stellate cells is the onset chopper (On-C), where brief chopping occurs at the onset, followed by little or no activity. It has been proposed that the stellate cells perform a temporal integration, in which the neuron would accumulate AN input activity across its many inputs and over time. Once 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 1983). They form a layer of cells parallel to the surface of the DCN, with their long axes perpendicular to the surface. These cells are associated with pauser or build-up patterns in response to

STBCFs. A pauser is characterized by an onset response, followed by a long pause (5-10 ms) and gradual buildup of firing. Build-up units have a similar pattern but lack the initial onset response. The timing of firing in pyramidal cells appears to be sensitive to inhibitory as well as excitatory inputs (Zhang and Oertel 1993, Young 1986, Oertel and Young 2004). This study targets responses of cells in the VCN, so pauser and build-up units were not looked at in detail.

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

Pitch of Harmonic Complex Tones

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 F0," occurs for F0s up to about 1400 Hz (Moore 1973). Consecutive components of a harmonic complex tone are said to be "resolved" when their frequencies are spaced far enough apart such that each falls in the pass band of a different auditory filter. When two or more harmonics fall within the same pass band of a single filter, they are said to be "unresolved." The bandwidths of the cochlear filters increase with increasing center frequency (CF) (Kiang et al 1965, Shera et al 2002), so low order harmonics are better resolved. Psychophysical data suggest that humans are able to resolve the first 6-10 harmonics (Bernstein and Oxenham 2003a, Plomp 1964) and that resolved harmonics produce a stronger pitch than unresolved harmonics.

Pitch Models

Debates regarding the origin of pitch processing date back to at least the 19th century (Seebeck 1841, Ohm 1843). 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 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 relied on either purely spatial or purely temporal

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

"Central pattern recognition" models (Goldstein 1973, Wightman 1973, and Terhardt 1974) propose that pitch may be derived by determining the frequencies of individual components and then selecting a "best-fitting" F0 from internally stored harmonic templates. Several models have been proposed to explain the formation of these templates (Shamma and Klein 2000). However convincing physiological evidence for the existence of the templates or how they may be formed is lacking and is currently a field of much interest. An advantage of these models is that they depend on harmonic resolvability for pitch estimates, but problems arise when trying to represent the inputs to the templates. If the inputs are based on AN fiber firing rates, high stimulus levels would cause the model to predict a degraded pitch due to rate saturation, but pitch discrimination performance remains relatively robust (Cedolin 2006, Bernstein and Oxenham 2005).

A complex tone consisting entirely of unresolved harmonics can elicit a pitch percept, though weaker than the percept resulting from resolved harmonics (Bernstein and Oxenham 2003b). The place models are not applicable for unresolved harmonics, but neural phase locking to the periodicity of the waveform (resulting in interspike intervals frequently corresponding to the waveform's period and its integer multiples) provides a cue to the pitch of the stimulus. Licklider's duplex theory of pitch perception (1951) is based on a frequency analysis performed by the cochlea followed by an autocorrelational analysis performed by neurons in the auditory system. The autocorrelation function can be implemented using delays (from synapses or a long and thin axon) followed by multiplication (coincidence detector neurons). Autocorrelation functions (ACFs) are also equivalent to an all-order interspike interval distribution (spanning all spikes, consecutive or not). A modern implementation of Licklider's theory uses pooled autocorrelation functions (ACFs), which sum across neurons receiving inputs from the entire length of the cochlea (Meddis and Hewitt 1991ab), to derive pitch estimates. Because phase locking degrades at high frequencies, the temporal models are able to predict an upper frequency limit to the "pitch of the missing F0" phenomenon. Furthermore, they predict pitch robust with stimulus level and work with both resolved and unresolved harmonics. Despite these consistencies with psychophysical studies, weaknesses of autocorrelation models include failing to predict stronger pitch for resolved harmonics (Carlyon 1998, Bernstein and Oxenham 2005),

underestimating the pitch strength of pure tones, and possibly requiring long neural delay lines or intrinsic oscillators, for which there is little physiological evidence.

Spatio-temporal Model

Convincing biological data is needed to either confirm or refute the spatial or temporal models. Recently Cedolin and Delgutte (2007) gave physiological evidence for a spatio-temporal representation of the pitch of harmonic complex tones with missing F0s in the cat AN. There is both physiological and theoretical evidence that 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 Kim 1975, Lyon and Shamma 1996, Shamma 1985a). For AN fibers tuned to CFs within the range of phase locking, this phase change is reflected in the relative timing of AN spikes at different cochlear locations and generates a 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 1985a). These central neurons could then provide inputs to a harmonic template mechanism.

Cedolin and Delgutte (2007) demonstrated that the locations of these rapid phase delays occur at places with CF corresponding to low order integer harmonic frequencies of the F0 of a complex tone (the resolved harmonics), and that the phase varied slower at places corresponding to non-integer harmonic values. This was first verified with a peripheral auditory model (Zhang et al 2001) and then demonstrated physiologically in single AN fibers in the cat.

Figure 1 shows the AN model's response at varying cochlear locations (y-axis, in "neural harmonic number," or CF / F0, from 0.6 to 1.6) to a pure tone stimulus at 1500 Hz in the form of period histograms as a function of normalized time (x-axis, in cycles). The model filters were fit to human psychophysical masking data (Glasberg and Moore 1990) for the examples in figures 1 and 2. The model fiber whose CF is equal to the tone frequency (neural harmonic number = 1) has a rapid phase change relative to fibers whose CFs are located further from the tone. The rate-place profile (the black curve in the right panel) represents the overall rate of each fiber by integrating the response of the model fibers over time. The mean absolute spatial derivative (MASD) profile (the red curve in the right panel) was constructed by integrating over time the absolute difference of the responses of two consecutive fibers. The MASD effectively represents

the result of a hypothetical lateral inhibition network operating on AN inputs from different CFs that could extract spatio-temporal cues (Shamma 1985b). Both profiles were then normalized by their maximum values and plotted against cochlear location (or neural harmonic number). The maximum at neural harmonic number = 1 in the rate profile indicates that the fiber with CF = F0 responds best compared to its surrounding fibers, while the same maximum in the MASD profile indicates that the phase change occurs most rapidly at this location.
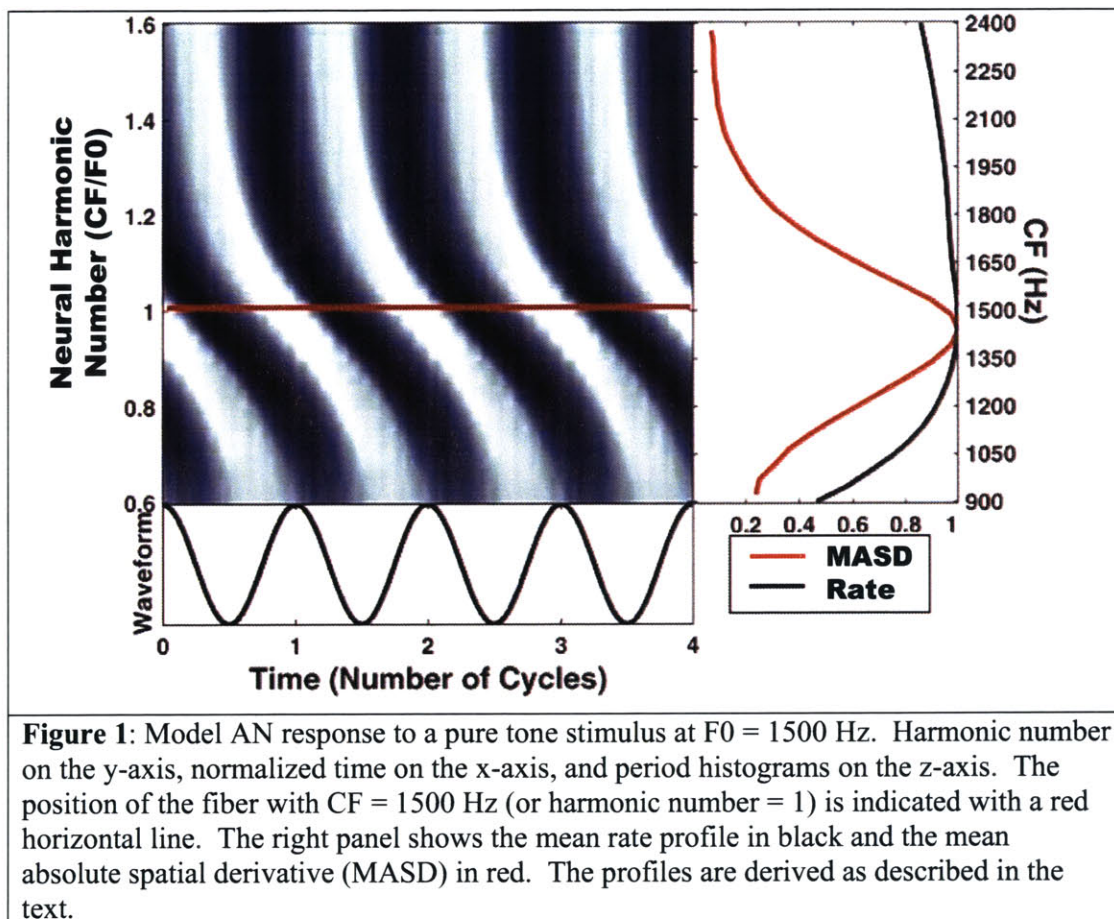

Figure 1: Model AN response to a pure tone stimulus at F0 = 1500 Hz. Harmonic number on the y-axis, normalized time on the x-axis, and period histograms on the z-axis. The position of the fiber with CF = 1500 Hz (or harmonic number = 1) is indicated with a red horizontal line. The right panel shows the mean rate profile in black and the mean absolute spatial derivative (MASD) in red. The profiles are derived as described in the text.

Figure 2 shows the peripheral model's response to a harmonic complex tone with missing F0 at 200 Hz and 50 dB SPL with the same format as in figure 1. The rate-place profile has maxima at integer neural harmonic numbers, indicating the better response of fibers with CFs near these locations (in black on the right panel), and less response in between. The strength of oscillation in the profile is effectively a measure of harmonic resolvability. The model also predicts that fibers with CFs at integer multiples of the pitch of the stimulus have rapid phase changes (slanted red lines) while fibers with CFs in between integer harmonic numbers have

slower phase changes (vertical yellow lines). As a result, the spatial derivative profile (in red on the right panel) has peaks at locations corresponding to integer harmonic numbers and valleys in between. Both profiles have peaks at integer harmonic numbers, but the rate profile's oscillation is weaker than the spatial derivative because the stimulus level is high enough to saturate the firing rates of the model fibers.
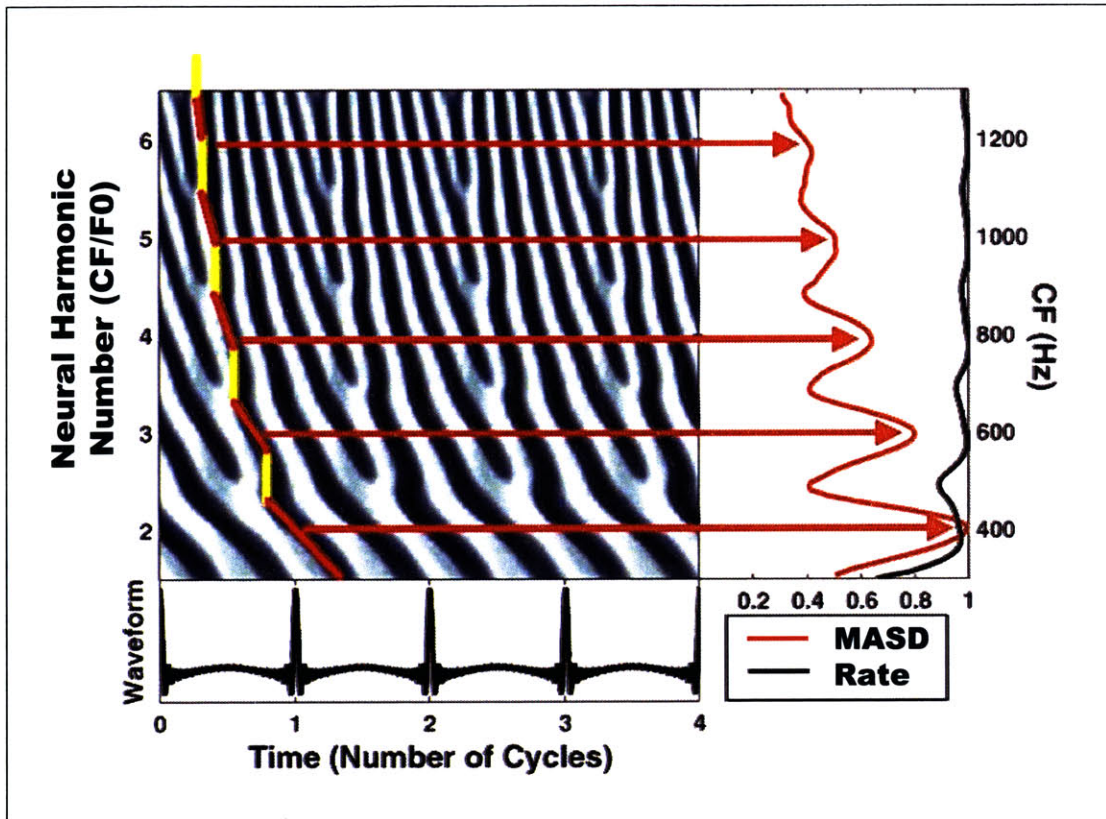

Figure 2: Taken from figure 2.1 in Cedolin (2006). Model AN response to a harmonic complex tone stimulus with missing F0 at 200 Hz and 50 dB SPL. More phase shift occurs at integer harmonic numbers (slanted red lines) while less phase shift occurs in between integer values (vertical yellow lines). The right panel shows that the rate profile (black) has poorer oscillations compared with the MASD profile (red).

The principle of local scaling invariance in cochlear mechanics (Zweig 1976) was used to infer the spatio-temporal response pattern to a given complex tone from a set of response patterns measured from a single neuron as a function of F0. The left panel of figure 3 shows the response patterns of a population of model fibers (using cat cochlear bandwidths) with CFs ranging from 750 to 2250 Hz to a harmonic complex tone stimulus with missing fundamental (fixed at 500 Hz). The right panel of the figure 3 shows the model response of a single AN fiber

(CF = 1500 Hz) to harmonic complex tone stimuli with varying F0 (ranging from 333 to 1000 Hz). The nearly identical response patterns, rate profiles, and spatial derivative profiles in the two panels support the use of scaling invariance to infer the response of an array of fibers to a complex tone from the response of a single fiber to complex tones with varying F0.
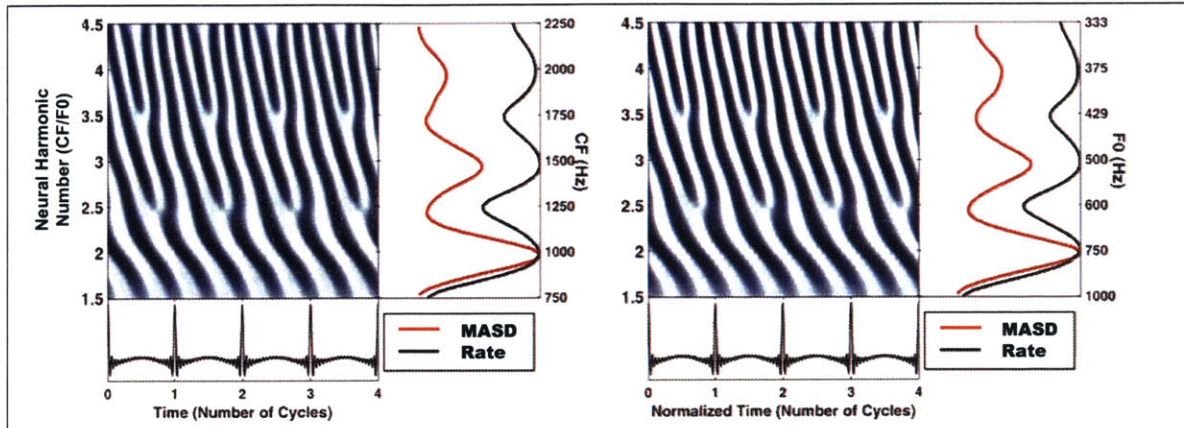

Figure 3: Taken from figure 2.2 in Cedolin (2006). Local scaling invariance. Left: response of array of AN fibers to same F0 stimulus; Right: response of single fiber to varying F0 stimuli. The similarity between the left and right panels illustrates that the spatio-temporal response pattern of an array of fibers to same F0 may be inferred from the pattern of a single fiber's response to stimuli with varying F0.

Figure 4 shows the response patterns of an AN fiber (CF = 1284 Hz) to a set of harmonic complex tones at three stimulus levels. As with the model data, period histograms are displayed as a function of neural harmonic number and normalized time. Rate and mean absolute spatial derivative (MASD) profiles are displayed on the right panels (see Methods for a description on the calculation of the profiles and their error bars). At the lowest level (19 dB relative pure tone threshold, or dBreTh), this fiber responds with peaks at integer neural harmonic numbers and valleys in between in both rate and MASD representations. As level increases (34 and 44 dBreTh), the rate profile saturates, degrading the pitch representation. In contrast, the MASD profile is relatively robust with level and thus better accounts for the psychophysical observation that the perception of pitch does not degrade with increases in level.

Compared with lower CF fibers, AN fibers with higher CFs are associated with relatively better cochlear frequency selectivity. Cedolin and Delgutte (2007) verified this by demonstrating that high CF AN fibers responded to harmonic complex tones with better harmonic resolvability (as indicated with stronger oscillations in their rate profiles) than lower CF fibers. In contrast to the rate, the MASD representation depends on phase locking and is
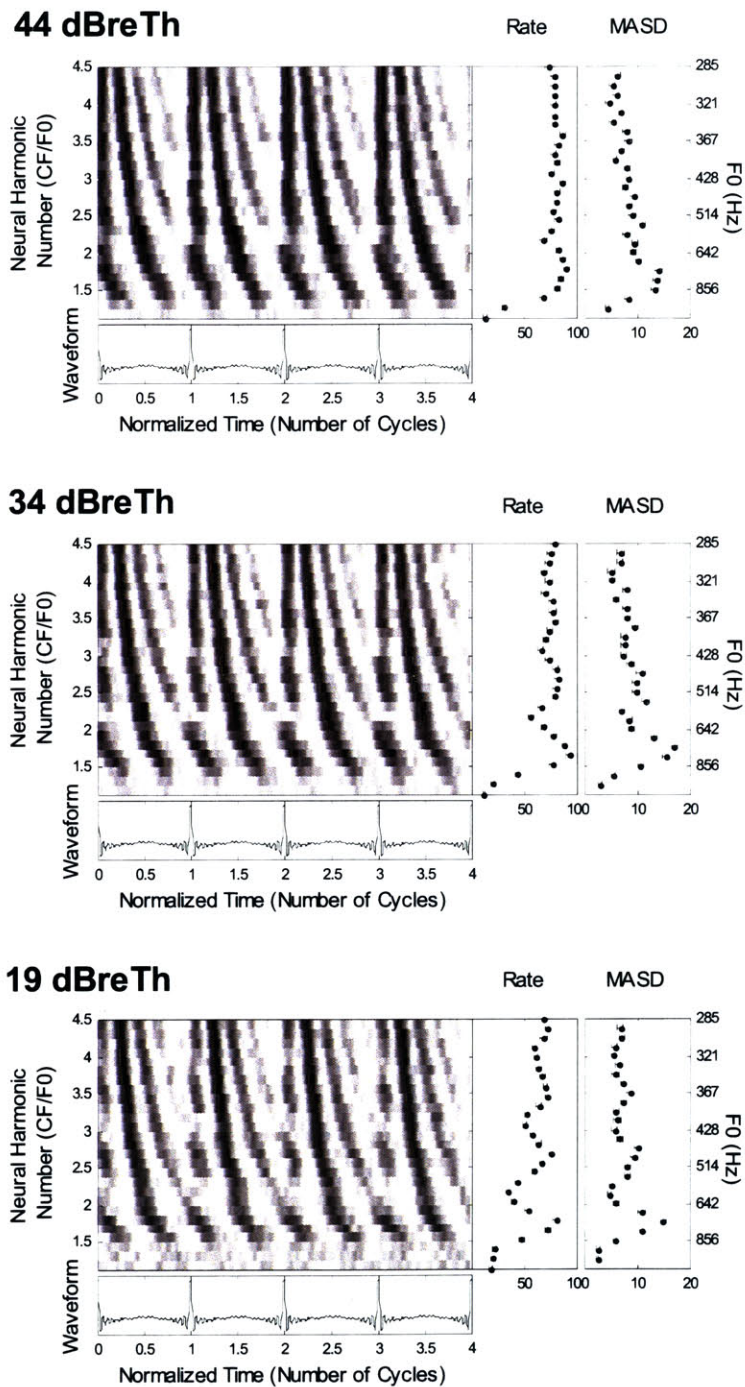
Figure 4: Data taken from Cedolin (2006). Response of an AN fiber with CF 1284 Hz to harmonic complex tones. As level increases, the fiber firing rate saturates, resulting in a degraded mean rate profile. MASD representation remains relatively robust with increases in level.

weaker for higher CF fibers where phase locking is poor, which is consistent with the existence of an upper limit for the pitch of a complex tone with missing F0.

By requiring the presence of both phase locking and resolved harmonics, the spatio-temporal representation of pitch combines the two traditional temporal and spectral theories and possesses the advantages offered by both while overcoming some of their limitations.

Purpose of the thesis

The AN provides spatio-temporal pitch cues, which can better account for psychophysical observations compared to other types of cues based on temporal or spatial representations (Cedolin and Delgutte 2007). However, it is not known whether higher level processes in the brain are making use of these 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 CN, where several unit types display the desired sharpened spectral and temporal responses.

The aim of this thesis is to evaluate sensitivity of single CN units to the spatio-temporal pattern of their AN inputs and to compare the responses of sensitive and insensitive units to harmonic complex tones. To measure spatio-temporal sensitivity, we use "Huffman sequences," which are complex transient stimuli designed to systematically manipulate the timing of AN fibers at neighboring cochlear locations (Carney 1990). If a CN unit receives AN inputs across different CFs and is sensitive to their relative timing, it could respond differently to manipulations of the Huffman sequences. Such a unit can be interpreted to be sensitive to the spatio-temporal pattern of its inputs. We hypothesize that sensitive CN neurons either have rate representations of pitch that are similar to the MASD representations derived in the AN, or demonstrate an enhancement of the MASD cue to the pitch of harmonic complex tones with missing F0.

Materials and Methods

Surgery and Experimental set-up

We recorded data from the AN of three cats and the CN of six cats. We also analyzed existing data from both the AN (1) and CN (4) collected by Cedolin and Dreyer. Methods were

similar to those described by Cedolin (2006) for the AN. Cats were anesthetized with Dial in urethane (75 mg/kg), with supplementary doses given as needed. The tympanic bullae were opened to expose the round window. After removal of the posterior portion of the skull, the cerebellum was either retracted to expose the AN, or partially aspirated to expose the CN. Throughout the experiment the cat was given injections of dexamethasone (0.26 mg/kg) to prevent brain swelling and Ringer's solution (50 mL/day) to prevent dehydration. The cat's temperature was maintained at 37°C.

The cat was placed on a vibration-isolated table in an electrically-shielded, sound-proof chamber. A silver electrode was positioned near the round window to record the AN compound action potential in response to click stimuli in order to assess the condition and stability of cochlear function.

Sound was delivered to the cat's ear through a closed acoustic assembly driven by an electrodynamic speaker (Realistic 40-1377). The acoustic system was calibrated to allow accurate control over the sound pressure level at the tympanic membrane. Stimuli were generated by either a 16-bit or a 24-bit digital-to-analog converter using sampling rates of 20 kHz, 50 kHz, or 100 kHz. Stimuli were digitally filtered to compensate for the transfer characteristics of the acoustic system.

Action potentials were recorded with either glass micropipettes filled with 2 M KCl for AN and CN recordings or parylene-insulated tungsten electrodes for CN recordings. The electrode was inserted into the dorsal surface of the AN or CN and advanced with a micropositioner (Kopf 650). The recorded signal was then bandpass filtered and input into a spike detector. Times of spikes, and in some CN experiments spike waveforms, were recorded and saved for offline processing.

To search for single units, click stimuli at around 55 dB SPL were used in the AN, and a 60 dB SPL noise stimulus was used for CN recordings. Once a unit was isolated, its frequency tuning curve was measured (Kiang et al 1970) to determine its CF and threshold. Spontaneous activity was then measured over 20 seconds. After each experiment, each AN fiber's threshold and spontaneous rate were compared against a standard data set (Liberman 1978). Any fiber whose threshold deviated by more than 2 standard deviations from the data set was not included in the analysis.

15

Stimuli

Short Tone Bursts at CF

After determining a CN unit's CF, threshold, and spontaneous rate, its response to short tone bursts at CF (STBCFs) with 25 ms duration were measured as a function of level usually from 5 to 55 dB above threshold in increments of 10-20 dB. The CN unit type was 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.

Huffman Sequences

To test whether CN neurons were sensitive to the spatio-temporal pattern of their AN inputs, we adopted the method of Carney (1990) using "Huffman sequences," which are complex acoustic stimuli with a flat magnitude spectrum and a smooth phase transition centered at the neuron's CF. Figure 5 shows the magnitude and phase spectra of the Huffman stimuli as well as their waveforms. The value "r" in these sequences is an inverse metric of the phase transition bandwidth, where low r (r = 0.85) corresponds to a broad transition and high r (r = 0.95) corresponds to a sharp transition. Manipulating the steepness of the phase transition allowed us to systematically vary the relative timing of spikes across the tonotopic array of fibers. The stimulus with broadest phase transition bandwidth excites adjacent AN fibers near the transition frequency more coincidentally, while the sharpest bandwidth stimulus causes longer delays between the relative timing of adjacent AN fibers.
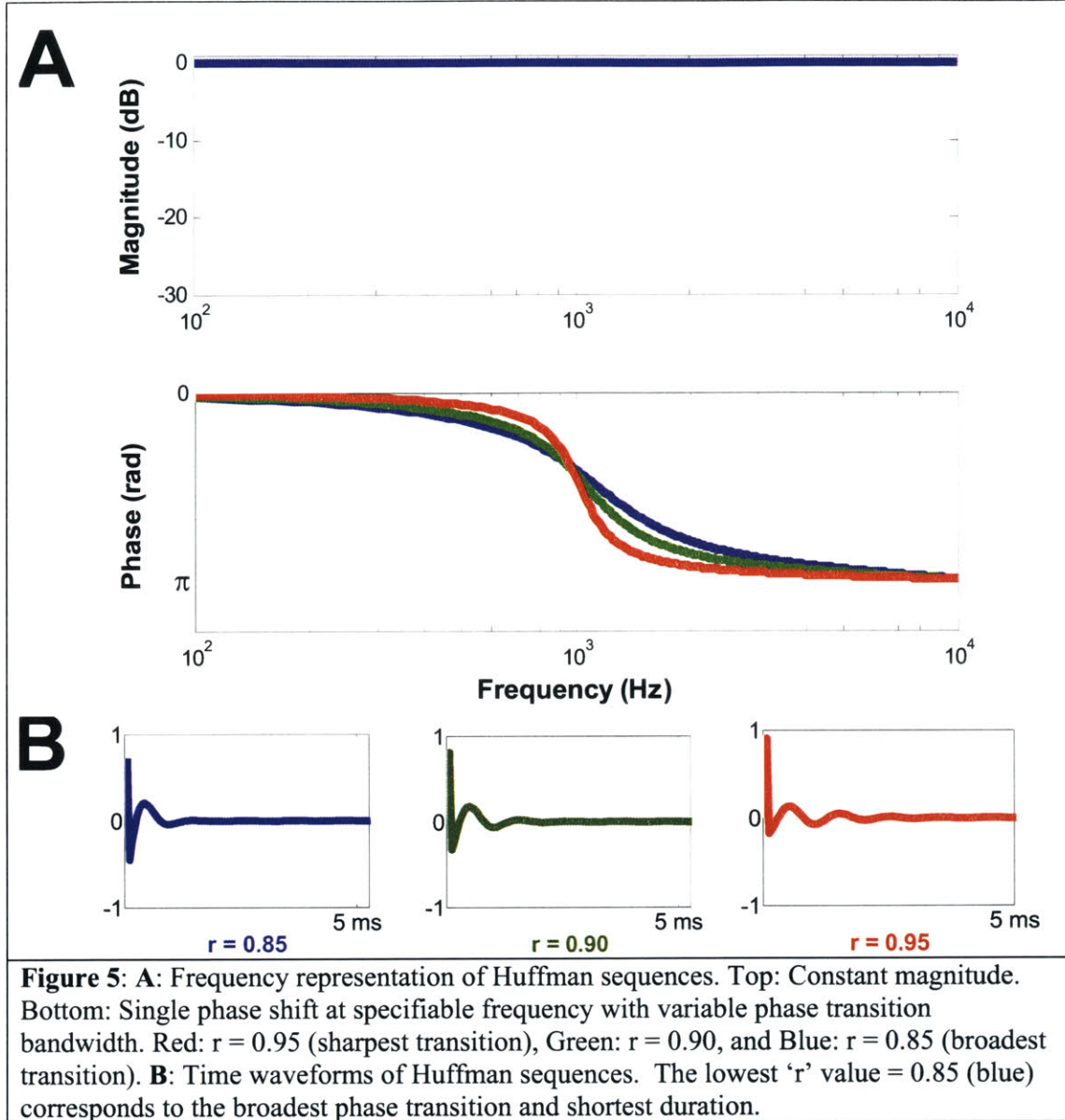
The Huffman stimuli were generated following methods described by Carney (1990). A positive impulse of unit amplitude was input into a filter with transfer function of the form

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

where

$$a_1 = \frac{2\cos\theta}{r} \quad , \quad a_2 = \frac{1}{r^2} \quad , \quad b_1 = 2r\cos\theta \quad , \quad b_2 = r^2 \quad , \text{ and } \theta = \frac{2\pi Ft}{Fs}$$

where Ft is the frequency of the phase transition and Fs is the sampling frequency. Fs was set to either 20, 50, or 100 kHz depending on the CF, and Ft was set to the neuron's CF. All of the

Figure 5: **A**: Frequency representation of Huffman sequences. Top: Constant magnitude. Bottom: Single phase shift at specifiable frequency with variable phase transition bandwidth. Red: r = 0.95 (sharpest transition), Green: r = 0.90, and Blue: r = 0.85 (broadest transition). **B**: Time waveforms of Huffman sequences. The lowest 'r' value = 0.85 (blue) corresponds to the broadest phase transition and shortest duration.

waveforms produced by this method were entirely damped out within 50 ms. Each trial consisted of a sequence of 6 stimuli (3 ascending r-values (0.85, 0.9, 0.95) and 3 descending r-values (0.95, 0.9, 0.85)) for a total duration of 300 ms. This was done to minimize any possible effect of the order in which the stimuli were presented. Spike times from the same r-value stimulus from the ascending and descending parts of the sequence were combined. Responses to these sequences were recorded for at least 100 trials (and typically between 200-300 trials) with no interruption. The sound pressure level of each stimulus was initially set to the approximate

threshold for response to Huffman stimuli (in many cases, this happened to be around 15-20 dB above pure tone threshold (dBreTh)) and then increased in 5 or 10 dB increments.

Harmonic Complex Tones

Methods for recording responses of CN units to missing fundamental harmonic complex tones were similar to those described by Cedolin (2006) for the AN. Stimuli were complex tones with equal-amplitude harmonics located at harmonic numbers 2-20 in 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 of the 10 s stimuli were repeated 20 times with no interruption. Spike times from the same F0 from the ascending and descending parts of the sequence were combined. Levels were typically initially set at 15-20 dB per component above pure tone threshold and then increased in 10 or 20 dB increments.

Data Analysis

CN Unit Classification

Based on its responses to 25 ms STBCF stimuli at several levels, each unit was classified as one of the following types: primary-like-with-notch (Pri-N), pauser/buildup, sustained chopper (Chop-S), transient chopper (Chop-T), phase-locker (PL), high-sync (HiS), low rate (LR), primary-like (Pri), flat (F) or onset (On). For consistency in unit typing, a decision tree was constructed to categorize units. Even though we were able to use the tree to confidently group most CN units into discrete categories, several units had complex response patterns that appeared to lie between two possible types. When possible, we used our judgment to decide the more likely type. However in some cases, we were unable to confidently select a group, and these units are labeled as unclassified.

First, each PSTH was manually examined for pauses, which occur in primary-like-with-notch (Pri-N) and pauser unit types. Pri-N units were defined by a pause in activity of 0.5 to 2 ms after the onset response, followed by sustained irregular discharge throughout the rest of the response. Examples of response patterns of two Pri-N units to STBCFs are shown in the top panels of figures 21 and 22. Pausers exhibited an onset response followed by a pause longer than those typically found in Pri-N units, 5-10 ms (Godfrey et al 1975) and then a gradual buildup. If there was little or no onset response followed by a late buidup of discharge, the unit was labeled as a build-up. Cells with pauser and build-up response patterns are usually found in the dorsal CN. Because this study was more interested in temporal processing in the ventral CN, recordings from these two unit types were not included in the analysis.

Unlike any other unit type, choppers have regular firing patterns. As discussed in Background, these units are further categorized as sustained choppers (Chop-S), which chop throughout the response, and transient choppers (Chop-T), which chop at the onset followed by irregular discharge.

Figure 6A shows the responses of a Chop-S unit (CF = 1725 Hz) to STBCFs at three levels. The left panel depicts the PSTH patterns, which are very regular throughout the duration of the response. The patterns show chopping intervals unrelated to the stimulus frequency, as well as poor phase locking. The CV (a metric for irregularity in the unit's discharge pattern) is the ratio of the standard deviation of the ISIs to their mean. The middle panels of figure 6A show CV as a function of time. The mean CVs, calculated from 12 to 20 ms after the stimulus onset, are very low for Chop-S units (< 0.35) and are generally constant throughout the response. Synchronization index (SI) is a metric for the tendency of a unit to fire at a particular phase of the stimulus and is calculated as the synchronized spike count divided by the total spike count (calculated from a period histogram with 100 bins / period). The right panel shows that this unit has low SIs at all three levels.

The chopping pattern of Chop-T units is restricted to the onset, and CV in these units increases with time. Figure 6B shows responses of a Chop-T unit (CF = 2197 Hz) to STBCFs. This unit chops during the first 5 ms of response, followed by irregular discharge. Because the responses are regular only at the beginning, Chop-Ts have CV values that increase with time and are slightly higher than those found in Chop-S units ($0.35 < CV < 0.5$). Similar to Chop-S units, Chop-Ts do not phase lock well to STBCF stimuli, as indicated by the low SIs on the right panel.
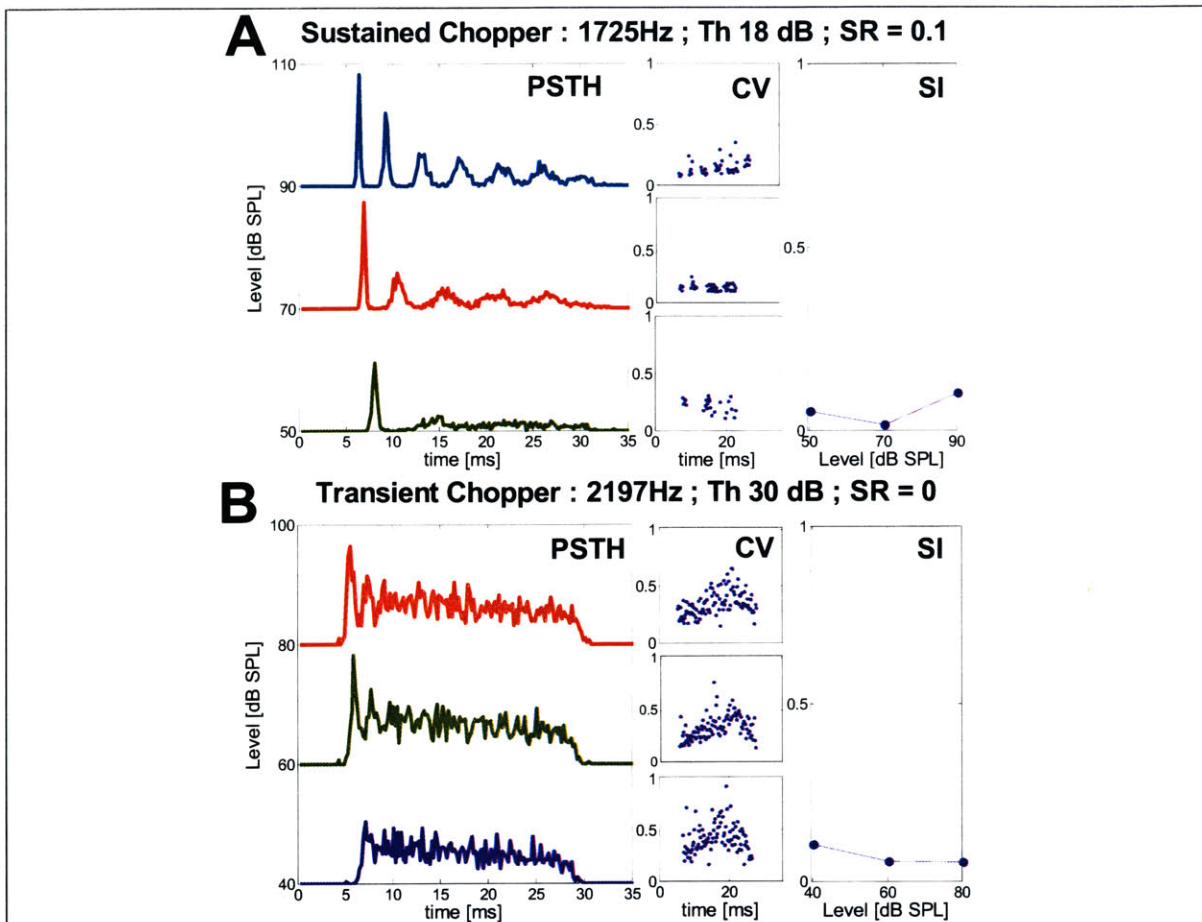
Figure 6: CN unit responses to STBCF stimuli. **A.** Chop-S with CF 1725 Hz. **B.** Chop-T with CF 2197 Hz. In the left panel of each figure, PSTHs are displayed at three levels (dB SPL). The coefficient of variation (CV) is plotted in the middle panel and was used as a metric of spike irregularity. Chop-Ts have slightly larger CV (0.35 - 0.5) values than Chop-S units (0 - 0.35). The low synchronization indices (SIs) on the right panels at all levels is consistent with poor phase locking in choppers.

The response patterns of low CF (< 1 kHz) units to STBCFs are characterized by such strong phase locking that other characteristics of the pattern that would normally be useful for classification are indiscernible. We classify many low CF CN units into two categories (phase locker (PL) and high-synchrony (HiS)) based on the degree of phase locking.

The top-left panel of figure 7A displays the PSTHs and period histograms of a PL unit with CF 504 Hz at four levels. The bottom left panel of the same figure shows the period histograms, which illustrate that most of this unit's spikes occur at a specific phase of the stimulus. The SI calculated from these histograms is plotted against level on the right panel. For

units with CF less than 1 kHz, SIs greater than 0.7 usually meant that the PSTHs displayed enough phase locking to make classification difficult.

HiS units make up a subset of the PL units that display better phase locking than typically found in AN fibers with the same CF. Figure 7B displays the responses of a HiS unit (CF = 496 Hz) to STBCFs. Characteristics of the PSTHs and period histograms are very similar to those of regular PL units, except that spikes are restricted to a smaller fraction of the stimulus cycle. The SI plot on the right panel indicates that the HiS unit phase locks better to the STBCF stimulus compared to the PL unit in figure 7A.
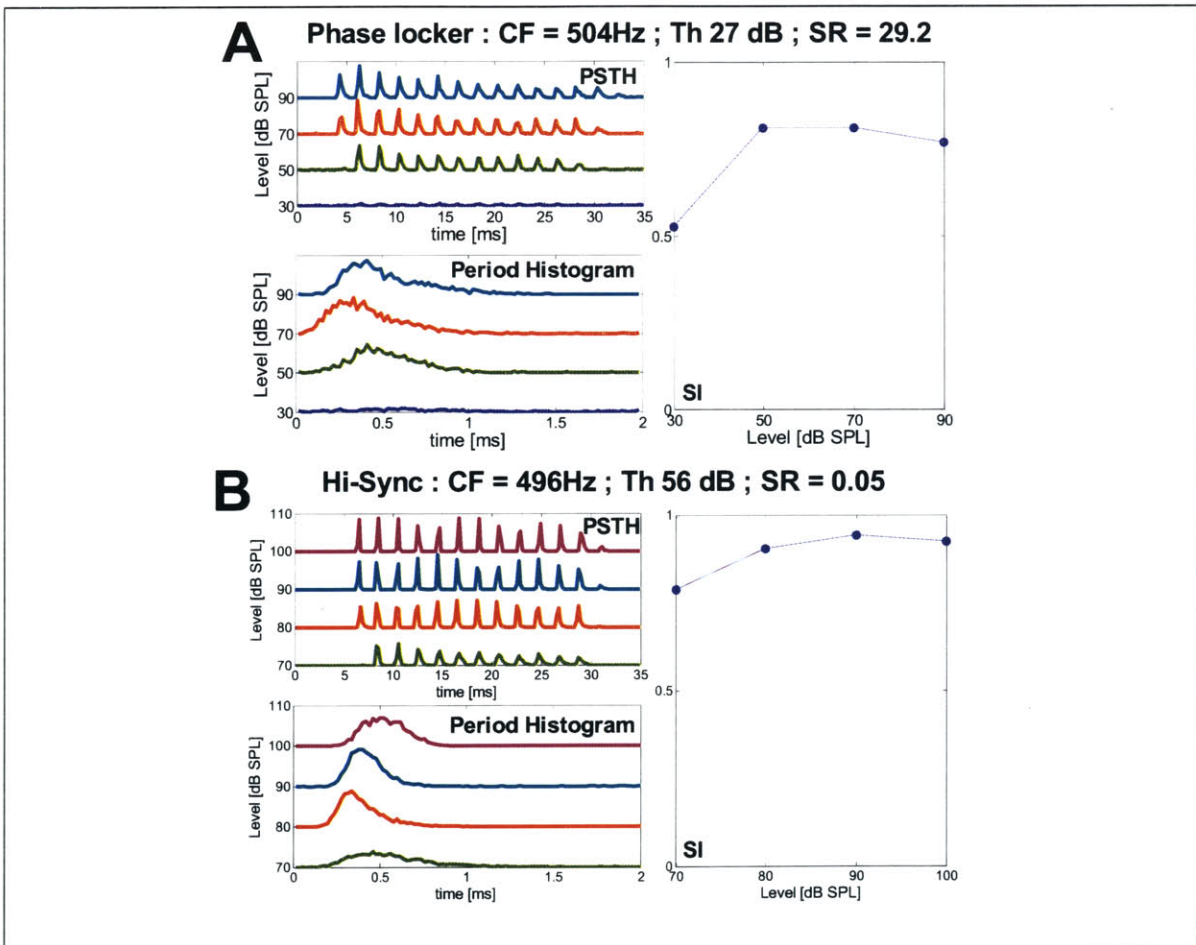

Figure 7. CN unit responses to STBCF stimuli. **A**. PL with CF 504 Hz. **B**. HiS with CF 496 Hz. In the top-left and bottom-left panels of each figure, PSTHs and period histograms are displayed at four levels (dB SPL). SI is plotted as a function of level on the right. PLs have similar SI values as those observed in the AN, while HiS units have higher SIs.

Low rate (LR) units are defined by having average spiking rates less than 100 spikes/s at all levels of response to STBCF stimuli.

The remaining units were classified as primary-likes (Pris), Flats, or onsets (Ons) based on the ratio of onset (first 10 ms of the response) to steady state (later 15 ms of the response) average firing rates. Flats had ratios less than 1.2, Pris had ratios between 1.2 and 5, and Ons had ratios greater than 5.

The PSTH shapes of Pri units are characterized by a high onset response followed by a gradual decline to a steady discharge until the end of the stimulus. This is similar to AN fiber responses to tone burst stimuli at the same frequency. Figure 8B shows the response of a Pri unit (CF = 1419 Hz) to STBCFs at three levels. The unit's onset and steady state ratios on the right panel are within the Pri range (1.2 to 5) at all three levels.

Figure 8A shows the STBCF responses of a Flat unit (CF = 1419 Hz) at three levels. This unit's PSTHs (on the left) have a rectangular envelope with no prominent onset response. This unit has nearly the same onset and steady state rates at all three levels recorded.

On-L units are characterized by a sharp onset peak followed by some later activity. Figure 8C shows an On-L unit's (CF = 1938 Hz) responses to STBCFs. The first spike latency histograms (bottom left) are very similar to the onset patterns of the PSTHs (top left), which also show some late activitiy. Thus, on average, for each stimulus presentation, this unit fired once at the onset and sometimes would fire again during the remainder of the stimulus. The right panel plots the onset and steady state rates, which at the higher two levels, have ratios greater than 5.

Figure 9A shows the maximum SI calculated for each CN unit (n=179) plotted as a function of frequency (which was always at CF). Unit types are distinguished by different symbols (as indicated in the legend of figure 9B). AN data (gray dots) were taken from Johnson (1980). All PL and HiS units had CFs less than 1000 Hz, and HiS units had SIs greater than 0.9, while PLs had SIs between 0.7 and 0.9. Generally, Pri and Flat units followed the trend of the AN data fairly closely, although some displayed lower SIs. Both types of choppers and Ons usually had higher CFs (>1 kHz) and had poorer phase locking than AN fibers and other CN units.

Figure 9B shows the standard deviation plotted against the mean ISI for each CN unit (n=179). Nearly all chopper units fall below the straight line, indicating they have CV values less than 0.5. Consistent with our classification scheme, Chop-Ts are closer to the line
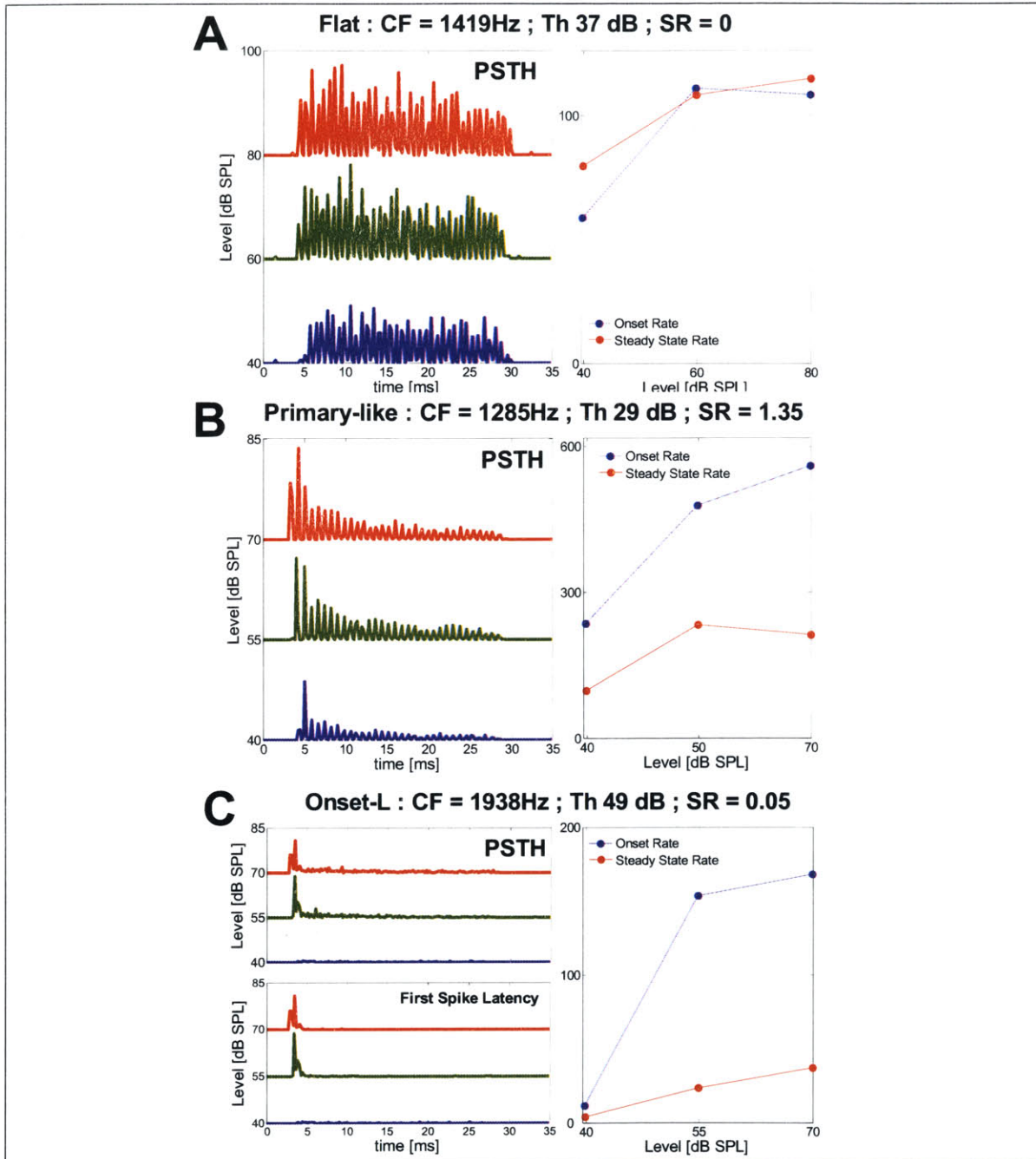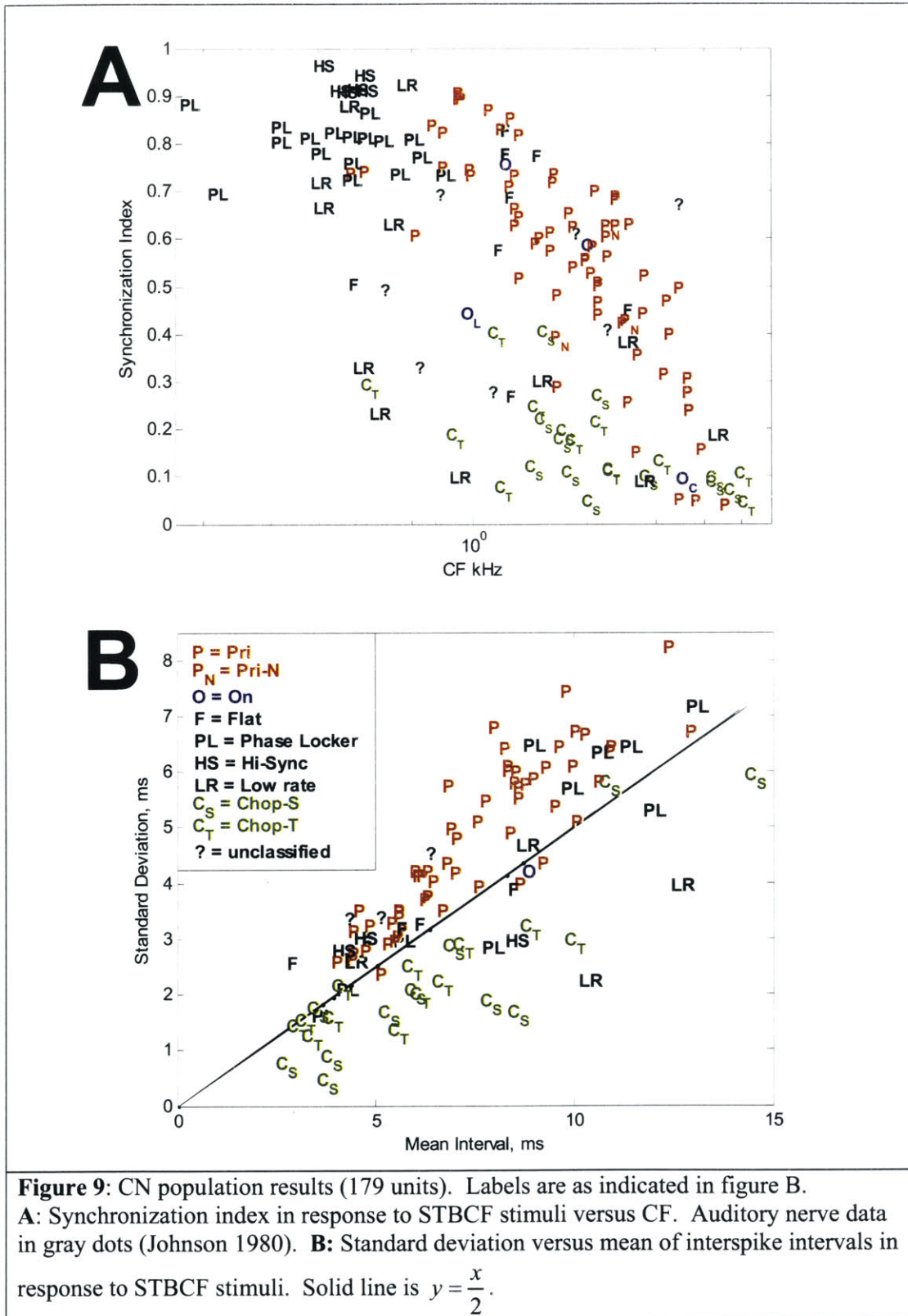
Figure 8: CN unit responses to STBCF stimuli. **A.** Flat with CF 1419 Hz. **B.** Pri with CF 1285 Hz. **C.** On-L with CF 1938 Hz. In the left panel of each figure, PSTHs are displayed at three levels (in dB SPL). The average rates were calculated over the onset (first 10 ms of the response) and the steady state (the remaining 15 ms of the response) and plotted against level in the right panels. Flats have onset to steady state rate ratios less than 1.2, Pris have ratios between 1.2 and 5, and Ons have ratios greater than 5. The left panel of figure C also shows the first spike latency histograms of the On-L unit.

Figure 9: CN population results (179 units). Labels are as indicated in figure B.
A: Synchronization index in response to STBCF stimuli versus CF. Auditory nerve data in gray dots (Johnson 1980). **B:** Standard deviation versus mean of interspike intervals in response to STBCF stimuli. Solid line is $y = \dfrac{x}{2}$.

(corresponding to CV values between 0.35 and 0.5) compared to Chop-S units, which fall further below (CV values less than 0.35). Furthermore, most of the other unit types fall above the line, indicating these units have CV values greater than 0.5 and have irregular discharges.

Huffman Sequences

Single unit recordings from AN fibers in response to the Huffman sequences formed the baseline for comparison with the CN data. Neural sensitivity to the rate of the phase transition was evaluated by looking for overall rate changes as well as temporal changes in the response PSTHs and comparing these changes to those found in AN fibers with similar CFs. We only analyzed units for which responses to Huffman stimuli were measured for at least two levels.

Several steps were taken to evaluate the overall rate sensitivity of an AN fiber or CN neuron to the phase transition bandwidth of the Huffman stimuli. After combining spike times from the ascending and descending parts of the stimulus, PSTHs were constructed with duration 150 ms (3 r-values, each with 50 ms duration) and bin width 0.1 ms. Three 15 ms windows were set for each PSTH corresponding to 2.4-17.4 ms after the onset of each stimulus. All AN and CN responses to the transient stimuli occurred strictly within these windows.

A two-way analysis of variance was performed on the total spike rate within the 15 ms windows with r-value and trial number as factors. This was followed by a multiple comparison test to order the three stimuli by their average rate responses and to evaluate the significance of this order. Each recording was classified as either non-monotonic (if the response to $r = 0.90$ was significantly greater or less than responses to the other two stimuli), monotonic (if the recording was not non-monotonic and the response to $r = 0.85$ was significantly greater than the response to $r = 0.95$, or vice versa), or flat (if the mean spike counts were not significantly different from one another) with a 95% confidence interval. Less than 3% of AN and CN recordings displayed non-monotonic responses. Because we use a rate difference metric (defined below) to compare responses to the $r = 0.85$ and $r = 0.95$ stimuli, only data that were monotonic or flat with changes in 'r' were considered for analysis.

An overall rate difference (RD) (Carney 1990) was computed for each recording by subtracting the rate in response to the stimulus with the sharpest phase transition (r=0.95) from the rate in response to the stimulus with the broadest phase transition (r=0.85). A normalized

rate difference (RD$_N$) was calculated by dividing the RD by the square root of the sum of the squares of the standard deviations of spike counts across trials for the two stimuli:

$$RD_N = \frac{r_{0.85} - r_{0.95}}{\sqrt{\sigma_{0.85}^2 + \sigma_{0.95}^2}}$$

We used a RD$_N$ (unit-less) instead of a RD (spikes/s) to facilitate comparison between units with different firing rates. This is particularly important in the CN due to diversity of unit types and their characteristics.

To compare temporal discharge patterns between units in the AN and CN, the average overall duration of the responses in the PSTHs were quantified. Each PSTH (with bin width 0.1 ms) was smoothed by convolution with a three-point triangular filter and windowed with the three 15 ms windows defined earlier. The mean and standard deviation of the smoothed PSTH outside the windows were calculated, and a threshold was set at five standard deviations above the mean. All local maxima of the smoothed response within the windows were detected as peaks. The total duration of the response was defined as the time between the threshold crossing before the first peak and the crossing after the last peak within each window.

In order for a CN unit to be considered sensitive to the phase transition bandwidth, changes in its responses to the three stimuli had to significantly differ from those observed in AN fibers with similar CFs. In short, a CN unit was said to be sensitive if its responses to at least two levels of Huffman stimuli resulted in RD$_N$ values that were significantly different from those measured in the AN at similar levels. Details specifying how this was done are described in Results.

Harmonic Complex Tones

Spike times in response to each 180 ms (200 ms minus the 20 ms overlap) duration complex tone were recorded and used for analysis. The non-scaling conduction delay typical 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 scaling invariance may be applied. In CN recordings, an additional 1 ms was subtracted from spike times to account for travel time between the AN and CN.

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), as shown in figure 2. The average discharge "rate-place" profile was obtained by integrating each period histogram over time at each F0 and plotting as a function of neural harmonic number. The mean absolute spatial derivative (MASD) profile approximates a differentiation by taking the first difference between adjacent rows and then integrating the absolute value of the difference over time. Both the rate and MASD profiles were smoothed by convolution with a three-point triangular filter.

Following methods by Cedolin (2006), in order to evaluate the statistical properties of the rate and MASD estimates, bootstrap resampling (Efron and Tibshirani 1993) was performed on each recording. One hundred resampled data sets were generated by drawing with replacement from the set of spike trains over all 20 trials. Responses to ascending and descending parts of the stimulus were bootstrapped independently of each other. The rate and MASD profiles were computed from each bootstrapped set, and the standard deviations across data sets were used as error bars.

To quantify harmonic resolvability, the following mathematical function was independently fit to each rate and MASD profile as a function of neural harmonic number (n = CF/F0) using a least squares method:

$$f(n) = A \cdot \cos(2\pi \cdot \delta \cdot n) \cdot e^{-\frac{n}{n_o}} + B \cdot e^{-\frac{n}{n_o}} + C$$

where the five parameters are: the amplitude (A), frequency (δ), time constant of decay (n_0), amplitude of the baseline (B), and the DC component (C). Based on the values of the parameters of the best fitting curves, we derived the "harmonic strength" (HS) for each bootstrapped set as a measure of the amplitude of the harmonic oscillations in the spatial pattern of the mean rate or MASD profiles, normalized by the median of the standard deviations across F0 values.

Results

Our results are based on recordings from 84 AN fibers in four cats, from which we recorded responses to Huffman sequences. In addition, we used AN recordings in response to complex tone stimuli from Cedolin (2006).

Though we recorded from 179 CN units in ten cats which we classified into types, we were only able to maintain contact long enough to record responses to Huffman or complex tone stimuli or both from 83 units (CFs 100 - 4060 Hz).

Huffman sequences

We recorded 552 responses to Huffman stimuli from 84 AN fibers (with CFs 132 - 3611 Hz) in four cats. Of these, 47 fibers had high SR (>18 spikes/s), 33 had medium SR (0.5 to 18 spikes/s), and 4 had low SR (<0.5 spikes/s). We also analyzed 260 responses from 53 CN units (with CFs 100 – 3533 Hz) in ten cats, comprised of 10 PL, 4 HiS, 28 Pri, 3 Pri-N, 1 On, 1 Flat, 2 Chop-T, and 4 Chop-S units. To be included in the analysis, response to Huffman stimuli had to be recorded at at least two levels.

Auditory Nerve responses

Responses of AN fibers to the three Huffman stimuli with varying phase transition bandwidths were quantified to provide a baseline for comparison with units in the CN. Figure 10 shows the response of an AN fiber with CF 545 Hz and SR 101 spikes/s to Huffman stimuli. The top panel shows the stimulus waveforms for the three r-values. Each column corresponds to one phase transition bandwidth, with the broadest transition on the left (r = 0.85) and the sharpest transition on the right (r = 0.95). Each row corresponds to one stimulus level (in dBreTh). Each panel represents the PSTH with duration 15 ms and bin width 0.1. RD and RD$_N$ at each level are labeled in the top right corner of each row. This format is repeated in all figures depicting fiber and unit responses to Huffman stimuli.

In figure 10, the peaks are separated by intervals of about 1 / CF. At each level, this fiber responded to the three stimuli with approximately the same number of peaks, though there is an additional small peak at the beginning of the response to r = 0.95 at the highest level (69 dBreTh). The rate response of this fiber increasingly favors the sharp-transition stimulus as level increases.

Figure 11 shows the responses of another AN fiber with a similar CF (440 Hz) but a low SR (1.2 spikes/s). At both levels, the response consists of several peaks separated by approximately 1 / CF, or 2.3 ms. The largest first peak occurs in response to the sharp-transition stimulus, and the largest second peak is in response to the broad-transition stimulus. This is consistent with the shape of the click stimulus waveforms in figure 5B, where the r = 0.95 stimulus is largest at the onset, and the r = 0.85 stimulus has the largest second maximum (which occurs near time 1 / CF). This fiber responds at the lower level (50 dBreTh) with 3 peaks for each of the three stimuli. At the higher level (60 dBreTh), the response to the sharp-transition
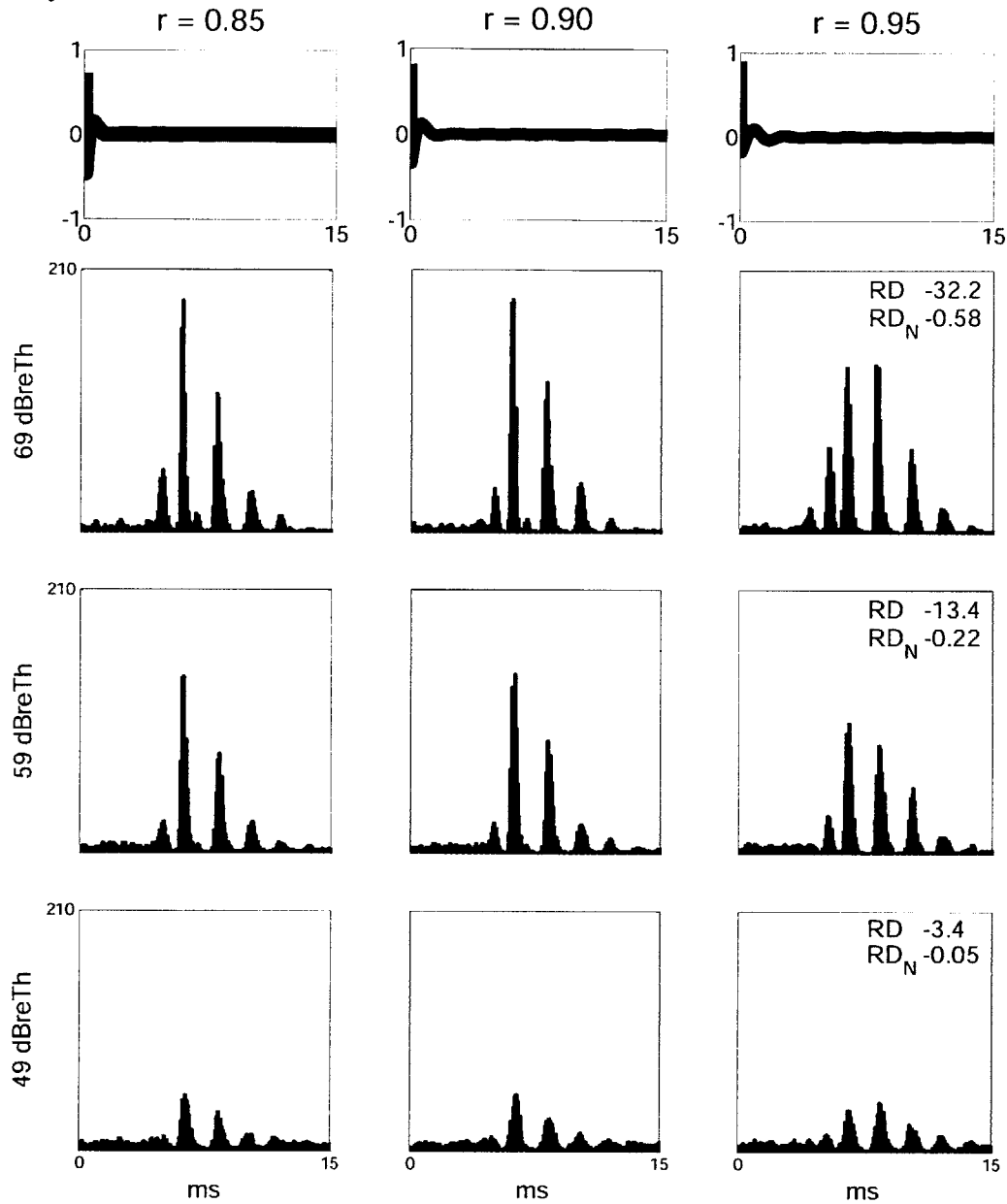
Figure 10: Response of an AN fiber with CF 545 Hz and SR 101 spikes/s to Huffman stimuli. Rows of panels represent PSTH responses (duration 15 ms with bin width 0.1 ms) at different levels (indicated on the left, in dBreTh). Columns represent responses to different phase transitions (indicated on the top, r = 0.85 corresponds to the stimulus with broadest phase transition and shortest duration). Displayed in the panels on the right are rate difference (RD) and normalized rate difference (RD_N) metrics as described in the text. Positive values indicate a better overall response to the r = 0.85 stimulus. The same format is repeated in later figures depicting unit responses to Huffman sequences.

Auditory nerve fiber: CF = 440 Hz, SR = 1.2

Figure 11: Response of an AN fiber with CF 440 Hz and SR 1.2 spikes/s to Huffman stimuli.
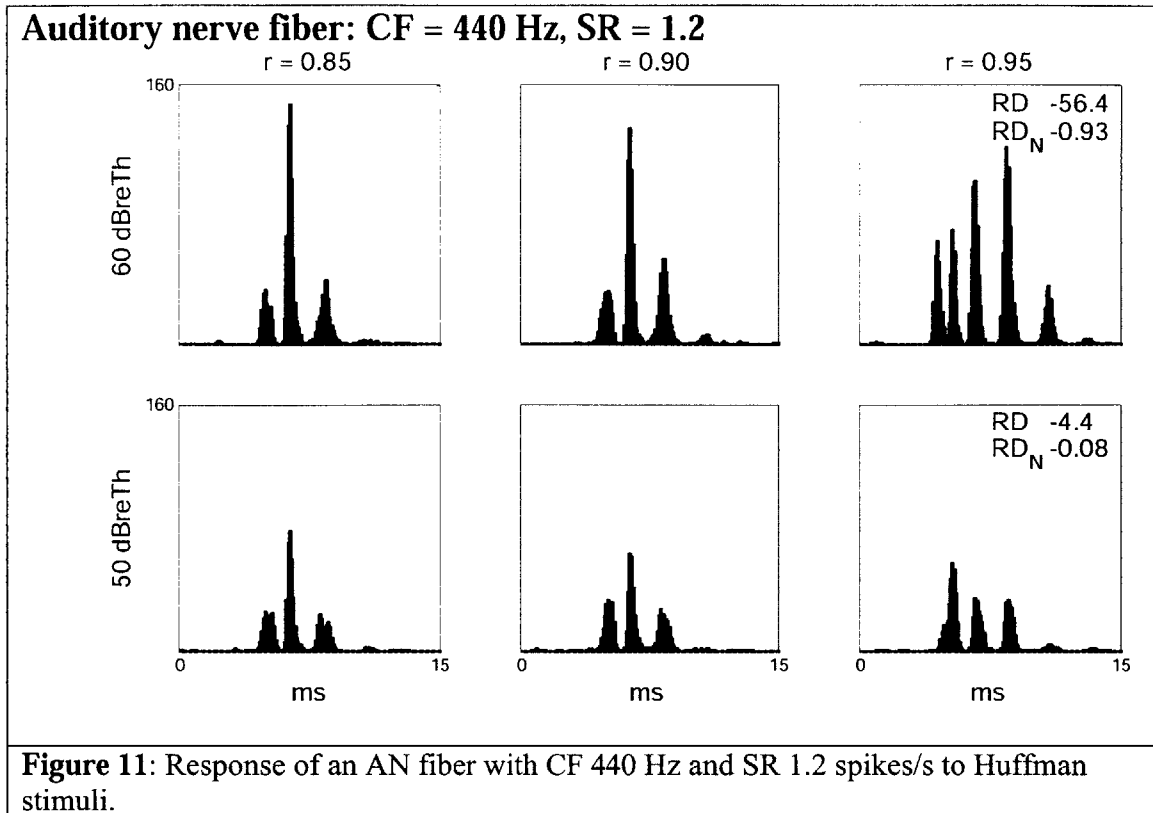
stimulus contains two additional peaks, which are missing in the response patterns to the other two stimuli. As a result, this fiber's rate responses favor the $r = 0.95$ stimulus at the high level but has no preference at the lower level. This fiber's changes with level are more pronounced than the high SR fiber depicted in figure 10.

Figure 12 shows responses of a higher CF (2552 Hz), high-SR (37 spikes/s) AN fiber at four levels. All of the responses have weaker peaks compared to those observed in the lower CF fibers. This is due to the degradation of phase locking in high CF fibers. Furthermore, the overall durations of the responses in figure 12 are shorter than those seen in figures 10-11, which is a reflection of the shorter stimulus waveform when the Ft is high. The RDs are very close to zero, suggesting that this high CF fiber prefers neither the broad- nor sharp-transition stimulus.

Figure 13 shows the RD_N as a function of level (in dBreTh) for all AN recordings. Ideally, these data should have been plotted as a function of level relative to threshold for Huffman stimuli. However, in our initial experiments, our sampling of levels was often too sparse to confidently estimate this threshold. Therefore we used levels defined in dB relative to
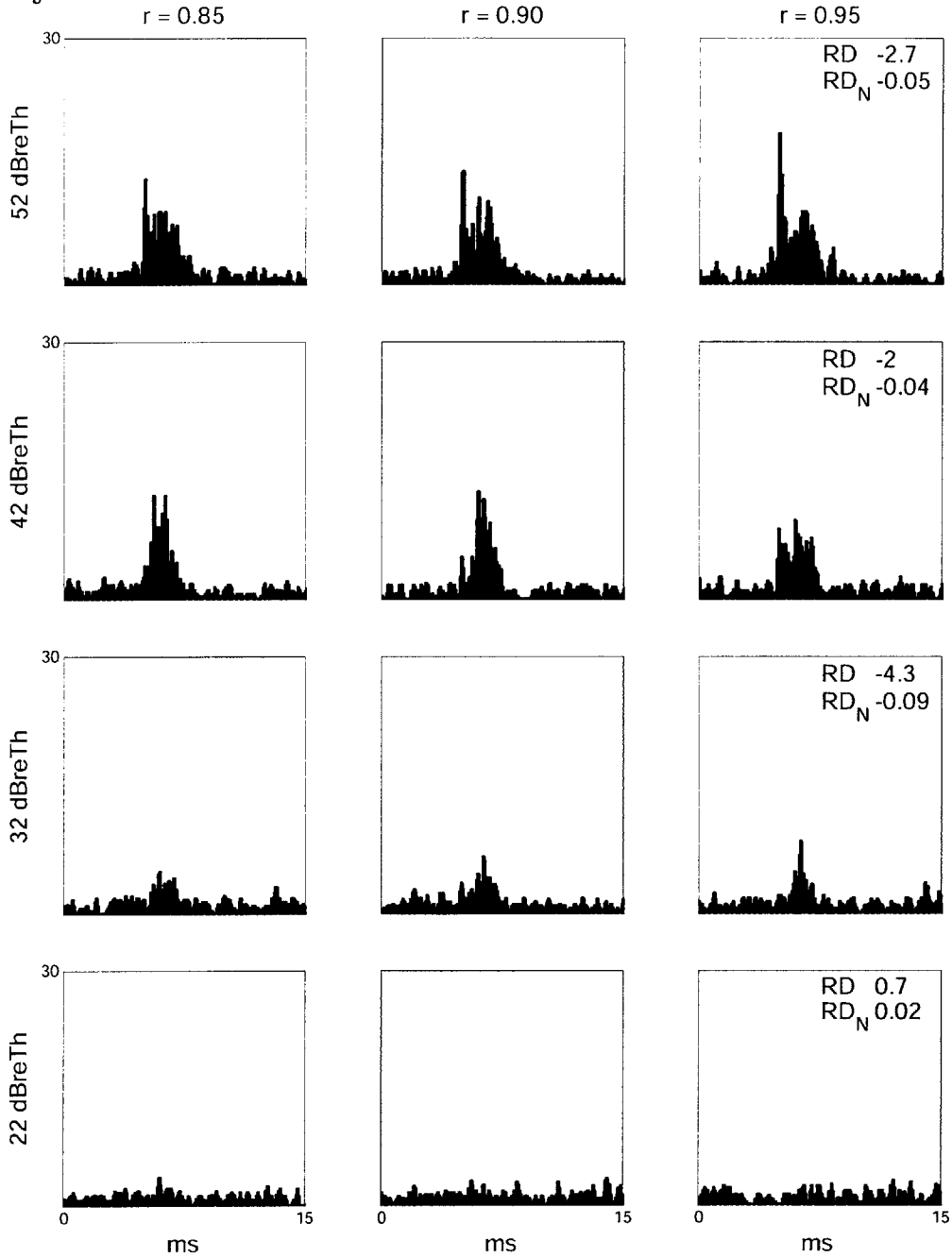
Figure 12: Response of an AN fiber with CF 2552 Hz and SR 37 spikes/s to Huffman stimuli.

the pure tone threshold at CF derived from tuning curve measurements. Actual Huffman thresholds are typically 15-30 dB higher than pure tone thresholds.

Data from fibers with low, medium and high SR are plotted with different symbols as indicated in the legend. Colored straight lines connect the data points corresponding to

31

recordings from the three example fibers (green: F10, blue: F11, and red: F12). Near threshold (< 40 dBreTh), RD_Ns were centered around zero, suggesting no preference for gradual or rapid phase transitions. When tested for dependence on the three stimulus conditions using ANOVA, 96% of the recordings below 40 dBreTh showed a flat effect.

    The stimulus with rapid phase transition (r = 0.95) is longer in duration. As level increased, RDs of most AN fibers became increasingly negative, indicating a preference for the sharp phase transition and/or a long duration stimulus. The longer duration responses of AN fibers to this stimulus combined with preferences in the rate responses to this stimulus suggest a saturation in the peak firing rate at high levels. This is most clearly seen for the fiber in figure 10, which responded to the r = 0.95 stimulus at the highest level with five peaks compared to the three-peak responses to the other two stimuli.
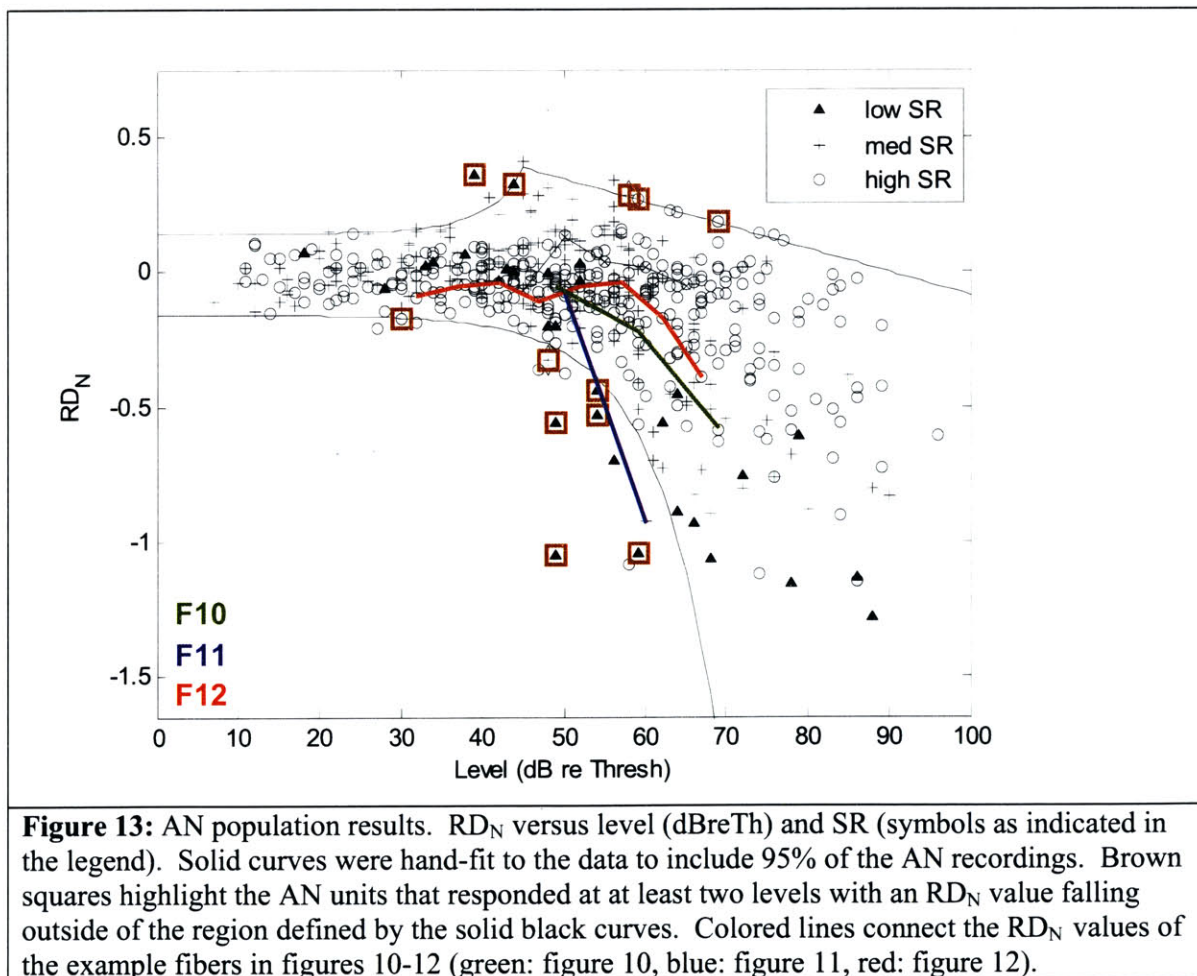

Figure 13: AN population results. RD_N versus level (dBreTh) and SR (symbols as indicated in the legend). Solid curves were hand-fit to the data to include 95% of the AN recordings. Brown squares highlight the AN units that responded at at least two levels with an RD_N value falling outside of the region defined by the solid black curves. Colored lines connect the RD_N values of the example fibers in figures 10-12 (green: figure 10, blue: figure 11, red: figure 12).

Figure 13 also shows that at levels greater than 45 dBreTh, fibers with lower SR tend to respond with more negative RD_N values compared to higher SR fibers. The blue line in the figure connects the data points from the medium SR fiber in figure 11, while the green and red lines both correspond to the high SR fibers in figures 10 and 12. The medium SR fiber responds better to the sharp-transition stimulus at comparable levels. In addition, the sharper slope of the blue line compared to the slower slopes of the blue and red lines suggest that lower SR fibers are more sensitive to changes in the level of the Huffman stimulus.

We fit line segments to the data in figure 13 such that just 5% of the AN responses lied outside the closed region defined by the fitted curves. Approximately the same number of AN responses fall above and below the top and bottom segments respectively. We further tried to fit the lines such that the levels of AN responses that fell outside this region had a fairly uniform distribution. 2 out of 4 low SR fibers, 1 out of 33 medium SR fibers, and 2 out of 47 high SR fibers had at least two Huffman responses that fell outside the region defined by the solid black curves in the figure. This result is summarized in Table I. The data points representing these

Table I

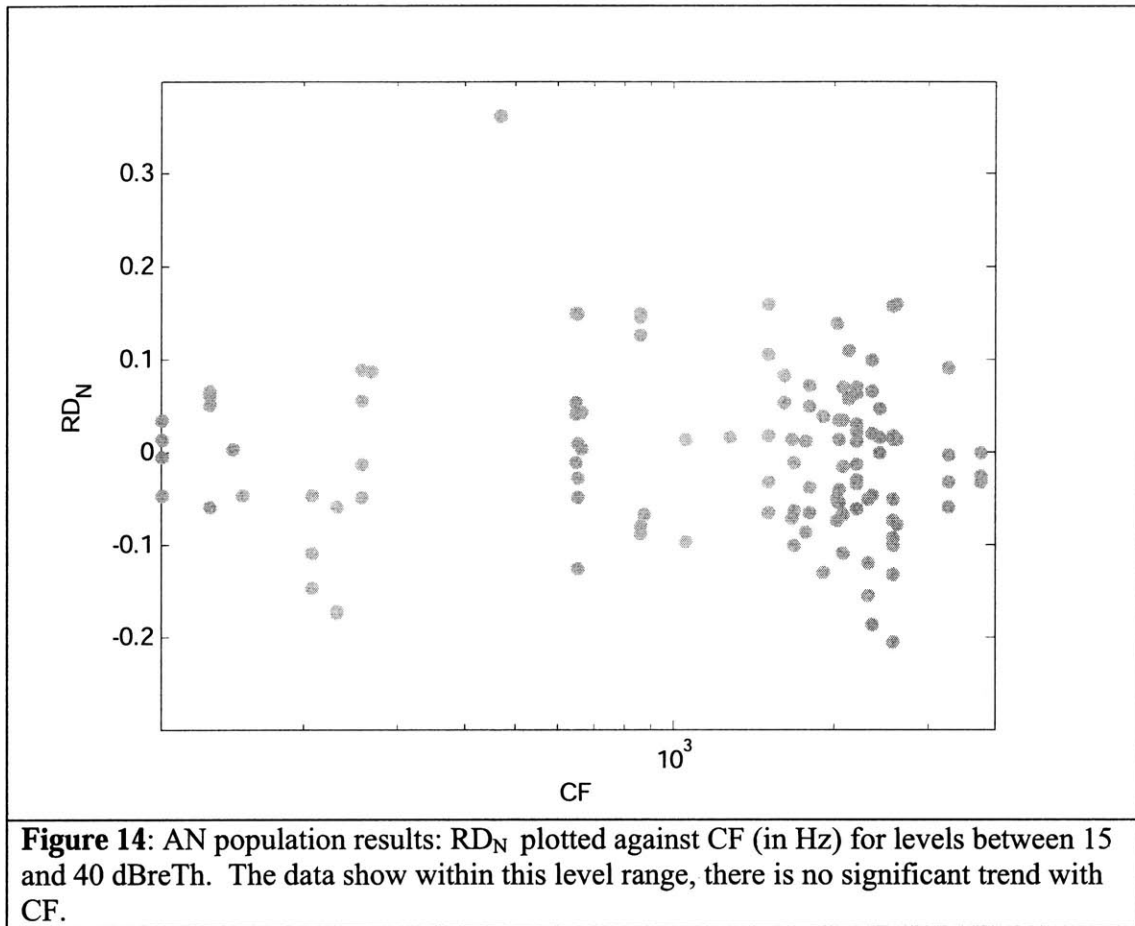|	Number Studied	Number "Sensitive"	Percent "Sensitive"
Low SR	4	2	50%
Medium SR	33	1	3%
High SR	47	2	4%
Total AN	84	5	6%

"sensitive" fibers have been highlighted in figure 13 with open brown squares.

Figure 14 shows RD_N as a function of fiber CF for levels between 15 and 40 dBreTh. There is no obvious dependence of RD_N on CF, justifying the pooling of data across CF.

Evaluating CN unit sensitivity

CN units were labeled as either phase sensitive ("PS", if its response at at least two levels fell outside the region defined in figure 13), or not phase sensitive ("NPS", if none or just one of its responses fell outside this region). PS units were further categorized by considering the direction in which they were sensitive.

The units that responded with more positive RD_N compared to AN fibers at at least two levels and never responded with more negative RD_N were defined as "PS+". These units showed a greater response to the stimulus with broad phase transition (r = 0.85), suggesting they prefer to

Figure 14: AN population results: RD_N plotted against CF (in Hz) for levels between 15 and 40 dBreTh. The data show within this level range, there is no significant trend with CF.

fire when their inputs are discharging in synchrony. This is consistent with a coincidence detection mechanism.

"PS-" units responded with more negative RD_N compared to AN fibers at at least two levels and never responded with more positive RD_N. Often, low-level recordings from these units were within the closed region, and as level was increased, the units' responses to the sharp-transition stimulus were greater.

A third type of phase sensitivity was defined as responding with a more positive RD_N than AN fibers at low levels and a more negative RD_N at higher levels, or "PS+/-." These units displayed characteristics consistent with a coincidence detector at low levels, but as level increased, the direction of sensitivity changed to be in the opposite direction.

Phase locker responses

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 PSTHs that are used to classify CN units. We therefore classified most low-CF units into two groups depending on whether the strength of phase locking exceeds that observed in the AN or not. We call these high-synchrony (HiS) and phase lockers (PL), respectively.

Figure 15 shows the responses to Huffman stimuli of one phase-locking CN unit (CF = 504 Hz). Responses of this unit to STBCFs are in figure 7A. The temporal response patterns to the three Huffman stimuli are similar in both the number of peaks and intervals between peaks, although there are subtle differences. The rate responses are also similar for the three stimuli, and the RD_Ns at all levels fall within the AN region, making this unit not phase-sensitive (NPS). This overall pattern is similar to that seen in most low-CF AN units.

Figure 16 shows results for another PL unit with a lower CF (100 Hz), with an extra panel at the top that shows the CN unit's response pattern to a STBCF. This unit shows a single wide peak in its PSTHs in response to each of the 3 Huffman stimuli, but the peak is clearly lower for the stimulus with sharp phase transition (r=0.95) than for other two stimuli. As a result of these height differences, RD_N is strongly positive at the higher two stimulus levels, making this unit PS+.

Only one other PL unit (CF = 205 Hz) had a single peak in its responses to Huffman stimuli but was NPS. It is not surprising that units with low CFs would respond to these transient stimuli with a single peak. The Huffman stimulus waveform has maxima that occur at intervals of approximately 1 / Ft. With such a low Ft (CF = Ft), the waveform would likely be damped out by the time of the next maximum after the onset. Single-peak responses were common in our Chop-S units, but none of these was PS+.

Hi-Sync responses

A subset of the PL group, HiS units respond to STBCF stimuli with better phase locking than found in AN fibers. Figure 17 displays the responses to Huffman stimuli of a HiS unit (CF = 475 Hz, responses to STBCFs are in figure 7B). The temporal response patterns for the middle two levels (42 and 47 dBreTh) show three peaks in response to the sharp-transition stimulus (r = 0.95) but only two peaks in response to the other two stimuli. As level increases from 42 to 52
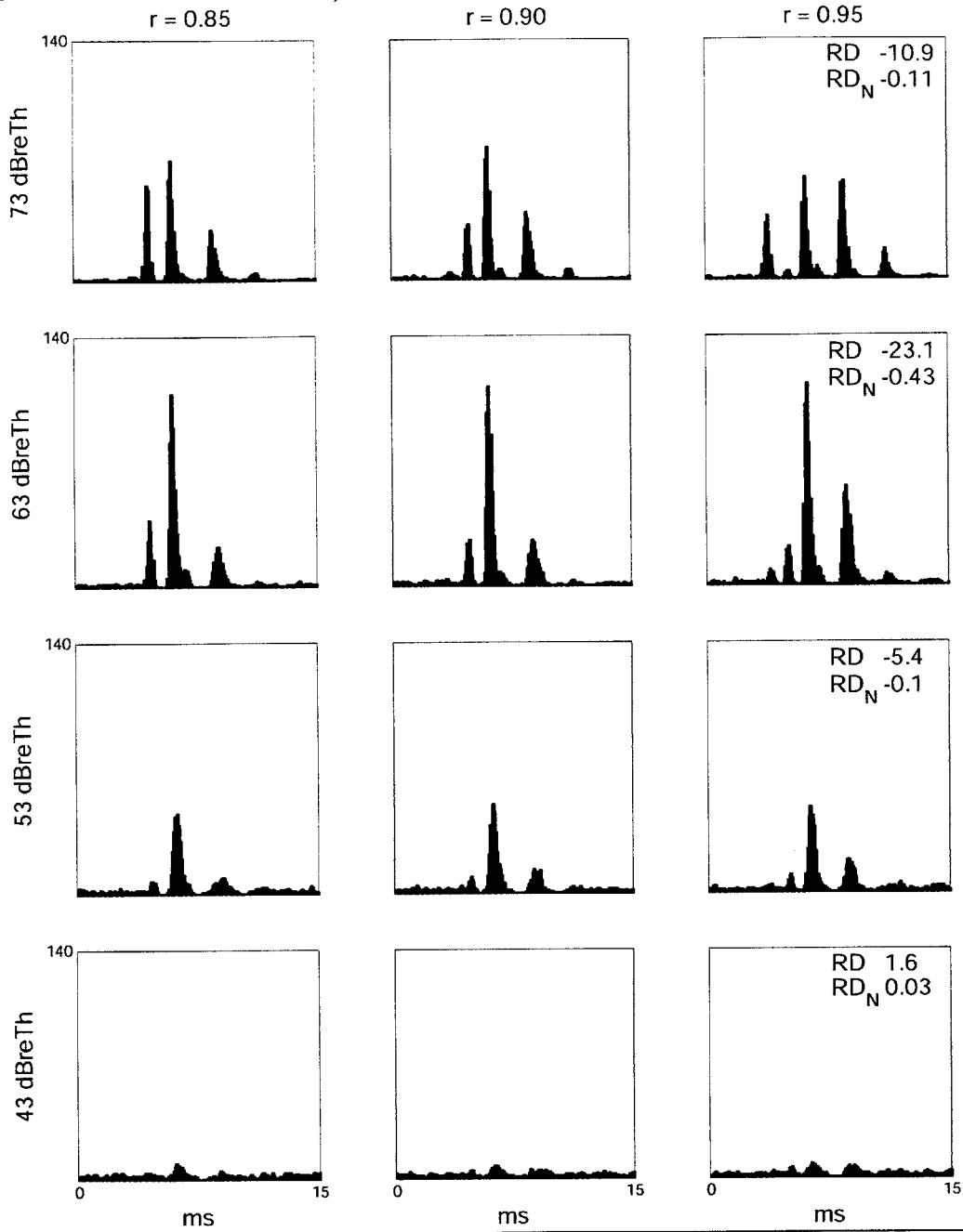
Phase Locker : CF = 504 Hz, NPS

Figure 15: Response of a NPS PL unit with CF 504 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 7A.
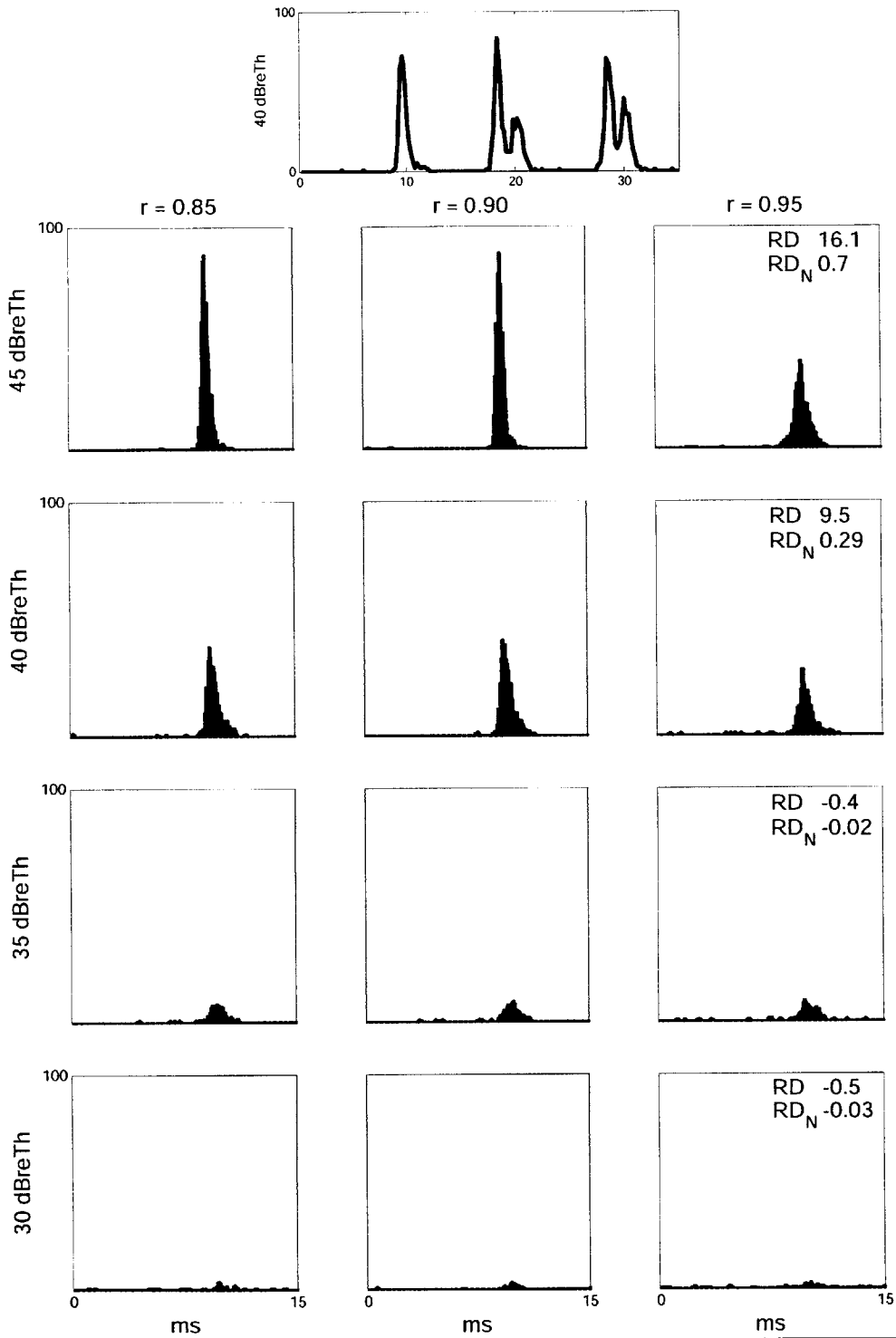
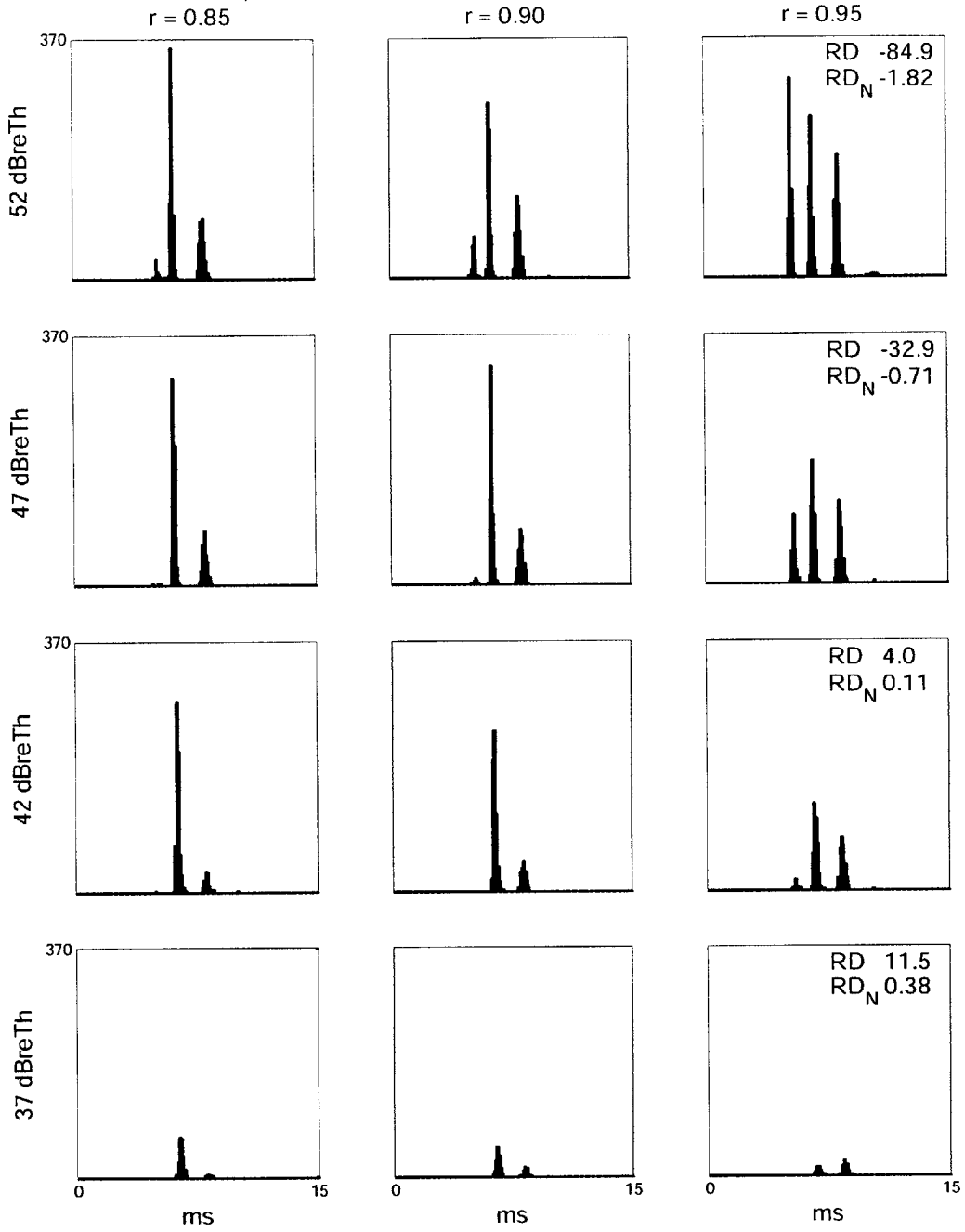**Figure 16:** Response of a PS+ PL unit with CF 100 Hz to Huffman sequences. Top panel: STBCF response.

Figure 17: Response of a PS+/- HiS unit with CF 475 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 7B.
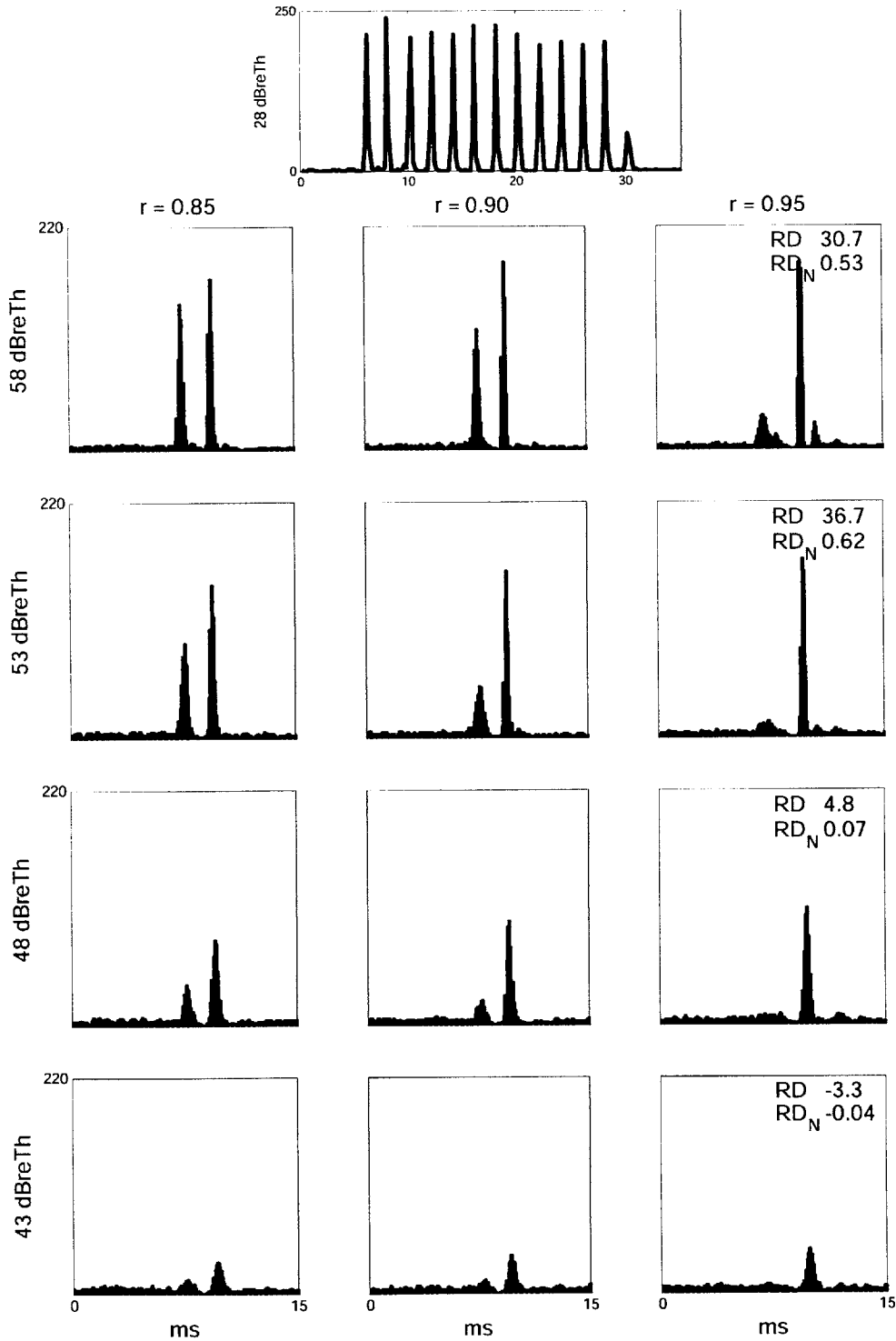
Figure 18: Response of a PS+ HiS unit with CF 501 Hz to Huffman sequences. Top panel: STBCF response.

dBreTh, the temporal response to the r = 0.95 stimulus shows the growth of a first peak that occurs before two later peaks. In comparison, the overall shape of the responses to the other two stimuli do not change as much with changes in level. The rate responses of this HiS unit are sensitive to level. Within a 15 dB range, the unit has reversed its sensitivity from preferring the broad phase transition at low levels (37 and 42 dBreTh) to the sharp phase transition at high levels (47 and 52 dBreTh), making this unit PS+/-.

Figure 18 shows results from another HiS (CF = 501 Hz, with the unit's response to a STBCF displayed in the top panel). In contrast to the other HiS example in figure 17, this unit responds to the longer duration stimulus with only one major peak rather than two of approximately equal size as observed in responses to the other two stimuli. Two recordings from this unit fell above the AN region, making it PS+. While AN fibers and all other CN types displayed decreasing RD_N with level (except for the PL unit in figure 16), this unit displayed increasing RD_N with increasing level. The behavior of the first peak of this HiS unit is similar to the single peak behavior of the PL unit in figure 16, where at high levels, the peak is much greater in response to the broad- versus the sharp-transition stimulus.

Out of 4 HiS units that we studied, 2 were PS and 2 were NPS. As with PLs, HiS units could respond to Huffman stimuli with a sensitivity in both directions that is highly dependent on level. One sensitive HiS unit was PS+, and the other was PS+/-. A third HiS unit (CF = 386 Hz, not shown) also favored the stimulus with sharp phase transition at high levels but was nevertheless NPS because the preference was too small.

Primary-like responses

Because they only receive one or two excitatory inputs from the AN, spherical bushy cells in the CN are expected to behave in a nearly one-to-one fashion with their inputs. These cells are associated with a Pri pattern in response to STBCFs. We therefore expect Pri units to respond to Huffman stimuli with patterns similar to those found in the AN.

Figure 19 shows the response patterns of a Pri unit (CF = 2500 Hz, with the unit's response to a STBCF displayed in the top panel) to two levels of Huffman stimuli. As in the higher CF AN fiber in figure 12, all of the responses have weaker peaks compared to those observed in lower CF fibers and units. The response at the highest level (50 dBreTh) has a

sharper onset peak than that observed at a lower level (40 dBreTh). At both levels, the overall shapes of the patterns in response to the three stimuli look very similar. In addition, the RD_N metric does not indicate any preference for any phase transition, making this unit NPS. These responses are similar to those of AN fibers with comparable CFs.

Figure 20 shows results from another Pri unit (CF = 1285 Hz, with the STBCF responses in figure 8B). The responses tend to have peaks placed at approximately 1 / CF intervals, though the interval between the first two peaks for the response to the r = 0.95 stimulus is slightly longer. Comparing responses at the two higher levels (51 and 61 dBreTh), the overall shape of the response patterns does not change much with level, except for the addition of a small later peak in responses to all three stimuli at 61 dBreTh. There are also slight differences between the responses to the different phase-bandwidth stimuli. At 61 dBreTh, the response to the sharp-transition stimulus has an additional later peak compared to the responses to the other two
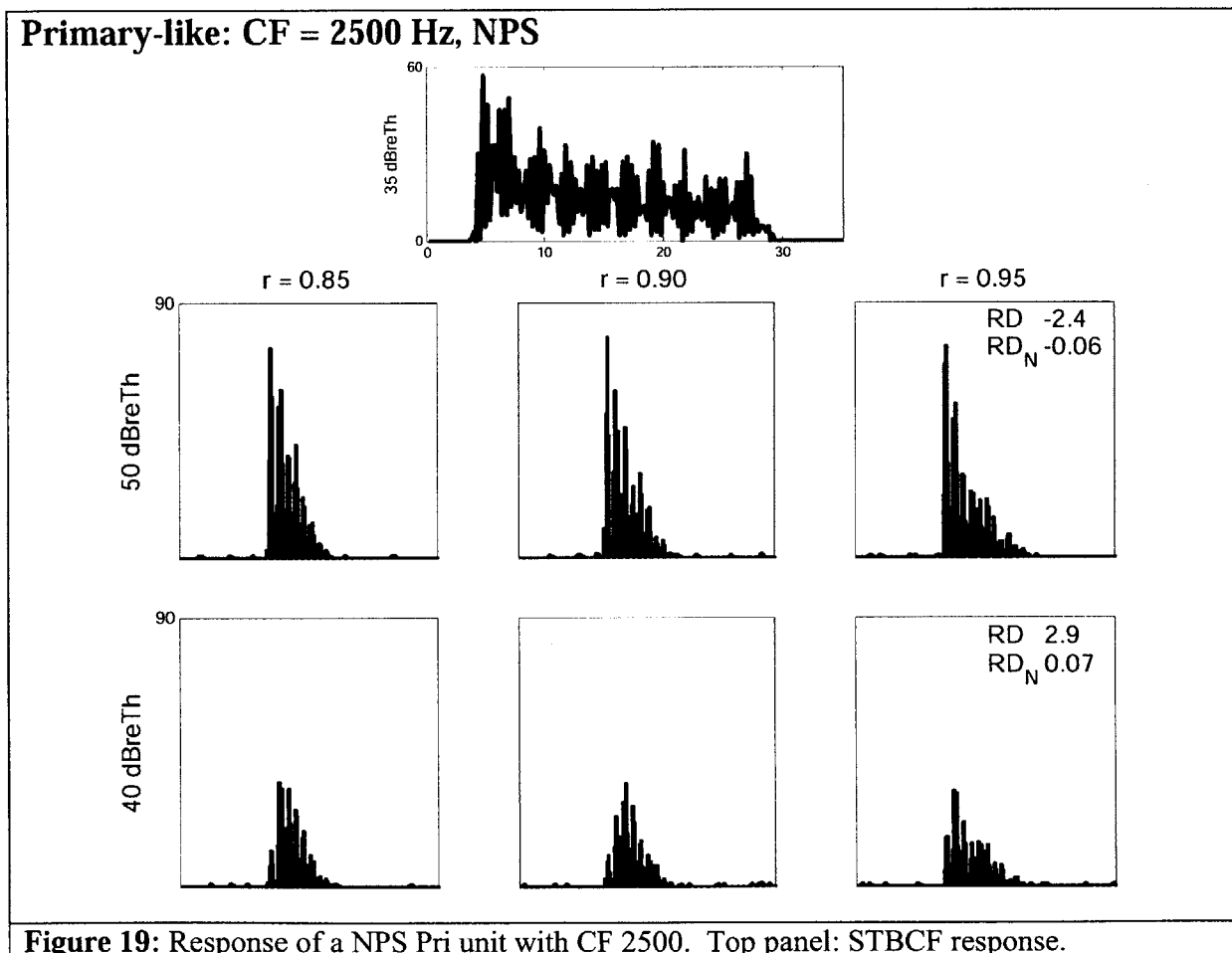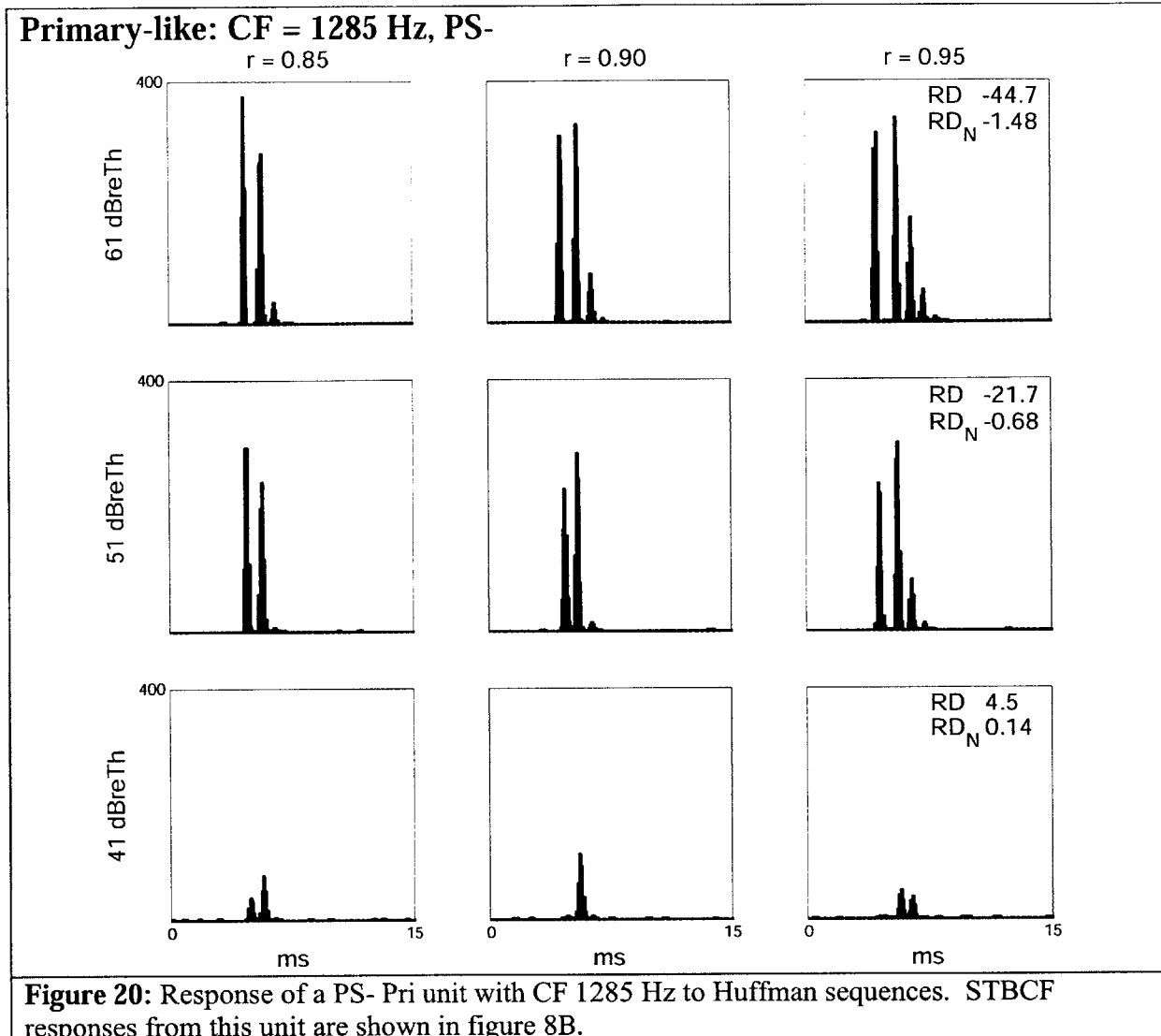

Primary-like: CF = 2500 Hz, NPS

Figure 19: Response of a NPS Pri unit with CF 2500. Top panel: STBCF response.

Figure 20: Response of a PS- Pri unit with CF 1285 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 8B.

stimuli. The rate responses indicate that this unit favors the sharp-transition stimulus at high levels and is therefore PS-. Our observations suggest that in this Pri unit, the rate difference became negative at higher levels due to the longer duration of the sharp-transition stimulus (causing a longer response duration).

Primary-like with notch responses

Globular bushy cells receive several converging inputs from the AN. These cells were of particular interest in this study because Carney (1990) reported that Pri-N units were especially

sensitive to changes in the phase spectra of Huffman stimuli compared to other unit types. She found that the direction of sensitivity was consistent with coincidence detection.

Figure 21 shows the responses of Pri-N unit (CF = 1375 Hz, with the response to a STBCF displayed in the top panel) to the three Huffman stimuli. In contrast to the Pri unit with a similar CF (figure 20), the Pri-N unit responds to the Huffman sequences with irregular spacing between peaks. This is most evident in the response to the sharp-transition stimulus at the highest level (46 dBreTh). The responses to the other two stimuli at this level are also more complex than at lower levels.

At 26 and 36 dBreTh, this unit responds with single peaks (except for a small second peak to the r = 0.95 stimulus at 36 dBreTh). The highest peak at the lower level (26 dBreTh) occurs in the response pattern to the broad-transition stimulus. 10 dB higher (36 dBreTh), the highest peak occurs in response to the sharp-transition stimulus. Furthermore, at these two levels, the highest peak is also the sharpest out of the responses to the three phase transitions, whether this occurs in the r = 0.85 response at 26 dBreTh or in the r = 0.95 response at 36 dBreTh. The lowest peak is the widest, whether this occurs in the r = 0.95 response at 26 dBreTh or in the r = 0.85 response at 36 dBreTh.

The rate responses at these levels are primarily reflective of the changes in the peak heights rather than peak widths. The positive rate differences at the two lower levels (16 and 26 BreTh) indicate that this unit slightly prefers the r = 0.85 stimulus, but this preference is not significantly different from that observed in the AN. At a level 10 dB higher (36 dBreTh), the best response of this unit changes direction, responding more to the r = 0.95 stimulus. This results in a negative RD_N, making this unit PS-.

Figure 22 shows the responses of another Pri-N unit (CF = 1610 Hz, with the response to a STBCF displayed in the top panel). As in figure 21, the temporal responses of this unit are characterized by single peaks at the lower level (20 dBreTh) and have more complex patterns at the higher level (40 dBreTh). Furthermore, the sharpest and highest peak at the lower level occurs in response to the r = 0.85 stimulus, while the widest and lowest peak occurs in response to the r = 0.95 stimulus. The rate responses indicate that this unit prefers the broad-transition stimulus at the lower level and the sharp-transition stimulus at the higher level. Sensitivity to the phase spectrum in both directions was significant, making this unit PS+/-.
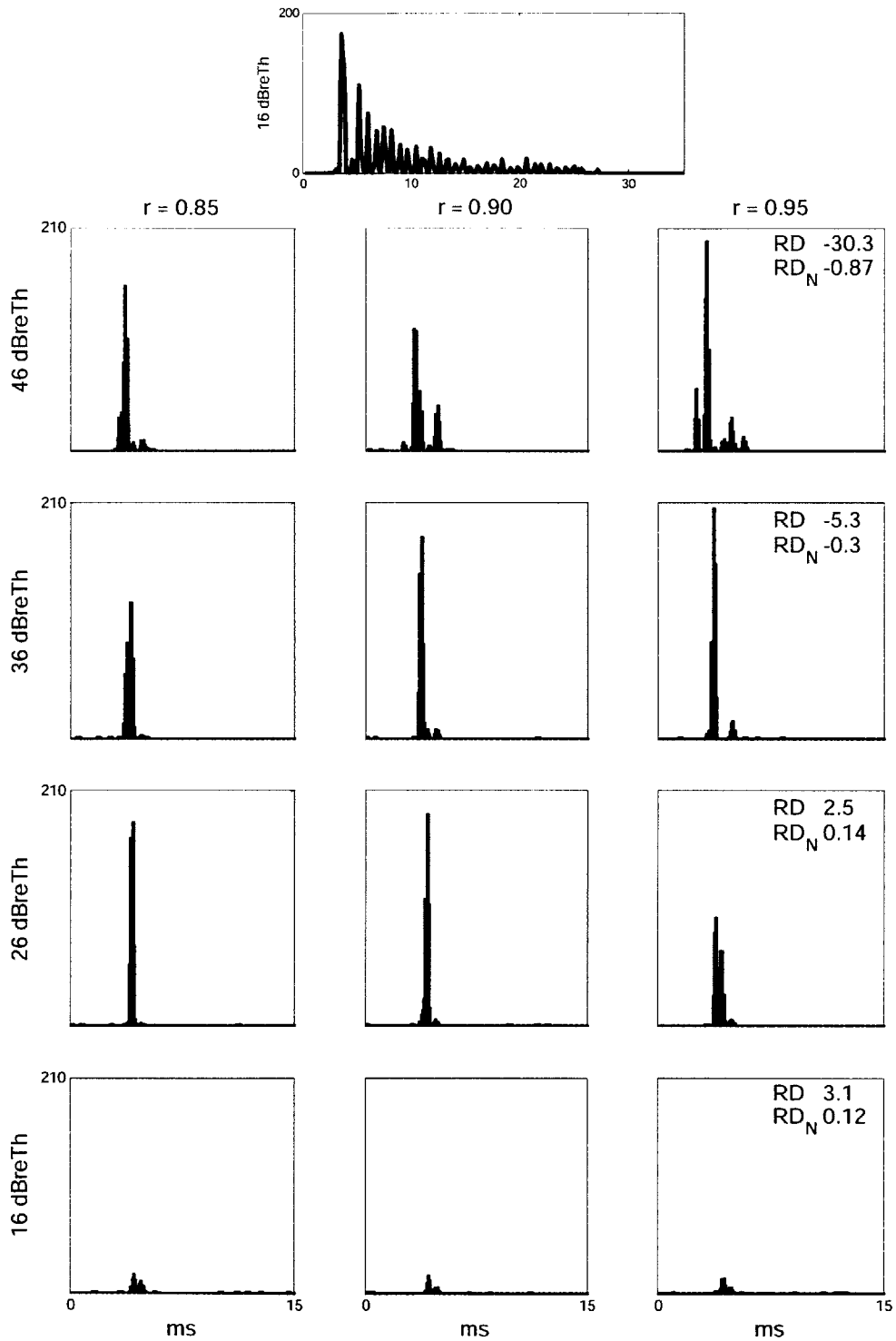
Primary-like-with-Notch: CF = 1375 Hz, PS-

r = 0.85   r = 0.90   r = 0.95

RD -30.3
RD$_N$ -0.87

RD -5.3
RD$_N$ -0.3

RD 2.5
RD$_N$ 0.14

RD 3.1
RD$_N$ 0.12

Figure 21: Response of a PS- Pri-N unit with CF 1375 Hz to Huffman sequences. Top panel: STBCF response.
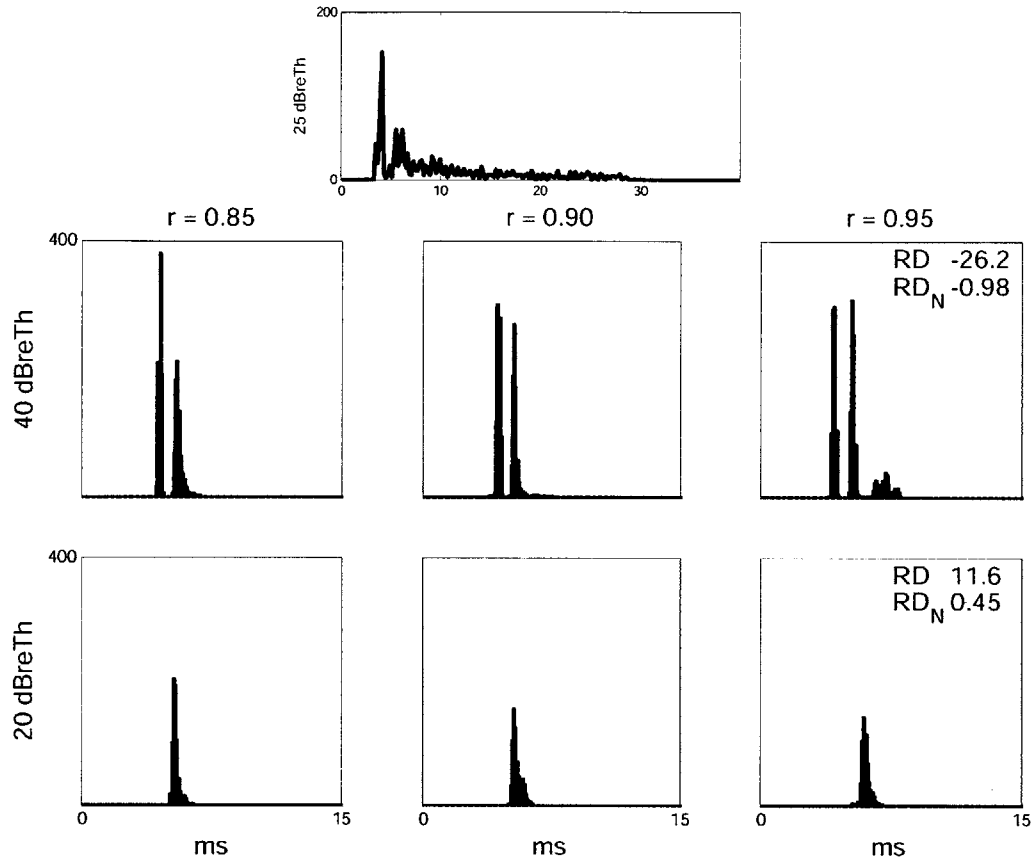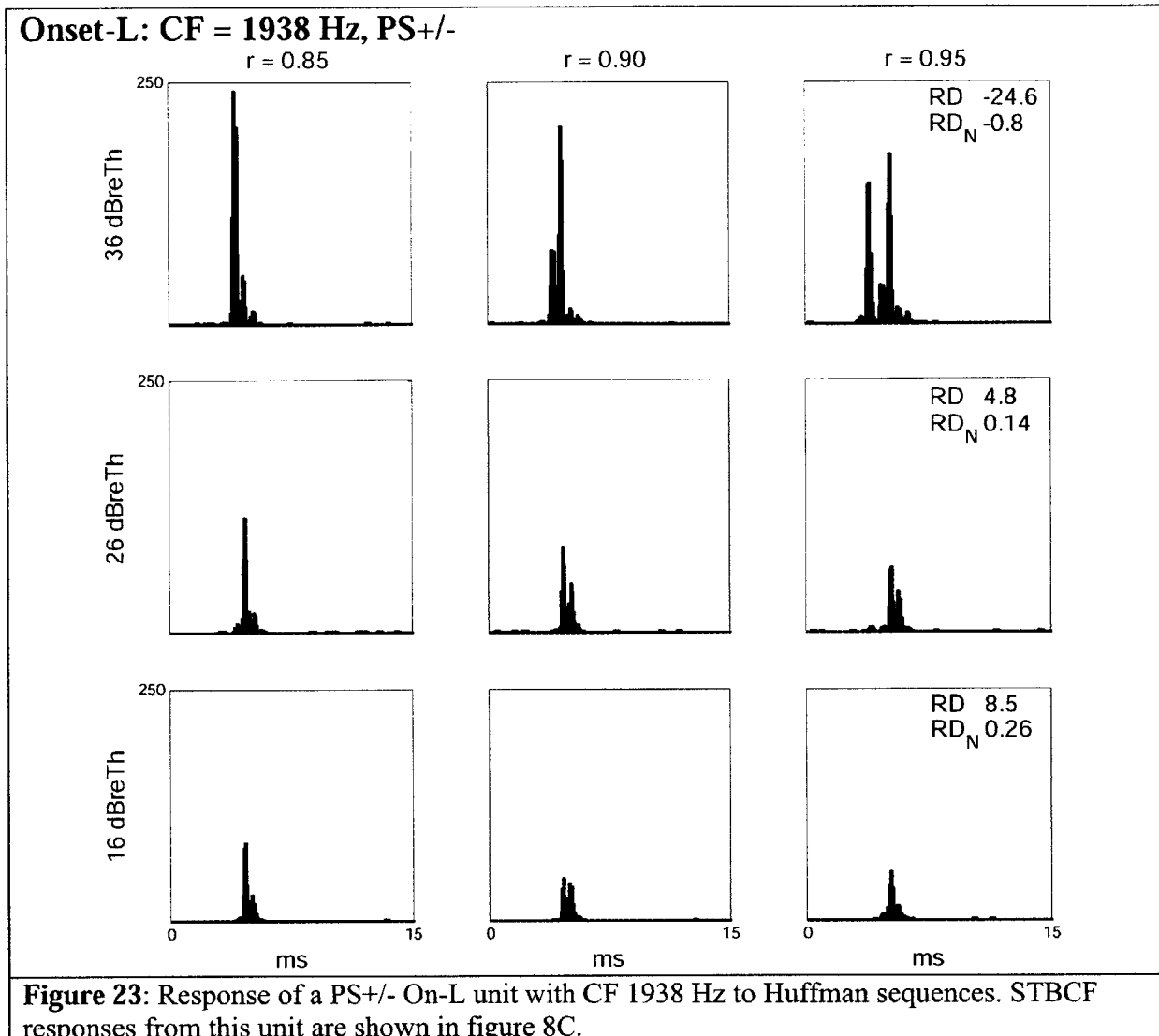
Figure 22: Response of a PS+/- Pri-N unit with CF 1610 Hz to Huffman sequences. Top panel: STBCF response.
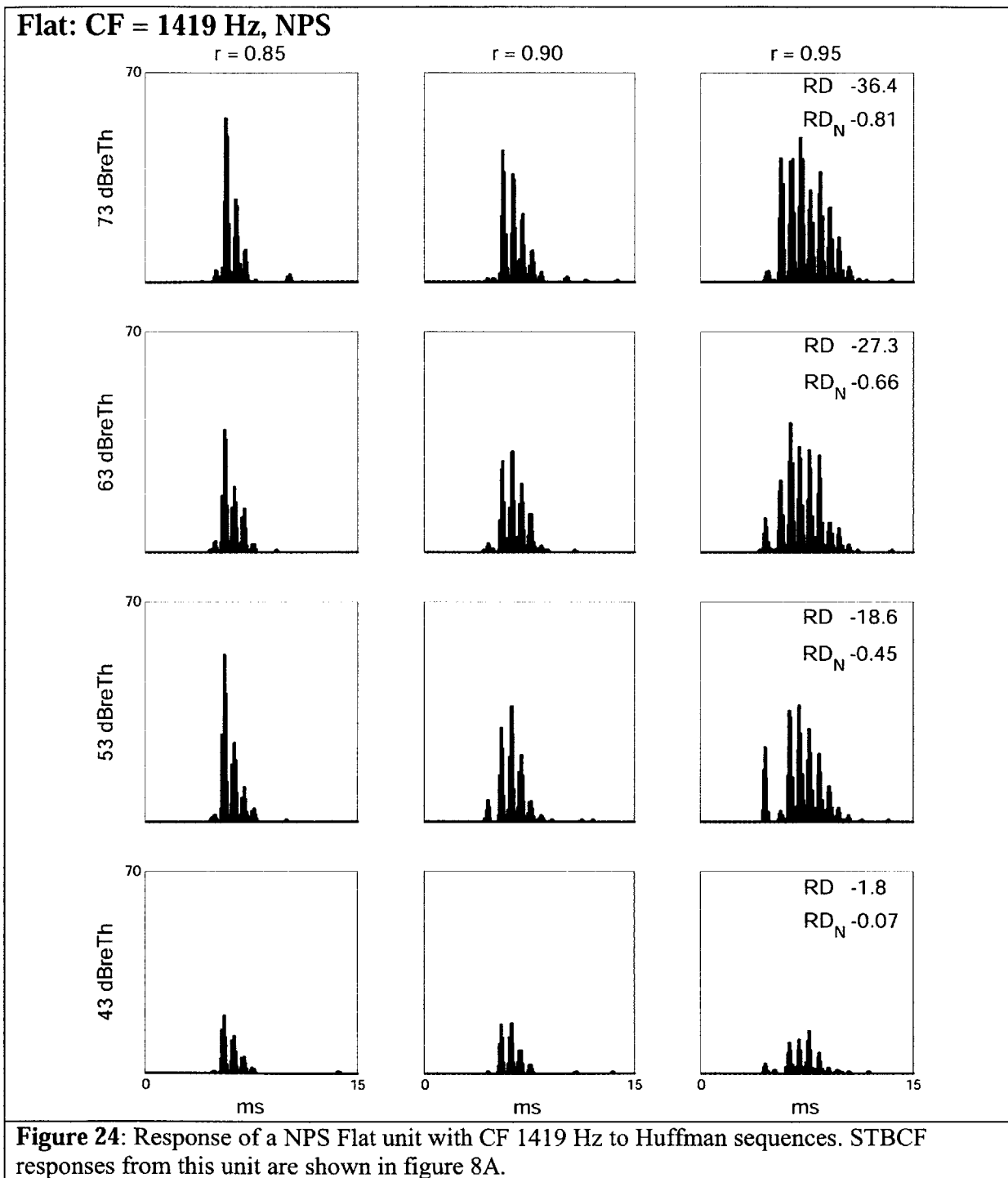
Onset responses

We studied the responses to Huffman sequences of one On-L unit (CF = 1938 Hz) at several levels. Figure 23 shows the responses of this unit (the STBCF responses are shown in figure 8C) at three levels. The strongest onset response (evaluated by considering the height of the first peaks in the PSTHs) occurs for the r = 0.85 stimulus at all levels. However, similar to what was observed in several AN fibers (see figure 11) and Pri units (see figure 20), the response to the stimulus with sharp phase transition shows a second peak at the highest level (36 dBreTh). This causes the overall rate response to have a preference for the sharp-transition stimulus at the highest level, making this unit PS+/-.

Onset-L: CF = 1938 Hz, PS+/-

r = 0.85   r = 0.90   r = 0.95

250   250   250

RD -24.6
RD$_N$ -0.8

36 dBreTh

RD 4.8
RD$_N$ 0.14

26 dBreTh

RD 8.5
RD$_N$ 0.26

16 dBreTh

0   15   0   15   0   15
ms   ms   ms

Figure 23: Response of a PS+/- On-L unit with CF 1938 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 8C.
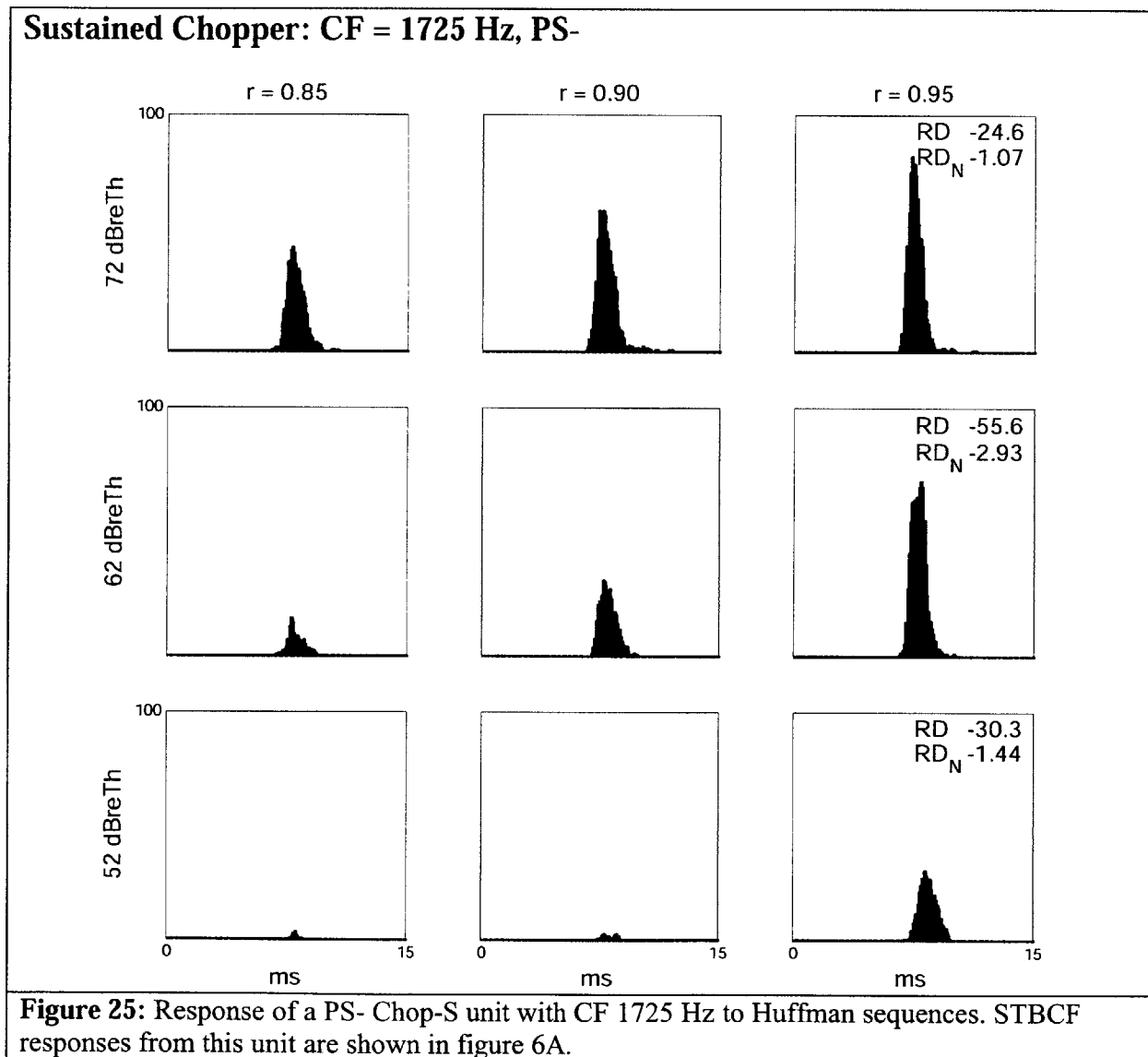
Flat responses

Figure 24 shows the responses of a Flat unit (CF = 1419 Hz, with the STBCF responses in figure 8A) to four levels of Huffman stimuli. The relative peak heights are similar between the three stimuli at all levels, but the responses to the r = 0.95 stimulus have longer duration compared to the other two stimulus responses. Responses to the sharper-transition stimuli (r = 0.90 and r = 0.95) sometimes have an additional peak that occurs before the first peak of the r = 0.85 response. This is most evident at 53 dBreTh. Furthermore, the time interval between the additional peak and the remainder of the response pattern is longer than the regularly spaced time intervals throughout the rest of the response. The RD values are negative at all levels, indicating

a slight preference for this unit to respond to the sharp-transition stimulus. However, none of the RD$_N$ values were significantly different from our AN observations, making this unit NPS.

Flat: CF = 1419 Hz, NPS

Figure 24: Response of a NPS Flat unit with CF 1419 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 8A.

Sustained Chopper responses

All four Chop-S units responded to the Huffman stimuli with a single wide peak. Figure 25 displays responses of a phase-sensitive Chop-S (CF = 1725, with the STBCF responses in figure 6A). The lack of peaks occurring at time intervals of 1 / CF in the temporal response patterns is consistent with weak phase locking in choppers. At all three levels, this Chop-S unit responds with a higher peak to the sharp- rather than broad-transition stimulus. This observation is reflected in the very negative RD_N values, making this unit PS-. In contrast to what was observed in AN fibers, the direction of sensitivity in this unit is caused by larger peak heights

Sustained Chopper: CF = 1725 Hz, PS-

Figure 25: Response of a PS- Chop-S unit with CF 1725 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 6A.

and not by a longer response to the r = 0.95 stimulus. The response patterns of this unit are similar to those recorded from a Chop-S unit by Carney (1990: figure 11a).

Two of the Chop-S units were sensitive, and both preferred the Huffman stimulus with sharp phase transition, making them PS-. The temporal and rate responses of the other two units were insensitive to changes in the phase spectrum of the stimulus (NPS).

Transient Chopper responses

Figure 26 shows the responses of a Chop-T unit (CF = 2197 Hz, with the STBCF responses in figure 6B). At the lowest level (50 dBreTh), the responses have single wide peaks in response to all three Huffman stimuli (similar to what was observed in the Chop-S unit in

Transient Chopper: CF = 2197 Hz, NPS

r = 0.85   r = 0.90   r = 0.95

RD   -60.2
RD$_N$ -1.23

RD   -57.6
RD$_N$ -1.3

RD   -4.3
RD$_N$ -0.14

Figure 26: Response of a sensitive Chop-T unit with CF 2197 Hz to Huffman sequences. STBCF responses from this unit are shown in figure 6B.
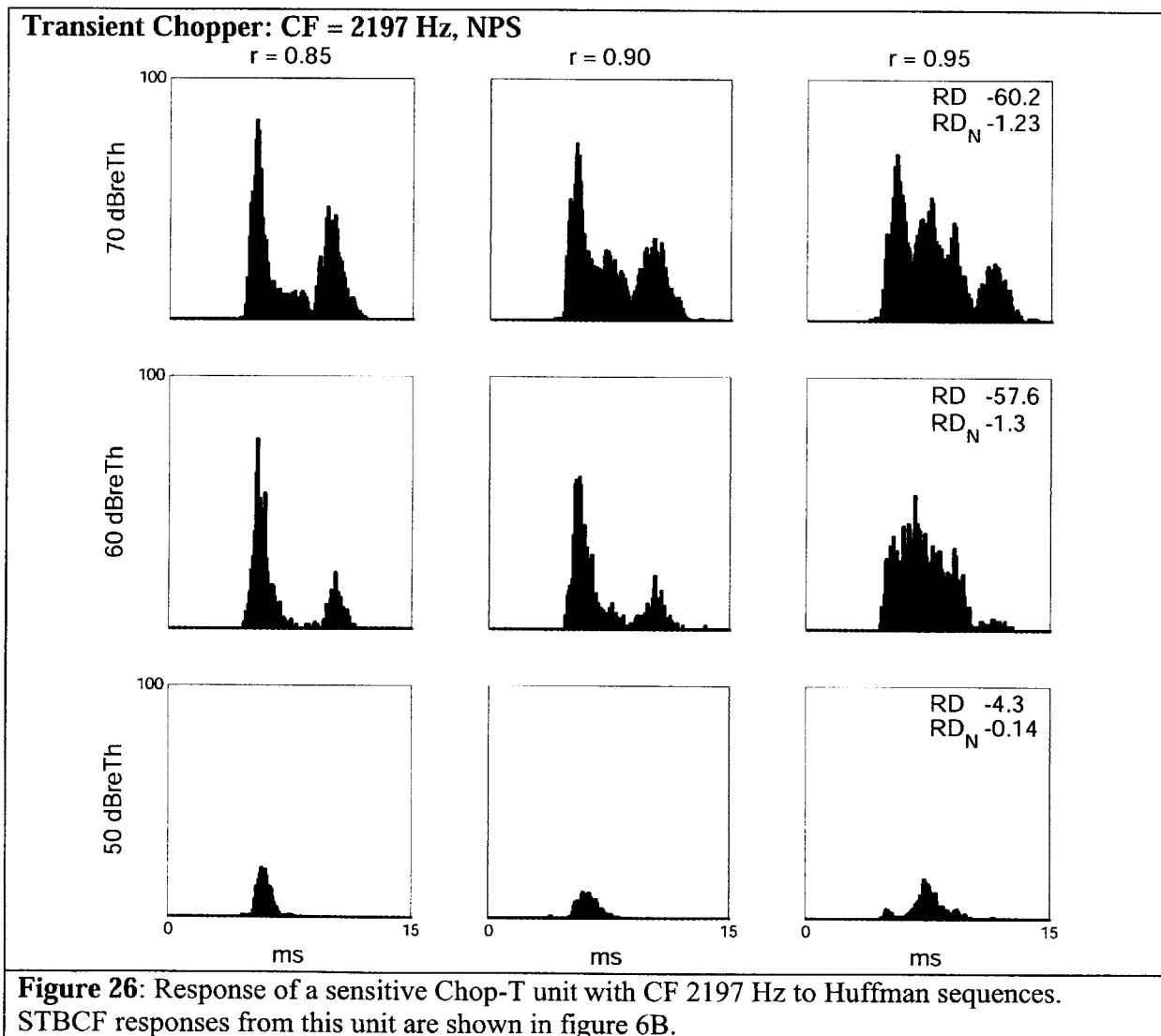
49

figure 25). At higher levels (60 and 70 dBreTh), the responses of this unit became multimodal and longer in duration. The rate responses are negative at all three levels, suggesting a preference for the sharp-transition stimulus. However, this preference is not significantly different from that observed in AN fibers, making this unit NPS. The temporal response patterns of this unit are similar to the Chop-T unit responses shown by Carney (1990: figure 11b).

Summary of rate responses in the CN

Figure 27 compares the RD_N of CN units (colored dots) with that of AN fibers (gray dots) as a function of level (dBreTh). Data from different unit types are plotted in separate colors and on different panels. Data collected from CN units that were sensitive to the phase transition bandwidth of the stimuli are highlighted with open colored circles. Colored lines connect data points corresponding to the same units, and the example CN units shown in figures 15-26 are labeled with "F" followed by the figure number.

Figure 27A shows PL data in blue and HiS data in red. Out of 10 PL units studied, 2 were sensitive to changes in the phase transition bandwidth, with one PS+ and the other PS-. At low levels, neither sensitive PL units favored any particular stimulus, but as level increased, one favored the stimulus with broad phase transition (PS+, F16) while the other favored the sharp phase transition (PS-).

Out of 4 HiS units studied, 2 were PS. As in PL units, HiS units could respond to Huffman stimuli with sensitivity in both directions that was highly dependent on level. One sensitive HiS unit is PS+ (F18) while the other is PS+/- (F17). A third HiS unit slightly favored the sharp-transition stimulus but was not sensitive. Figure 27A shows that recordings from all other PL and HiS units fell within the closed region defined by 95% of the AN data (except for one PL recording at 20 dBreTh).

Figure 27B summarizes data recorded from Pris (in blue) and Pri-Ns (in red) and compares these units with AN fibers. Out of 28 Pri units, 7 were sensitive to the phase transition bandwidth of the Huffman stimuli. All of the sensitive Pri units displayed more negative RD_{NS} than AN fibers at comparable levels (6 were PS-, and 1 was PS+/-). At lower levels (<40 dBreTh), all of the Pri unit recordings were within the insensitive region. At higher levels (>40 dBreTh), nearly all of the recordings (sensitive and insensitive) from Pri units were shifted
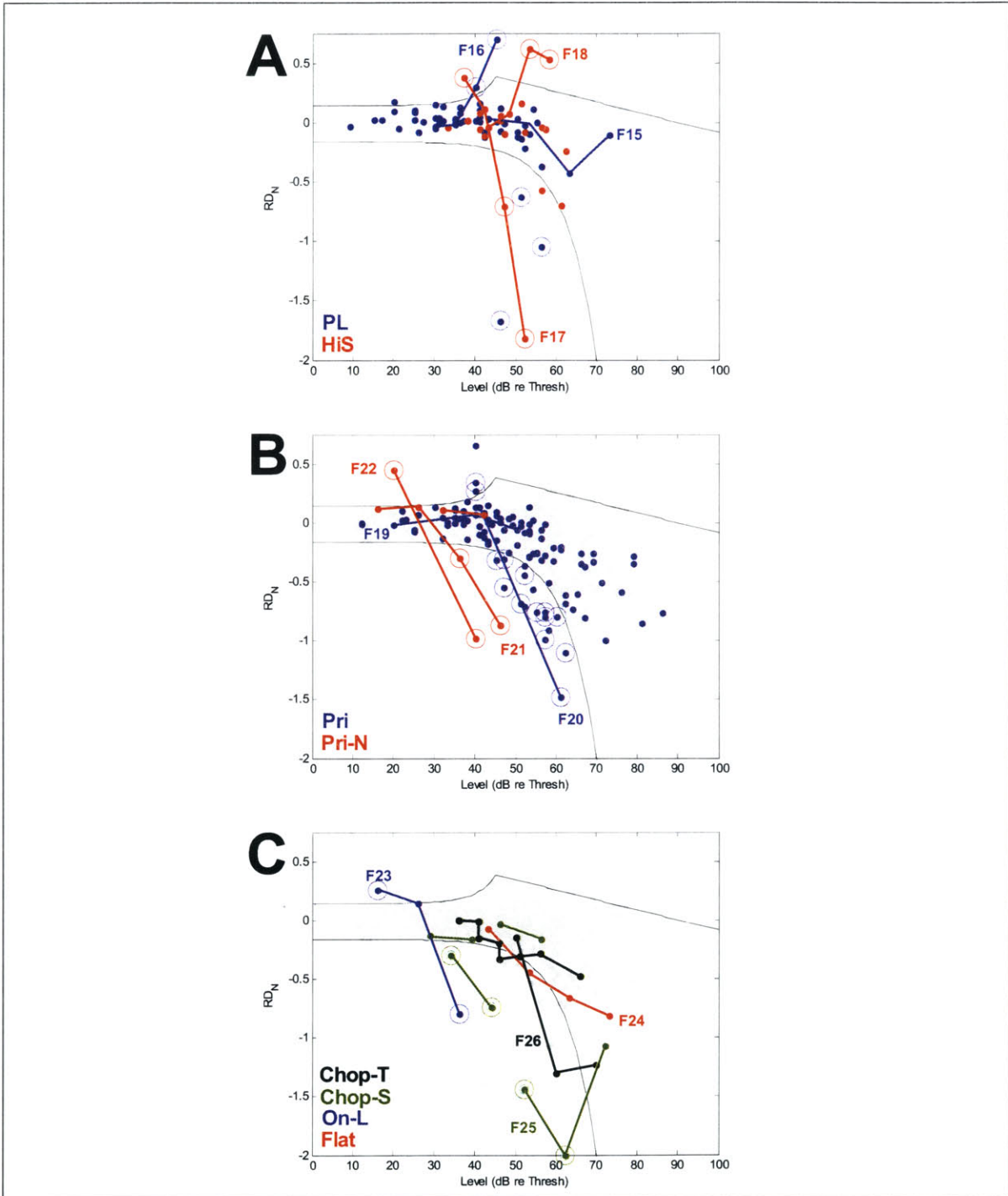
Figure 27: CN population results. **A.** PL and HiS, **B.** Pri and Pri-N, **C.** Chop-T, Chop-S, On-L, and Flat. AN data are plotted in gray dots. CN data are plotted in colored dots. Open colored circles indicate CN recordings that fall outside the closed region defined by 95% of the AN data. Colored lines connect between data points corresponding to the same units. Data recorded from the example units in figures 15-26 are labeled.

towards more negative RD_{NS}, indicating Pris better respond to the sharp-transition stimulus than AN fibers.

Two out of 3 Pri-Ns were sensitive. One Pri-N was PS- (F21), and another was PS+/- (F22). Figure 27 shows that the sensitive Pri-N units respond with RD_N values that are significantly more negative than observed in AN, PL, HiS, and Pri units at comparable levels.

Figure 27C shows data from 2 Chop-T units, 4 Chop-S units, 1 On-L, and 1 Flat. Neither Chop-T unit (black) was sensitive to the phase changes in the stimuli, though one unit (F26) slightly favored the sharp-transition stimulus at high levels. 2 Chop-S units (green) were sensitive, and both preferred the sharp-transition stimulus (PS-). The On-L (blue) whose responses are shown in figure 23 was PS+/-. Recordings from this unit resulted in similar RD_N values at comparable levels as the PS+/- Pri-N (15-50 dBreTh) in figure 27B (F22). The Flat unit (red) had one recording that barely fell outside the AN region and was NPS.

Table II summarizes the total number of CN units studied, the number of sensitive units, and the direction of their sensitivity. For comparison, AN data are shown in italics in the first row of the table. 5 out of 84 (6%) AN fibers were "sensitive" to the phase transition bandwidth of the stimuli. The distribution of the direction of "sensitivity" of AN fibers was fairly uniform across our defined categories, as expected from how we drew the boundaries of the NPS region.

The table shows that 10 out of 16 sensitive CN units were PS-, meaning a better response to the sharp-transition stimulus than AN fibers at the same levels. 2 sensitive units (one PL and

Table II

	Number Studied	Number PS+	Number PS-	Number PS+/-	Number Sensitive
AN	*84*	*1 (1%)*	*2 (2%)*	*1 (1%)*	*5 (6%)**
PL	10	1 (10%)	1 (10%)	0 (0%)	2 (20%)
HiS	4	1 (25%)	0 (0%)	1 (25%)	2 (50%)
Pri	28	0 (0%)	6 (21%)	1 (4%)	7 (25%)
Pri-N	3	0 (0%)	1 (33%)	1 (33%)	2 (67%)
Chop-T	2	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Chop-S	4	0 (0%)	2 (50%)	0 (0%)	2 (50%)
On	1	0 (0%)	0 (0%)	1 (100%)	1 (100%)
Flat	1	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Total CN	53	2 (4%)	10 (19%)	4 (8%)	16 (30%)

* 1 AN fiber was sensitive but did not fit any of the three categories.

one HiS) were PS+, which responded better to the broad-transition stimulus. The remaining 4 sensitive units were sensitive in both directions (PS+/-). These units preferred the broad-transition stimulus at low levels and the other at high levels.

Temporal response patterns

We observed that the responses of CN units to Huffman stimuli often have shorter duration compared to the responses of AN fibers with similar CFs and at comparable levels. To quantify this, the duration of each Huffman response was measured as described in the methods and normalized by the CF period. Figures 28A-B compare the average duration data between two CN unit types (black) and AN fibers (pink) as a function of level (in dBreTh). Each column represents a different phase transition stimulus. The lines and error bars are the average and standard deviation of the response duration for 30 AN and 16 Pri recordings per bin in figure
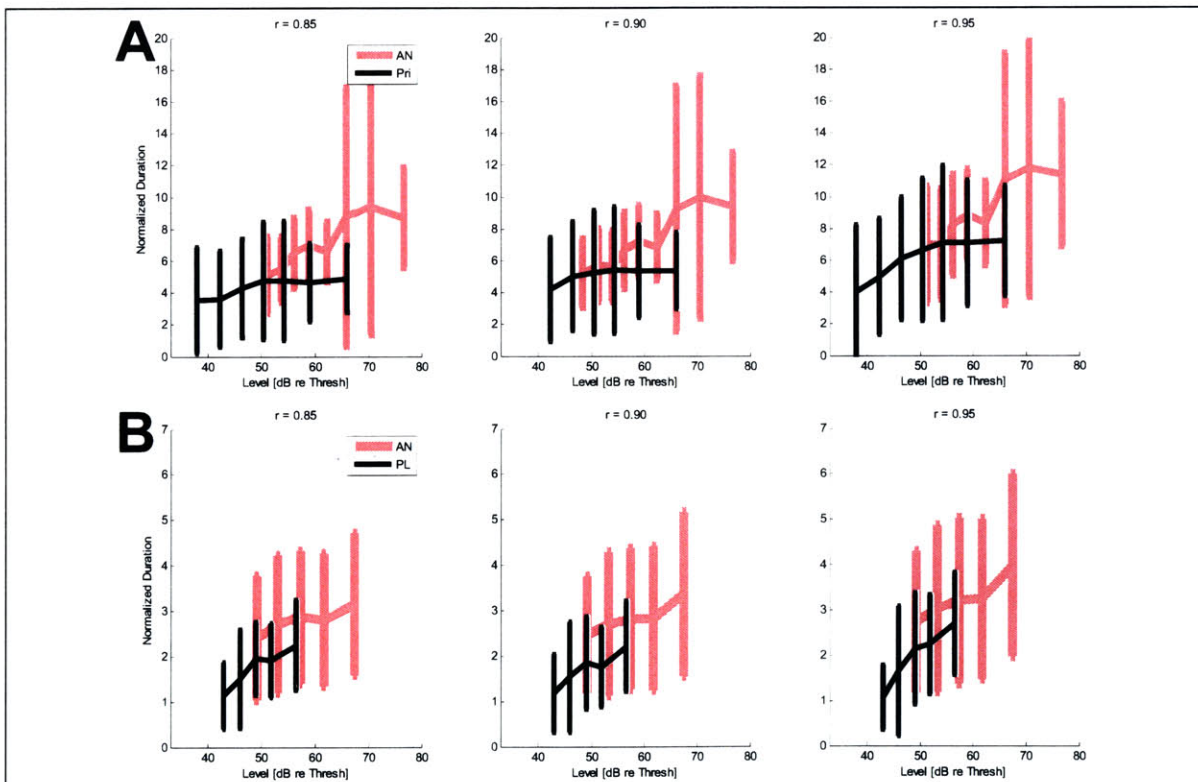

Figure 28: Duration versus level (dBreTh) of unit responses to Huffman stimuli. Each line corresponds to the average (middle of each line) and standard deviation (half length of each line) of binned data sets. Bins overlapped by 50%. **A.** Pri (black) and AN (pink) data. Bin widths are 16 (Pri) and 30 (AN) data points. CF was restricted between 1 and 4 kHz. **B.** PL (black) and AN (pink) data. Bin widths are 10 (PL) and 30 (AN) data points. CF was restricted to be less than 1 kHz.

28A and 30 AN and 10 PL recordings per bin in figure 28B. In figure 28A, the CF range was limited to between 1 and 4 kHz, while the CF was limited to less than 1 kHz in figure 28B. The x-coordinate corresponds to the mean of the levels (in dBreTh) of the recordings included in the bin. Adjacent bins overlapped with one another by 50%.

In response to all three stimuli and in both CF ranges, AN fibers respond with longer duration as level increases. At high levels, AN fibers also respond with longer durations to the longer duration stimulus (r = 0.95) compared to the r = 0.85 stimulus.

At lower levels (<55 dBreTh), Pri units respond to all three Huffman stimuli with durations similar to those measured in the AN. At higher levels (>55 dBreTh), Pri units have shorter durations than AN fibers. Furthermore, as level increases beyond 55 dBreTh, the Pris do not continue to respond with longer durations. This is most clear in response to the sharp-transition stimulus. These observations suggest that Pris have a sharpened temporal response compared to that of their inputs and do not necessarily respond in a one-to-one fashion.

Figure 28B shows the duration data for PL units and AN fibers with CFs below 1 kHz. In contrast to the Pri data, PLs respond with longer durations as level increases. Furthermore, at all levels and for all three stimuli, the temporal responses of PLs are slightly shorter in duration than in AN fibers Both sets of data have durations that increase with level at approximately the same rate. This suggests that the responses of low frequency CN units to transient stimuli are temporally sharpened compared to those of their inputs, and that this sharpening may not depend on level.

Harmonic complex tones

We recorded 178 responses to complex tone stimuli from 64 CN units (with CFs 100 – 4060 Hz) in 10 cats. These included 25 Pri, 5 Chop-S, 8 Chop-T, 2 Ons, 2 Pri-N, 5 Flat, 5 LR, 10 PL, and 2 HiS units. For comparison, we used the AN data of Cedolin (2006) obtained in response to the same stimuli. These data consisted of 173 responses from 94 AN fibers in 6 cats.

Primary-like responses

Our findings for Pri units in response to complex tone stimuli are similar to those in the AN. The three panels in figure 29 show response patterns of three Pri units with different CFs. Responses are displayed as period histograms as a function of normalized time (time * F0, or
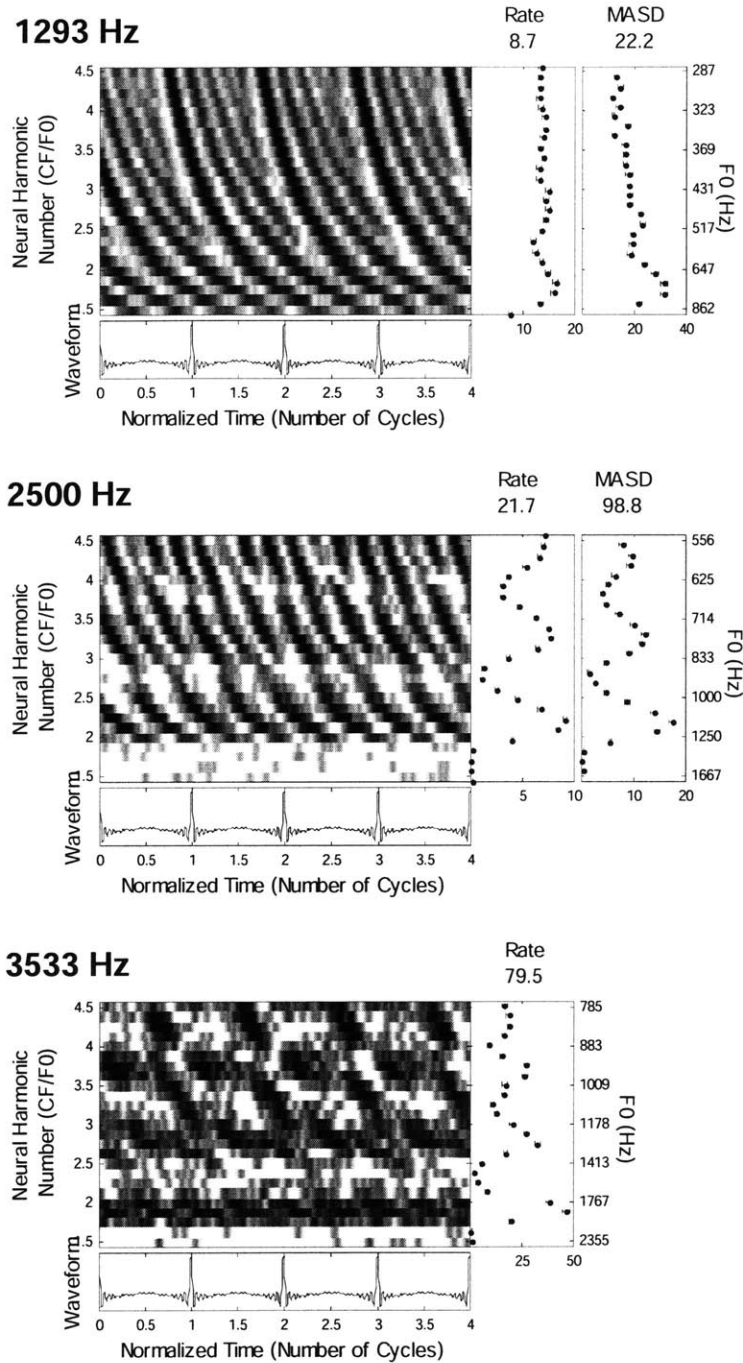
Figure 29: Responses of three Pri units to harmonic complex tone stimuli. The Pri unit with highest CF (3533 Hz) does not phase lock consistently in response to each tone in the stimulus, so the MASD profile was not plotted for this unit.

cycles) and neural harmonic number (CF/F0). The stimulus waveforms are plotted against the normalized time axis in the bottom of each panel. The rate and MASD profiles (except for the response of the Pri with CF 3533, for which only the rate profile is shown) are also plotted as a function of harmonic number (or F0 in Hz) and are shown to the right. Harmonics were said to be resolved if the rate or MASD profiles displayed clear peaks at integer neural harmonic numbers (when one harmonic coincides with the CF). The "harmonic strength" (HS) is a measure of the amplitude of the oscillation in the rate and MASD profiles and is derived as described in Methods. The HS is expressed as a signal-to-noise ratio and is displayed above each profile.

The top panel of figure 29 shows the response of a Pri unit with CF 1293 Hz. The response latencies of this unit only varied rapidly when the CF was twice the F0. The resulting rate and MASD profiles do not display strong cues to resolved harmonics above the second. This is similar to what is seen in the AN and consistent with relatively poor frequency selectivity at low CFs.

The middle panel shows results from a Pri unit with higher CF (2500 Hz), which has better harmonic resolvability than the lower CF Pri (as seen in the strong oscillations in rate response). In addition, the response latencies vary rapidly at integer harmonic numbers (resulting in pronounced peaks in the MASD profile) and vary less in between (resulting in valleys in the MASD). At this CF, the MASD cue is stronger than the rate cue as shown by the harmonic strength metric.

The bottom panel shows the response of a third Pri with CF 3533 Hz. In this unit, frequency selectivity has further improved, resulting in a strong oscillation in the rate profile and a high HS. On the other hand, because phase locking has begun to break down at this high frequency and only seems to occur in response to the stimuli with lower F0s, the MASD profile could not be estimated reliably and is not shown. These trends with rate and MASD representations in units with different CFs are consistent with AN observations (Cedolin 2006).

We recorded responses of CN units to complex tone stimuli at different levels to test for an extraction or an enhancement of the spatio-temporal cues. Figure 30 shows the responses of a Pri unit (CF = 2300 Hz) to complex tone stimuli at two different levels. At the lower level (25 dBreTh), both the rate and MASD profiles have strong oscillations, resulting in high HS values displayed above each profile. However at 45 dBreTh, the firing rate of this unit has saturated,
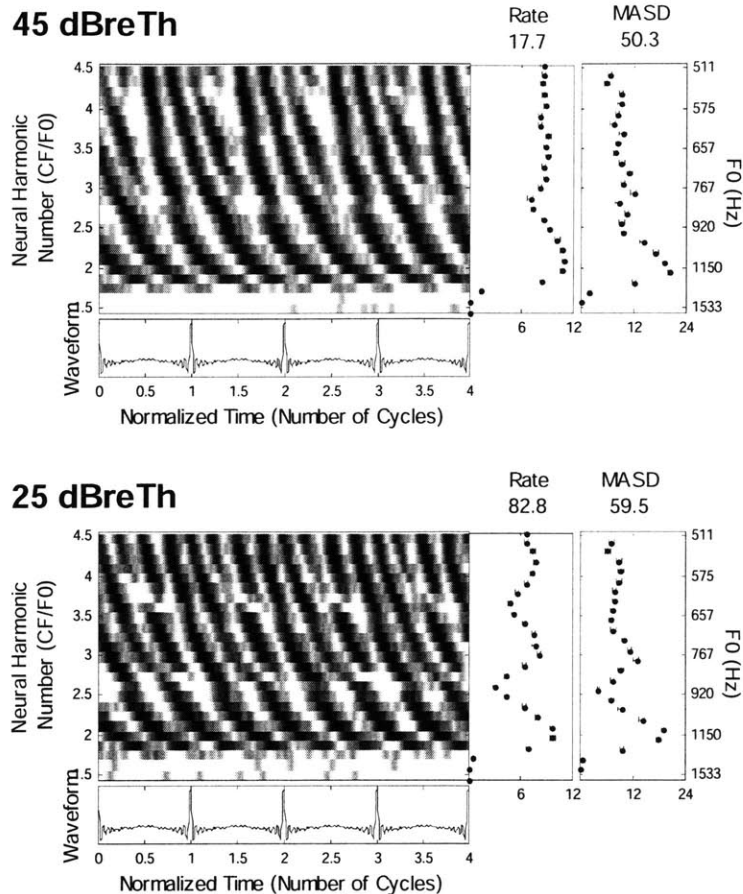
Figure 30: Responses of a Pri unit with CF 2300 Hz to harmonic complex tone stimuli at two levels.

resulting in less pronounced oscillations and a degraded rate cue. The MASD cue remains relatively robust at this higher level, though its profile has much weaker oscillations in response to the lower F0 tones than at a lower level. Thus the MASD representation in Pri units is more robust with increases in level than the rate representation. The same trends have also been shown in the AN (Cedolin 2006).

Phase locker and Hi-Sync responses

Low CF (<1000 Hz) CN units were typically either PL or HiS. Many of these responded to complex tone stimuli with phase locking patterns similar to those found in low CF AN fibers. Frequency selectivity at these low frequencies is poorer than at higher frequencies, and the

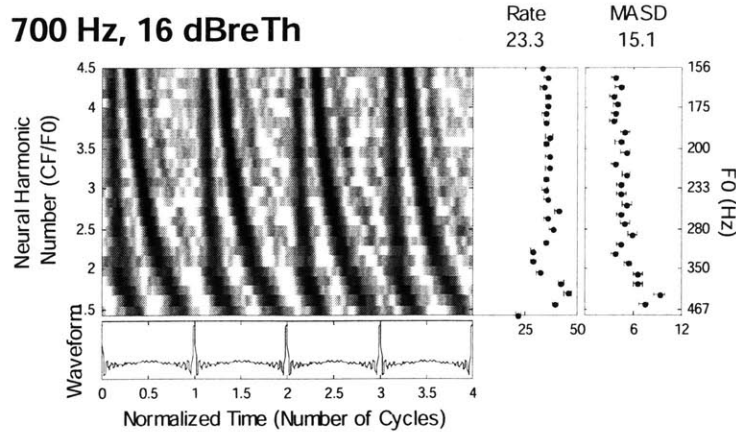**Phase Locker**

700 Hz, 16 dBreTh

Rate 23.3   MASD 15.1

Figure 31: Response of a PL unit with CF 700 Hz to harmonic complex tone stimuli.

response latencies of these low CF units tend to vary uniformly with neural harmonic number. As a result, neither the rate nor the MASD profiles show strong cues to resolved harmonics. Figure 31 shows the response of a typical phase-locker (CF = 700). The shape of the period histograms look similar to those found in AN fibers with similar CF. There were no obvious differences between the response patterns of PL and HiS units.
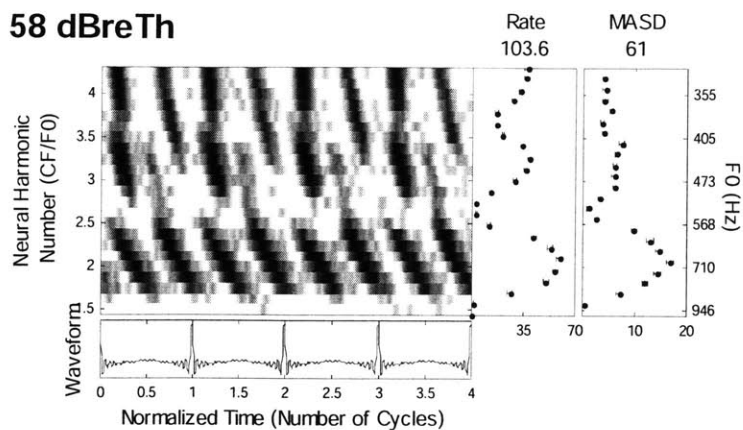
Flat responses

Figure 32 shows the responses of a Flat unit (CF = 1419 Hz) to three levels of harmonic complex tone stimuli. This is the same unit for which we showed response to STBCFs in figure 8A and to Huffman stimuli in figure 24. Contrary to what has been observed in the AN and other CN units, the rate cues of this unit improve with level, while the MASD does not change much with level. This was the only unit in our sample that displayed a significant increase in the rate strength with level. The other 4 Flat units (not shown) responded to complex tone stimuli in a fashion similar to Pris and AN fibers.
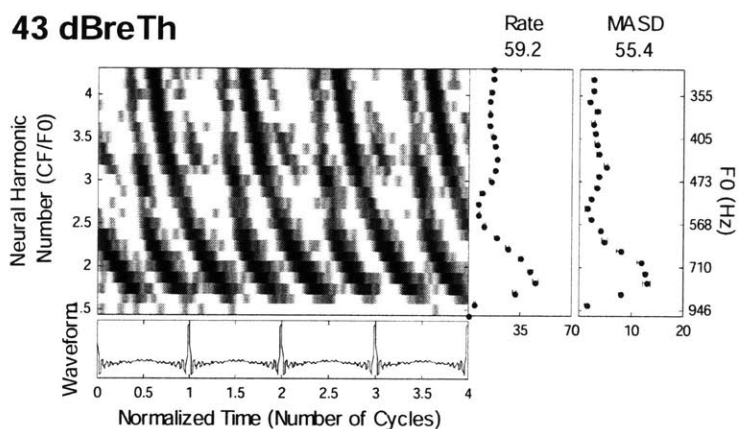
Chopper responses

Chopper units in the CN weakly phase lock to periodic stimuli over 1 kHz. None of the chopper units responded to the complex tone stimuli with strong phase locking, except in a few cases for complex tones with low F0s. As a result, the MASD profiles of most chopper
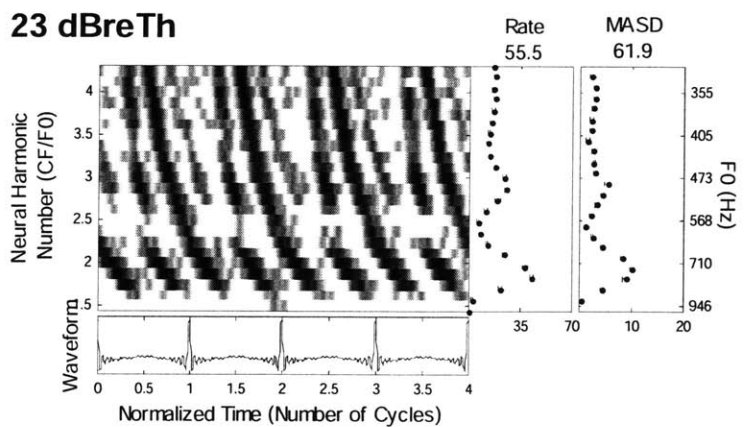
Figure 32: Responses of a Flat unit with CF 1419 Hz to harmonic complex tones. This unit's responses to STBCFs are shown in figure 8A and to Huffman stimuli in figure 24.

recordings are not indicative of the variation of response latencies, and only the rate profiles are shown.

Figure 33 shows the response patterns of a Chop-S unit (CF = 4060 Hz) to harmonic complex tone stimuli at three levels. At the lowest level (22 dBreTh), this unit responds preferentially at integer neural harmonic numbers, resulting in strong oscillations in the rate profile and a high HS (103.9). As level increases, the unit responds to complex tones with all F0 values, resulting in weaker oscillations in the rate profile and a degradation of the rate cue.

Figure 34 shows results from a Chop-T unit (CF = 1149 Hz) in response to complex tone stimuli at three levels. This unit has good harmonic resolvability at the two lower levels (0 and 9 dBreTh). However, at the highest level (19 dBreTh), the unit's rate profile has less pronounced oscillations due to a saturation in the firing rate, and results in a poorer rate cue.

Summary data for complex tones

The six panels in figures 35A-B compare CN and AN population results for the HS associated with the AN mean rate (light blue), AN MASD (pink), and CN mean rate (various colors) representations. HS values are plotted against CF (on the x-axes) and level (in dBreTh, on the three panels in each figure). CN data are plotted separately for various unit types. Pri (blue) and Pri-N (red) units are shown in figure A, and Chop-S (green), Chop-T (black), On (blue) and Flat (red) units are shown in figure B. Data recorded from units that were sensitive to the phase transition bandwidth of Huffman stimuli are marked with open circles.

For medium and high levels, low CF AN fibers (< 2 kHz) typically have stronger MASD than rate cues. As CF increases for all three level ranges, the rate cue becomes stronger than the MASD. Furthermore, both cues decrease with level increases, but the MASD is relatively more robust compared to the rate representation.

There is considerable overlap between the AN and CN data. However, in figure 35A, a few low CF (1-3 kHz) Pris (several of which are phase sensitive to the Huffman stimuli) have stronger harmonic oscillations in their rate profiles compared to those of AN fibers. In fact, Pris with these CFs have rate cues similar in value to the MASD cues in the AN and sometimes display rate cues even stronger than AN MASD cues. The triangles at the top of both figures indicate data for which harmonic strength exceeded 150. This only occurred in several high CF

Sustained Chopper: CF = 4060 Hz

62 dBreTh
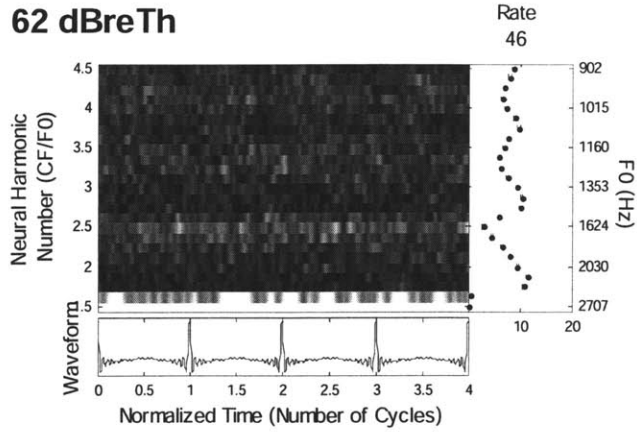
42 dBreTh
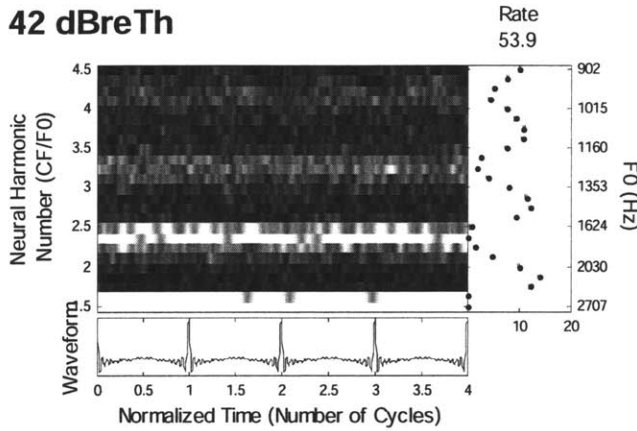
22 dBreTh

Figure 33: Responses of a Chop-S unit with CF 4060 Hz to harmonic complex tone stimuli. The MASD profiles are not plotted for this unit because there is no apparent phase locking in the response patterns.
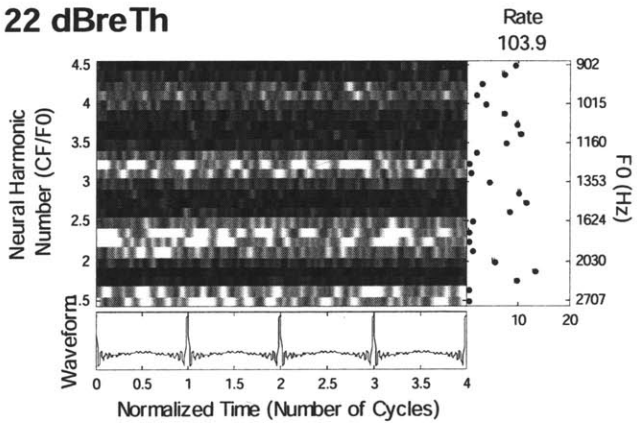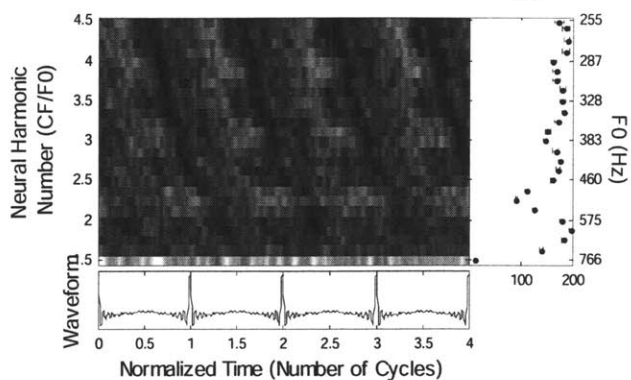
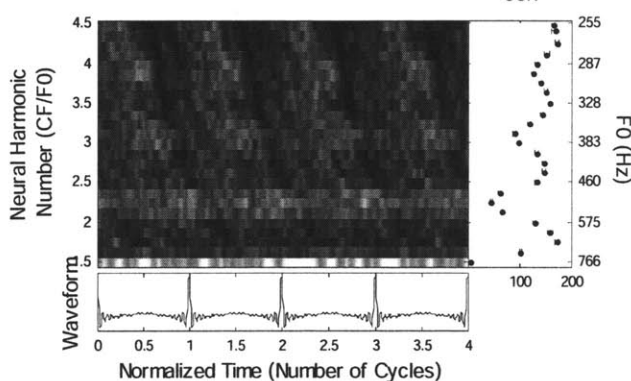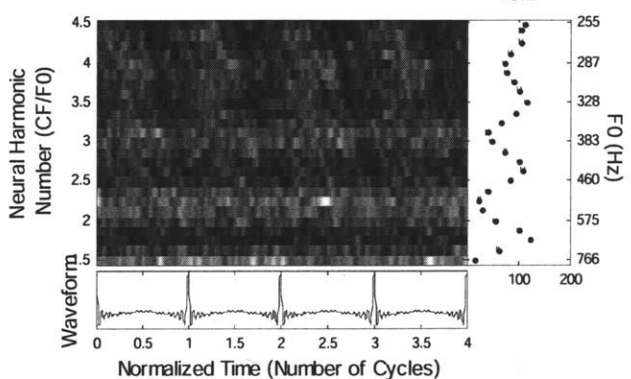**Figure 34:** Responses of a Chop-T unit with CF 1149 Hz to harmonic complex tone stimuli. The MASD profiles are not plotted for this unit because there is not phase locking in the patterns in response to all F0 tones.

AN fibers and two Pri units with lower CFs (1456 and 2075 Hz). This may suggest several Pri units have better frequency selectivity than AN fibers at similar CFs.

Figure 35B shows that at low levels (left panel), two Chop-T units (between 1 and 2 kHz, in black) responded with stronger rate representations than those observed in the AN and other unit types. The Chop-T with lower CF (1149 Hz) corresponds to the unit whose responses are

Figure 35: Comparison of CN unit rate representations of pitch to AN data taken from Cedolin (2006). Light blue: rate representation of AN fibers. Pink: MASD representation of AN fibers. Triangles: data points that exceeded a harmonic strength of 150. **A.** Blue: Pri, Red: Pri-N. **B.** Green: Chop-S, Black: Chop-T, Blue: On-L, Red: Flat. In both figures, open circles mark data from CN units that were phase sensitive to the Huffman stimuli.

shown in figure 34. All three recordings in response to complex tones were made at low levels (< 20 dBreTh). Figure 35B also shows that the flat unit (CF = 1419 Hz, whose responses are shown in figure 32) has very strong rate representations at all levels. This unit is very robust with level, and its rate cue in the highest level range greatly exceeds AN rate, AN MASD, and other CN rate representations in units with the same CF.

The other CN unit types (Pri-N, Chop-S, and On) have rate responses that fall within the range of AN data. There was no apparent correlation between phase sensitivity to the Huffman sequences and the relative HS values.

Discussion

The present study was motivated by previous results from the AN that a spatio-temporal representation of pitch based on the MASD better accounts for pitch psychophysically than the traditional rate-place representation. Specifically, the MASD is more robust with level and predicts the upper frequency limit to pitch. The current study looked for evidence that the spatio-temporal cues available in the AN are extracted in the cochlear nucleus and possibly converted to rate-place cues.

We have recorded from CN units with transient complex stimuli designed to vary the spatio-temporal pattern of AN firing in order to test for sensitivity to the relative timing of AN inputs in the CN. In addition, we have recorded CN responses to harmonic complex tone stimuli to determine whether rate responses of CN units are more robust with level than those of AN fibers. We hypothesized that CN units that are sensitive to changes in the phase spectrum of Huffman sequences would extract spatio-temporal cues in the AN to the pitch of complex tones, and give robust rate responses to harmonic complex tones.

Responses to Huffman stimuli

The majority of the CN units that we recorded from were insensitive to the manipulations of the phase pattern of the Huffman stimuli. We defined units to be sensitive if they displayed changes in their responses to the stimuli that were significantly greater than those observed in the AN. CN unit sensitivity was further categorized depending on the direction of sensitivity. Most of the units that were phase-sensitive were PS-, indicating better response to the sharp-transition

stimulus than AN fibers. Several unit types were PS-, and possible mechanisms for this sensitivity may differ among types as discussed later in this section.

We have also found a small number of CN units that prefer to respond to the broad phase transition Huffman stimulus, which excites AN fibers innervating neighboring cochlear regions more coincidentally than the sharp-transition stimulus. This was observed in PS+ units at all levels and in PS+/- units at lower levels. This suggests that PS+ and PS+/- units may receive converging afferent inputs from AN fibers tuned to different frequencies and be sensitive to the relative timing of these input discharges. Carney (1990) proposed that this is consistent with a coincidence detection mechanism operating on inputs from different CFs at the level of the CN. In general, unit types with well-timed onset responses may also be consistent with a coincidence mechanism (Young et al 1988), but do not require differing CF inputs.

Auditory nerve responses

In order to rigorously evaluate CN unit sensitivity to manipulations of the temporal discharge pattern of its inputs, AN responses to the Huffman stimuli were collected to establish a baseline to which CN data were compared. Our results show that near threshold levels, RD_{NS} were centered around zero, and rate responses were nearly the same to the three stimuli, suggesting no preference for gradual or rapid phase transitions (figure 13). In contrast, Carney (1990) found that AN fibers preferred the Huffman stimulus with slow phase transition (r=0.85) at low levels. As explain in Methods, we used a single positive impulse as our input sequence to generate the Huffman stimuli. By having an additional negative impulse in her input, her stimuli effectively truncated responses to the longer duration stimulus, causing a higher RD and explaining this difference between the results.

The stimulus with rapid phase transition (r=0.95) is longer in duration, and AN fibers have longer duration responses to these stimuli compared to the stimuli with broader transitions. We found that as level increased, most AN fibers displayed an increasingly negative RD_N. If the peak heights in response to all three Huffman stimuli continued to increase with level, the overall response to all three would increase. However at high levels, we find an overall preference for the longer-duration stimulus, suggesting that saturation of the peak firing rate occurs. Carney's (1990) results were consistent with this observation. She found that the sensitivity of the fibers

usually changed directions as level increased, preferring the longer duration stimulus and resulting in a negative rate difference.

For CN units with responses similar to those of AN fibers (all Pris and most PLs), Carney (1990) defined CN sensitivity by comparing RDs near threshold levels. In this study, we define sensitivity for Pris and PLs by comparing RD_{NS} between the AN and these units at comparable levels (in dBreTh).

Primary-like responses

Because they receive only one or two primary afferent inputs, spherical bushy cells in the CN are not expected to be sensitive to the temporal pattern of their inputs compared to other cell types. Our result that 7 out of 28 (25%) of the Pri units are sensitive to the Huffman stimuli (6 PS- and 1 PS+/-) is therefore somewhat surprising. Figure 27B shows that at high levels (>40 dBreTh), the overall RD_{NS} of all Pri units are shifted towards more negative values compared to AN results. This could be due to several reasons. Comparison of temporal discharge patterns for AN fibers versus Pri units suggests a possible explanation for this different behavior.

Our results suggest a temporal sharpening of both the overall response durations and sharpening of onset patterns between AN fibers and Pris. Figure 28A shows that the normalized duration of AN fiber responses increase with level. At lower levels (<50 dBreTh), the response durations of Pri units also increase with level and are, on average, very close to those measured in the AN. However as level increases above 50 dBreTh, the Pri response durations no longer continue to increase and, on average, have shorter overall durations compared to AN fibers. This result indicates that temporal sharpening may occur between the AN and Pris that affects the overall response durations to transient complex stimuli at high levels.

Another type of temporal sharpening that may occur between AN and Pri units is in the onset response. When comparing the response patterns of an AN fiber (Figure 12) and a Pri unit (Figure 19) with similar CFs to Huffman stimuli at similar levels, the Pri has a more sharpened onset response at 50 dBreTh than the AN response at 52 dBreTh. If the Pri unit in figure 19 receives inputs from fibers resembling the one portrayed in figure 12, more precise timing of the stimulus onset occurs at the synapse. These results indicate that several forms of temporal sharpening may occur at the synapses between AN fibers and Pri units.

The direction of sensitivity in most of the Pri units (PS-) suggests that they do not behave as cross-frequency coincidence detectors. Instead, a simple coincidence detector might achieve this type of temporal sharpening and does not require that the AN input fibers are tuned to different CFs.

Another possible factor that may influence the timing of responses in Pri units is inhibition. AN fibers excite spherical bushy cells via large endbulb synapses, which also receive inhibitory inputs from the DCN and other VCN cell types (Wu and Oertel 1986, Oertel et al 1990), as well as from higher order nuclei, such as the trapezoid body (Schofield 1991, 1994, Covey et al 1984, Warr and Beck 1996). A study in the gerbil AVCN demonstrated that for a large number of spherical bushy cells, inhibition enhances the temporal precision of the coding of the stimulus onset (Kopp-Scheinpflug et al 2002). These mechanisms may also influence response to Huffman stimuli.

Even though units in the CN are labeled as phase sensitive or not phase sensitive, our results do not indicate that sensitivity is bimodal. In particular, Pris seem to have a range of sensitivity, as well as a range of prominence of onset responses (as defined by our onset-to-steady-state-rate ratio in response to STBCFs). Coincidence detectors are thought to be associated with prominent onset response patterns. However our data do not indicate a correlation between phase sensitivity and a high onset-to-steady-state-rate ratio. The only Pri unit that responded with sensitivity in the direction consistent with cross-frequency coincidence detection (PS+/-) had onset-to-steady-state-rate ratios of 1.2 in its responses to STBCFs, which was minimum in the range defined for Pris. Furthermore, all other sensitive Pri units (6 PS-) had ratios less than 1.6, suggesting that Pris with flatter responses to STBCFs are more likely to be phase sensitive. However, low onset-to-steady-state-rate ratios were also seen in phase-insensitive units, including many other Pris as well as the Flat unit. A variety of mechanisms may contribute to the diversity in responses among Pri units, suggesting these units may not behave in a one-to-one fashion with their AN inputs, as is commonly thought.

Phase locker and Hi-sync responses

PL and HiS units displayed a variety of characteristics in response to the Huffman stimuli. Figure 27A shows that both types included units that resulted in increasing and decreasing RD_N values with level as well as units that were phase-insensitive. The variety of

responses is consistent with the idea that both of these groups are likely comprised of several cell types. Additional information such as the units' locations within the CN would be helpful for further classification, but was not looked at in detail.

The sensitivity of some PL and HiS units to the Huffman sequences suggests that they may receive AN inputs from different CFs and are sensitive to their relative timing. We observed one PS+ PL and one PS+ HiS unit, which were unique in our sample of AN and CN units in that RD_N increased with level. The sensitivity of these units suggests that there may be a subclass of PLs that fire preferentially when their inputs are discharging in synchrony, as expected for a cross-frequency coincidence detector.

Primary-like-with-notch and Onset-L responses

For CN units with Pri-N, onset, and chopper responses, Carney (1990) assessed phase sensitivity by qualitative analysis of the response discharge patterns and found that Pri-N and Chop-T units were particularly sensitive to the manipulation of the temporal pattern of their inputs. We defined phase sensitivity in these units primarily by using the RD_N metric while also considering the overall shape of the PSTHs.

Globular bushy cells receive a number of converging AN inputs, have well-timed onset response patterns, and are associated with nonlinear membrane properties. In response to Huffman stimuli at low levels, Pri-N units responded slightly more to the stimulus with broad phase transition ($RD_N > 0$). As level increased, the sensitivity changed direction, and units responded more to the longer stimulus. The changes in sensitivity were not a result of longer response durations to the sharp phase transition stimulus. Instead, the changes came from differences in the onset responses, as seen in the PSTHs in figure 21 at 26 and 36 dBreTh. These results are qualitatively similar to what Carney (1990, figure 8a) observed in Pri-N units, and contrast to results from the AN, where the increase in response duration is primarily responsible for the negative RD_N at high levels.

CN units with response type On-L may result from recordings from bushy cells and can form a continuum with Pri-N units. It is possible that our Huffman recordings from the On-L unit are from a globular bushy cell. Both its temporal and rate responses to Huffman stimuli as shown in figures 23 and 27C were similar to those of the two Pri-N units in figures 21, 22, and 27B.

Chopper responses

Like globular bushy cells, stellates receive converging AN inputs and have well-timed onset responses. A striking feature in the firing patterns of chopper units that separates them from other types is their regularity. Unlike PLs, choppers tend to fire at regular intervals that do not correspond to their CF. In addition, their long latencies suggest that choppers may perform an integration over their AN inputs and over time before firing (upon reaching a "threshold") and resetting their membrane potentials (van Gisbergen et al 1975).

The multimodal responses of the Chop-T to Huffman stimuli in figure 26 suggest that if stellates perform temporal integration over their AN inputs, the Chop-T unit reaches "threshold" more than once during its response. For each stimulus presentation, most of this unit's first spikes occurred in the first peak of the multimodal PSTH. This result is consistent with click responses in chopper units, where PSTHs typically display 1 to 3 peaks, and the times of the first, second, and third spikes correspond to the first, second, and third peaks respectively (Bourk 1976). The two Chop-T units in this study displayed no preference for any of the Huffman stimuli at low levels but responded better to the stimulus with sharper phase transition at higher levels. However, these differences were too small for these units to be considered phase-sensitive.

Unlike other types, the responses of chopper units to Huffman stimuli do not have peaks spaced at time intervals of 1 / CF. The long latency, single wide peak patterns of the Chop-S responses indicate that these units prefer to fire within a certain time window after the stimulus onset. On average, the Chop-S unit in figure 25 only fired once per stimulus presentation. It is possible that this cell performs an integration of AN inputs over time such that it fires if "threshold" is reached. If that is true, then differences between the multimodal Chop-T and single-peak Chop-S responses suggest that either "thresholds" differ between the Chop-T and Chop-S units, or the membrane potentials are reset differently, such that Chop-Ts are more likely to fire again after an initial response.

Both sensitive Chop-S units were PS-, though since all of our recordings were made at relatively high levels (figure 27C), it is possible that if additional recordings were made at lower levels, we cannot rule out that the sensitivity of these units would be PS+/-. However, chopper units are known to have lower STBCF thresholds and higher click thresholds than AN fibers (Bourk 1976), partly justifying our use of higher level stimuli for these units. More importantly,

considering their long latencies and regular firing patterns, it is not likely that Chop-S units behave as coincidence detectors, which require precise temporal coding of inputs rather than intrinsic cell properties that strongly influence response patterns.

Harmonic complex tones

In order to extract the spatio-temporal pitch cues, 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 remain unclear. One possible mechanism is a lateral inhibition network of neurons that prefer a non-coincident relationship among their inputs. The spatial derivative operation is meant to mimic the response of hypothetical neurons that would have rate-place profiles which represent the difference in firing pattern between AN fibers innervating adjacent cochlear regions (Shamma 1985b). Maxima in the MASD profile of AN fibers indicate locations of rapidly varying phase changes. Another possibility is a cross-frequency coincidence detector, where a neuron would fire best when its inputs occur in synchrony. In the absence of rate cues (where the firing rate is saturated), a cross-frequency coincidence detector neuron would have maxima in its rate-place profile where the MASD profile has minima (where response latencies of its inputs are more uniform), and vice versa.

We have recorded responses of phase-sensitive and phase-insensitive CN units to harmonic complex tones and searched for a correlation between phase-sensitivity and enhancement or degradation of their rate representations of pitch compared to those in the AN.

Pri responses were qualitatively similar to those previously recorded in the AN. Harmonic resolvability in the rate profiles of Pri units increased with increasing CF, and the rate representations of pitch in these units tended to degrade with level. We found several Pris that better represented pitch in their rate responses at medium and high levels compared to AN fibers. These included both phase-sensitive and insensitive Pris. There is not an obvious correlation between phase-sensitivity and robust pitch representations in Pris.

In response to complex tone stimuli, the phase-insensitive Flat unit displayed a significantly better rate representation of pitch with increases in level. This was never observed in any other unit. Furthermore, this unit's rate cues were significantly stronger than either rate or MASD cues in the AN in the highest level range. This result indicates that there are cells in the

CN that represent pitch in their rate responses in a way that is more consistent with psychophysical observations than the AN. However, the Flat unit's insensitivity to the phase spectrum of Huffman sequences suggests that the improved rate representation did not result from spatio-temporal sensitivity. Instead, it is possible that this unit is using some other type of cue to enhance its representation of pitch. The response patterns of this unit to Huffman stimuli are similar to those of low SR AN fibers. It is possible that the Flat unit receives inputs from these fibers, which have wide dynamic range and could account for the robust rate representation.

All PL and HiS units displayed poorly resolved harmonics in their rate responses and low HS values, which is consistent with poor frequency selectivity in low CF fibers. None of our recordings from Pri-Ns or On-Ls in response to harmonic complex tones indicated that the rate representation of pitch in these units was stronger than in AN fibers with similar CFs. One unit of each type was PS+/-, but neither displayed responses to the complex tones that were robust with level. It is difficult to interpret our finding that several low-frequency CN units were sensitive to the Huffman sequence manipulations in the context of harmonic resolvability because the periphery has poor frequency selectivity and thus does not have strong spatio-temporal cues.

A study of vowel representation in Pris and choppers demonstrated that Chop-T units had rate-place representations of the vowel spectrum that were robust with level (Blackburn and Sachs 1990). However our results indicate that both types of choppers (Chop-S and Chop-T) have degraded rate representations with increases in level. Both have HS values similar to those of AN fibers with the same CFs and in the same level ranges. Out of 5 Chop-S units from which we recorded response to complex tones, two were PS- and had rate representations stronger than the other Chop-S units. This result suggests that even though it is unlikely that spatio-temporal cues are being extracted in these units, there could be a distribution within the Chop-S group of units that correlates relative phase-sensitivity with stronger pitch representations. Of course, this is based on a very small sample of units and should be verified with more recordings with the same stimuli or with other stimuli. While it is unlikely that Chop-S units extract the spatio-temporal pitch cues in the AN, physiological and modeling studies have shown that these units may play an important role in temporal pitch processing (Wiegrebe and Winter 2001, Wiegrebe and Meddis 2003).

Our findings suggest that several CN units are able to do some sort of temporal processing. However, the lack of detailed anatomical information about the convergence of AN fibers and possibly inhibitory inputs onto cells in the CN has limited our ability to interpret these results. A model that accounts for what is known about this convergence for each unit type and can predict single CN unit responses to complex stimuli would be extremely useful for further understanding of what type of processing occurs at the level of the CN.

Conclusions

Our results from AN and CN responses to Huffman stimuli suggest that only a small fraction of CN units are sensitive to the relative timing of their AN inputs across different CFs. We hypothesized that the rate response to harmonic complex tones of phase-sensitive units would provide a pitch representation that is robust with level. We first evaluated CN unit sensitivity to the spatio-temporal pattern of their inputs with transient complex stimuli. AN responses to these stimuli justified pooling over a wide range of CFs but indicated that level needed to be carefully monitored to determine the sensitivity of CN units. Only a few units were found to have sensitivity in the direction consistent with cross-frequency coincidence detection. Our small data set of CN unit responses to harmonic complex tone stimuli indicate that a few CN units provide better rate representations of pitch compared to AN fibers with similar CFs. However there is little evidence for a correlation between phase sensitivity and a robust rate representation of pitch in CN units.

In this study, we defined CN sensitivity to the Huffman stimuli primarily by using the RD_N metric, which is an overall measure of response. However we have found that the temporal response patterns can clearly differ for Huffman sequences with different phase transition bandwidths. More work can be done on refining the definition of "sensitivity" by taking into account some of these temporal characteristics. In addition, a modeling study that manipulates CN cell properties (such as the number of inputs and their CFs) would give valuable insight into the mechanisms which give rise to such a variety of responses to Huffman stimuli and harmonic complex tones. This study focused on the hypothesis that CN units could represent the spatio-temporal pitch cues from the AN in their rate responses. It is also possible that instead, CN units enhance or at least preserve these cues and that they are extracted at later stages. A study

examining the available pitch cues in the CN may suggest where to look next for the pathway that processes pitch.

Acknowledgments

References

Bernstein J. G. (**2005**). Pitch perception and harmonic resolvability in normal-hearing and hearing-impaired listeners. (*PhD dissertation.* Massachusetts Institute of Technology).

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

Bernstein J. G., Oxenham A. J. (**2003b**). Effects of relative frequency, absolute frequency, and phase on fundamental frequency discrimination: Data and an autocorrelation model. *J Acoust Soc Am.* **113**:2290.

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 /ɛ/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. *J. Neurophysiol.* **63**:1191-1212.

Bourk T. R. (**1976**). *Electrical Responses of Neural Units in the Anteroventral Cochlear Nucleus of the Cat* (PhD thesis). MIT, Cambridge, MA.

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.

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. *Neuroscience,* **4**:1909-1923.

Cariani P. A., Delgutte B. (**1996a**). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.* **76**:1698-1716.

Cariani P. A., Delgutte B. (**1996b**). Neural correlates of the pitch of complex tones: II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J. Neurophysiol.* **76**:1717-1734.

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., 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. (**2007**). Spatio-temporal representation of the pitch of complex tones in the auditory nerve. *In Hearing – From Basic Research to Applications.* Edited by Kollmeier. New York (in press).

Cedolin L. (**2006**) Neural Representations of Pitch: Role of Peripheral Frequency Selectivity. (*PhD dissertation,* Massachusetts Institute of Technology).

Covey E., Jones D. R., Casseday J. H. (**1984**). Projections from the superior olivary complex to the cochlear nucleus in the tree shrew. *J. Comp. Neurol.* **226**:289-305.

Cutler A., Dahan D., van Donselaar W. (**1997**). Prosody in the comprehension of spoken language: a literature review. *Language and Speech.* **40**:141-201.

Efron B., Tibshirani R. J. (**1993**). An introduction to the bootstrap. New York: Chapman & Hall.

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

Fletcher H. (**1940**). Auditory patterns. *Reviews of Modern Physics* **12**: 47-65.

Glasberg B. R., Moore B. C. J. (**1990**). Derivation of auditory filter shapes from notched-noise data. *Hear. Res.* **47**:103-138.

Godfrey D. A., Kiang N. Y., Norris B. E. (**1975**). Single unit activity in the posteroventral cochlear nucleus of the cat. *J. Comp. Neurol.* **162**:247-268.

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.

Javel E., McGee J. A., Horst W., Farley G. R. (**1988**). Temporal mechanisms in auditory stimulus coding. *In Auditory Function: Neurological Bases of Hearing.* Edited by G. M. Edelman, W. E. Gall, and W. M. Cowan. Wiley: New York. pp. 515-558.

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.

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

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.

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. Ciba Found Symp:* 241-273.

Kopp-Scheinpflug C., Dehmel S., Dorrscheidt G. J., Rubsamen R. (**2002**). Interaction of excitation and inhibition in anteroventral cochlear nucleus neurons that receive large endbulb synaptic endings. *J. Neurosci.* **24**:11004-11018.

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.

Licklider J. C. R. (**1951**). A duplex theory of pitch perception. *Experientia.* **7**:128-133.

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.

Meddis R., Hewitt M. J. (**1991a**). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. *J. Acoust. Soc. Am.* **89**:2866-2882.

Meddis R., Hewitt M. J. (**1991b**). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity. *J. Acoust. Soc. Am.* **89**:2883-2894.

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

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. New York, NY: Springer-Verlag.

Nolan F. (**2001**). Intonation in speaker identification: an experiment on pitch alignment features. *Forensic Linguistics.* **9**:1-21.

Oertel D., Wu S. H., Garb M. W., Dizack C. (**1990**). Morphology and physiology of cells in slice preparations of the posteroventral cochlear nucleus of mice. *J. Comp. Neurol.* **295**:136-154.

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

Osen K. K. (**1969a**). Cytoarchitecture of the cochlear nuclei in the cat. *Journal of Comparative Neruology.* **136**:453-484.

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.

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.

Pfeiffer R. R. (**1966**). Classification of response patterns of spike discharges for units in the cochlear nucleus: Tone burst stimulation. *Experimental Brain Research.* **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.

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

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. Stockholm: Samson und Wallin. p. 345.

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.

Rhode W. S., Smith P. H. (**1986**). Encoding timing and intensity in the ventral cochlear nucleus of the cat. *J. Neurophysiol.* **56**:261-286.

Scheffers M. T. M. (**1983**). Sifting vowels: auditory pitch analysis and sound segregation. (*PhD dissertation*, Rijksuniversiteit te Groningen, The Netherlands).

Schofield B. R. (**1991**). Superior paraolivary nucleus in the pigmented guinea pig: separate classes of neurons project to the inferior colliculus and the cochlear nucleus. *J. Comp. Neurol.* **312**:68-76.

Schofield B. R. (**1994**). Projections to the cochlear nuclei from principal cells in the medial nucleus of the trapezoid body in guinea pigs. *J. Comp. Neurol.* **344**:83-100.

Schouten J. F. (**1938**). The perception of subjective tones. *Proc. Kon. Akad. Wetenschap.* **41**:1086-1093.

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

Seebeck A. (**1841**). Beobachtungen uber einige Bedingungen der Entstehung von Tonen. *Annalen fur Physik und Chemie.* **53**:417-436.

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 Natl Acad Sci USA.* **99**:3318-3323.

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., 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.

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.

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

Summerfield A., Assmann P. (**1990**). Modelling the perception of concurrent vowels: vowels with different fundamental frequencies. *J. Acoust. Soc. Am.* **88**:680-697.

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

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

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

Young E. D. (**1986**). Response characteristics of neurons of the cochlear nuclei. *In: Hearing Science*, edited by C. I. Berlin. San Diego, CA: College-Hill. P. 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.

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

Warr W. B., Beck J. E. (**1996**). Multiple projections from the ventral nucleus of the trapezoid body in the rat. *Hear. Res.* **93**:83-101.

Wiegrebe L., Meddis R. (**2003**). The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. *J. Acoust. Soc. Am.* **115**:1207-1218.

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

Wu S. H., Oertel D. (**1986**). Inhibitory circuitry in the ventral cochlear nucleus is probably mediated by glycine. *J. Neurosci.* **6**:2691-2706.