Title

Effects of continuous background noise on the representations of tonal and speech stimuli in cat primary auditory cortex

Author

Wong, Sarah Wenlong Chuang

Publication Date

1998

Effects of Continuous Background Noise on the Representations of Tonal and Speech Stimuli in Cat Primary Auditory Cortex

by

Sarah Wenlong Chuang Wong

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Speech and Hearing Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

[redacted]

Date **University Librarian**

Degree Conferred: . . . .

Acknowledgements

Many people and opportunities helped in the creation of this dissertation research and writing. I would like to thank my parents for teaching me the value of having a good education and the meaning of hard work. I would like to thank my husband George for the much needed words of encouragement and support during times when the work was becoming overwhelming and too much for me to bear.

With regards to the content in this dissertation, none of it would have been possible without the instruction and assistance of my research advisor Dr. Christoph Schreiner. To my dissertation committee members, I would like to thank Dr. Ervin Hafter for his comments which helped to provide me with a renewed perspective of research in the hearing sciences and to Dr. Michael Merzenich for his review of my research work. I would also like to thank Dr. Patricia Leake and members of her laboratory for allowing me to use their resources while I was conducting my behavioral training experiments.

Certainly fellow graduate students and post-docs in the Schreiner laboratory and other members in the Keck Center have been instrumental at various stages of my graduate studies in providing assistance with data collection, data analysis, and other technical support. I would like to remember the following colleagues for helping me make progress with my work: B. Calhoun, M. Brosch, B. Bonham, L. Miller, H. Read, H. Attias, M. Escabi, X. Wang, S. Kumar, R. Beitel, and P. Bedenbaugh.

And finally, I would like to remark about my time as a graduate student in the program of Speech and Hearing Sciences here at the University of California in San Francisco. Though it was a longer journey than what I had anticipated, the opportunities to explore various clinical and basic science aspects of speech and hearing sciences have provided me with a well-rounded education in the field. Therefore, I would like to thank the faculty and administration of the University of California at both San Francisco and Santa Barbara for their contribution to my graduate education.

Abstract

Effects of continuous background noise on the representations of tonal and speech stimuli in cat primary auditory cortex.

by

Sarah Wenlong Chuang Wong

Cortical representations of sound have been studied using tonal and complex stimuli such as animal vocalizations and speech sounds. From animal studies, it is known that cortical representations of tonal stimuli demonstrate level-dependent behavior in firing rate, bandwidth, and latency. Yet, little is known about cortical representations of speech at different sound intensities; moreoever, less is known about the effects of continuous background noise on cortical representations of tonal and speech sounds. The aim of these experiments was to study the effects of continuous background noise on these cortical representations of tonal and speech sounds in the primary auditory cortex (AI) of cats.

First, the effect of continuous background noise on the frequency specificity of single unit neurons in AI was probed by inducing a 20 dB-threshold shift and measuring the resulting changes in the receptive field. Response differences attributable to monaural/binaural stimulation and stimulus interaural phase differences were considered in order to study the range of background noise effects on the response properties. Second, the effect of a fixed level of continuous background noise on the spatial relation of multiple-unit receptive field properites in AI was examined. Lastly, the effect of continuous background noise on the cortical representation of speech was probed by studying the temporal firing patterns of multiple-units in AI. Concurrently, response differences from changing signal level were examined to study the range of intensity effects on neural activity. Animals were trained to make discriminative

behavioral responses to speech sounds in order to study the effects of stimulus relevancy.

From these experiments, it was found that continuous background noise inversely alters the spectral bandwidth of single units and the spatial distribution of multiple-unit spectral bandwidth. Rate and temporal representations of speech sounds were found to be dependent on intensity and signal-to-noise levels, whereas the spatial-temporal firing pattern was more robust to changes in level, especially signal-to-noise level. These experimental findings reveal how simple perturbations in contextual backgrounds have profound impact on spectro-temporal receptive field properties of single unit neurons and neural assembly responses to tonal and speech stimuli in AI.

Christoph Schreiner 6-8-95

Christoph E. Schreiner, M.D., Ph.D., Dissertation Committee Chairperson Date

Table of Contents

Chapter 3: The cortical representation of speech syllables and the effects of continuous background noise on its representation. . . . . Page 147

List of Tables

List of Figures

Introduction

Listening to relevant sounds embedded in noisy backgrounds is what we as humans experience everyday as we acoustically interact with the people and objects in our environment. The simple act of listening and objectively responding to these meaningful sounds is undoubtedly a complex physiological, perceptual, and cognitive task whose processing occurs and is completed within a matter of milliseconds. How important the task of signal and noise segregation is for maintaining our normal hearing capacity becomes evident when patients with only a mild-to-moderate hearing loss complain about their reduced ability to communicate in noisy environments. By studying how relevant sounds such as speech are represented by neural activity in the brain and how noise interferes with these representations, one can begin to understand the neural algorithms that are used to achieve a sound reception that is sufficient for successful communication which leads to meaningful social behavior.

Although complex sounds are first transformed acoustically by the auricle and ear canal of the outer ear, mechanically by three bony ossicles composing the middle ear, and then hydrodynamically in the cochlea of the inner ear, it is the electrical impulses transmitted along neural lines following the neurochemical synaptic release of the inner ear haircells that comprise the initial stages in the study of neural representations of sound. Up through the level of the cortex, neurally-represented sounds are filtered systematically according to frequency (spectral) and timing (temporal) information from the acoustic signal. In the auditory cortex, the representation of sounds has been considered modifiable based upon the degree of auditory experience an individual has acquired. Another line of evidence demonstrating a difference between cortical and subcortical processing of sounds is found in humans with cortical word deafness or other abnormal central auditory neuropathy, who have normal hearing abilities for simple sounds, but manifest the inability to discriminate and recognize complex sounds such as words. Therefore, the cortical transformation and representation of sound information must necessarily be found in the neural encoding of complex sounds across a wide range of intensities and signal-to-noise levels.

Neural representations of speech sounds have been studied in humans through the use of imaging techniques such as PET (positron emissions topography), MRI (magnetic resonance imaging), and MEG (magneto-encephalography). The use of non-invasive techniques imaging techniques poses great limitations in understanding neural representations of speech on the cellular level which is required from an auditory neuroscience perspective. Thus, the use of an animal model such as the cat, for which studies of neural encoding of simple and complex stimuli at several stages in the auditory system have revealed a basic understanding of how sounds are processed, can allow for the exploration of speech representations at a finer grain of neural detail than would otherwise be found in non-invasive human studies.

One method in studying cortical representations of speech in variant background contexts is to examine first the cortical representations of tonal stimuli in similar contexts. Neural responses to simple sounds, such as tone-bursts, can provide a basic, although not always complete, characterization of the filter properties leading up to and including the studied neuron. The frequency response area for pure tones of cortical neurons allows several measures of the receptive field of a neuron. By measuring the frequency response areas of many neurons within one section of the primary auditory cortex, one can study the topography of response properties and the effects of contextual backgrounds on the topography. Systematic organizations in the topography of simple sounds may provide a basis for structuring the cortical representation of speech and provide for its generalization across varying signal intensities and signal-to-noise levels.

Everyday sounds are typically processed using two ears, i.e. binaurally. It has been shown that binaural processing contributes substantially in the extraction of signals from noise. There are interaural disparities in the signals that reach both ears that need to be accounted for when studying the representations of binaurally-presented sounds. These interaural disparities include phase differences which have been shown to affect low-frequency sounds mainly and level differences which affect high-frequency sounds. Speech sounds are complex sounds whose transitions and formant structure are composed of mainly low-frequency sounds. Thus, to test whether cortical representations of speech sounds may be influenced by interaural disparities in phase information, it is necessary to examine the impact of interaural phase differences on frequency response properties.

Studying sounds that are processed in the presence of varying background contexts may aid in understanding a neurophysiological basis for speech intelligibility in the presence of background noise. Moreover, finding answers to questions of how the cortex represents auditory signals in the presence of background noise is paramount to helping those with speech and hearing impairments that may not necessarily be attributed to peripheral hearing loss.

The aim in this dissertation research was to study the effects of continuous background noise upon representations of simple and complex sounds in the primary auditory cortex of the cat. Four experimental questions were asked.

1. From a single unit study perspective, are there spectral tuning differences between monaural and binaural processing of tone bursts in noise stimuli?

2. From a single unit study perspective, is the binaural configuration of the tone bursts in noise stimuli reflected by differences in the spectral receptive field properties?

3. From a population coding perspective, are maps of spectral receptive field parameters invariant to the influence of continuous background noise?

4. From a population coding perspective, how is the representation of speech syllables of different signal intensities and signal-to-noise levels reflected in the spatial-temporal firing patterns of neurons in auditory cortex?

Combined, the answers to these questions provide important insights into the role of primary auditory cortex in the processing and representation of complex sounds.

Chapter 1: A single unit study of the effects of continuous background noise on the representation of tonal stimuli

1.0.0 Introduction

The effect of continuous background noise on tone-evoked cortical cell activity in primary auditory cortex has mostly been studied in the contralateral stimulation condition (Phillips, 1985; Phillips, 1990; Phillips and Cynader, 1985; Phillips and Hall, 1986). Previous studies have shown that continuous noise affects cortical cell activity differently than non-continuous, gated wide-band noise stimuli (Phillips and Cynader, 1985). Defined by their firing behavior at high intensity levels, monotonic (maintaining excitatory activity throughout the cell's dynamic range) and non-monotonic neurons (exhibiting inhibitory activity at high SPL) in the cortex respond differently to wide-spectrum noise (Phillips, et al, 1985) such that monotonic cells respond to noise stimulation while non-monotonic cells often do not. Neither class of cells responds in a sustained fashion to continuous noise. Furthermore, the effect of continuous noise masking of tonal stimulation is similar in both monotonic and non-monotonic neurons. Continuous background noise has been found to shift the dynamic range of cortical neurons thereby elevating response threshold and prolonging spike latency. At masked thresholds in excess of 15 dB, slope of the tone intensity function steepens (Phillips and Hall, 1986). The neural mechanism by which excitatory and inhibitory events are controlled during tone-in-continuous-noise stimulation are not known. Since monotonic and non-monotonic neurons behave in a like manner to tone-in-continuous-noise stimulation, it has been suggested that neural responses are modified by a process that is common to both cortical cell groups such as a spectrally general adaptation, which may only partially be conveyed by short-term adaptation originating from the auditory periphery (Phillips, 1985).

The relevancy of studying the effect of background noise on cortical activity to binaural acoustic stimulation can only equivocally and implicitly be understood in terms of our knowledge of complex sound processing in primary auditory cortex.

Psychophysical evidence from binaural masking level difference (BMLD) studies have already implicated certain binaurally-configured signal-in-noise conditions (antiphasic signals in correlated noise or antiphasic noise in correlated signals) to be more effective in producing unmasking effects (Hirsch, 1948; Licklider, 1948) which yield greater signal detection through a mechanism hypothesized to be different from that underlying sound localization (Henning, 1974). The ability to separate signals from background noise may have something to do with lateralization of components within the waveform (Henning and Gaskell, 1981). Physiological studies in the inferior colliculus (IC) of guinea pigs have provided evidence for a neural basis for sensitivity differences due to changes in binaurally-configured stimuli (Caird, Palmer, and Rees, 1991; Jiang, McAlpine, and Palmer, 1997; McAlpine, Jiang, and Palmer, 1996a).

It has also long been known that the physiological mechanisms underlying low-frequency binaural hearing, as reflected at the level of the auditory cortex (Brugge, et al., 1969; Brugge and Merzenich, 1973; Reale and Brugge, 1990), use interaural disparities in signal phase (time) for source locating computation originating in the brainstem (Chan, Yin, and Musicant, 1987; Kuwada and Yin, 1983; Yin, Chan, and Carney, 1987). Binaural interaction neurons (EE/EI), defined by the comparison of neural activity from combined binaural inputs versus monaural input, have been observed in the brainstem (Goldberg and Brown, 1969; Yin and Chan, 1990) and in the cortex (Hall and Goldstein, 1968; Phillips and Irvine, 1983), and hypothesized to be spatially segregated in the cortex (Imig and Adrian, 1977; Middlebrooks, Dykes, and Merzenich, 1980). Simplified, summation neurons (EE) manifest response strengths similar to and stronger than monaural stimulation, although, subgroups within EE neurons such as those with low characteristic frequencies (CF) may show response strengths which vary as a function of interaural phase disparity of the signals (Reale and Brugge, 1990). Excitatory-inhibitory (EI) neurons manifest suppressed neural activity compared to monaural stimulation.

The ability to discriminate speech sounds in a free-field environment in the presence of background noise is a task that humans deal with on a daily basis. The discrimination of signals embedded in background noise has been considered a neural event occurring at the level of the cortex in the central auditory system. Patients with cortical deafness who have normal detection thresholds but have trouble recognizing and discriminating words (Phillips and Farmer, 1990) have provided one line of evidence for discrimination abilities residing at the cortical level. Another line of evidence arises from physiological masking studies which reveal a greater salience of the neural response to the masked signal in the cortex (Phillips and Hall, 1986; Phillips, 1990) than in the periphery due to the reduction of temporal activity to stimulus inputs with each ascending station of the auditory central nervous system (Delgutte, 1990; Phillips and Cynader, 1985). Psychoacoustic studies in mammals have indicated that average rate coding, and not temporal phase-locking (Dye and Hafter, 1980), may be used in the central nervous system in performing frequency discrimination in background noise in low-frequency ranges (Hienz, Sachs, and Aleszczyk, 1993; Sinnott and Brown, 1993). The temporal resolution capability of the cortex is known to be low, usually below 30 Hz; thus, some form of spatial-temporally distributed discharge pattern, as observed in other sensory cortices (Recanzone, Merzenich, and Schreiner, 1992), is thought to represent complex sounds in primary auditory cortex (Wang et al., 1995). It is hypothesized, as well, that the discrimination of signals in the context of background noise may be reflected by different spatial-temporal neural activity patterns in the primary auditory cortex.

In this current series of experiments, the question of how continuous background noise affect the cortical representation of tonal stimuli was approached by comparing receptive field properties of neurons in quiet and noisy backgrounds under monaural and binaural stimulation conditions. This approach is methodologically different than a previous study (Phillips and Kelly, 1992) which attempted to find differences in neural activity from auditory evoked potentials to binaurally-configured tone-in-continuous background noise stimulation. The intention of the current study is to sample a population of mostly low-frequency neurons and to determine the effect of continuous background noise on binaural spectral receptive fields of primary auditory neurons in AI while examining differences between 1) monaural versus binaural stimulation and 2) the impact of interaural signal phase on spectral parameters.

1.1.0 Methods

1.1.1 Surgery

A total of seven adult cats were used in this single-unit recording study. The primary auditory cortex from the right cortical hemisphere was studied.

Initially, cats were anesthetized by an intramuscular injection of ketamine HCL (10 mg/kg) in combination with the tranquilizer acepromazine maleate (0.28 mg/kg). Following cannulation of the brachial vein in the right foreleg, sodium pentobarbital (initially 25-29 mg/kg; titrated to effect) was administered with an IV injection to begin the surgical procedures. An injection of dexamethasone (0.14 mg/kg) to prevent brain edema together with atropine (1 mg/kg) to reduce salivation was administered subcutaneously. After attaching the EKG to monitor the heart rate and respiration and inserting the rectal temperature probe to monitor the animal's body temperature (which was maintained around 37.5 °C by wrapping a temperature-controlled heating pad around the animal's lower body), a tracheotomy was performed and a semisoft tracheal tube introduced to facilitate unobstructed breathing. The tracheotomy was also important for reducing sounds from breathing through an open mouth while the head of the animal was held by a standard orbital-maxillar-palatal holder.

For the craniotomy, the skin over the medial/posterior half of the skull was reflected and the bone over the primary auditory cortex was removed. Agarose was used to cover the exposed cortical surface to assist in maintaining the stability of the recording site.

After surgery, subsequent delivery of the anesthetizing agents during the recording study were infused intravenously. The anesthetic agent used in the single unit recordings was sodium pentobarbital (titrated to effect). During surgery and stimulus-response recordings, cats were maintained at a surgical level of anesthesia evaluated by on-line monitoring of vital signs including heart rate, body temperature, and flexion and corneal reflexes.

1.1.2 Neural Recordings

Extracellular recordings from each sampled location in the auditory cortex were obtained through the insertion of tungsten microelectrodes (1 to 2 Mohms) mounted on a single hydraulic microdrive (Kopf) while presenting acoustical stimuli to the ears. Triggered spike recordings were pre-amplified (Differential Amplifier DAM-50), bandpass-filtered (1-3 kHz) (Dual Variable Filter, Type VBF 8), windowed through a spike discriminator (BAK Dis-1), amplified, sampled at a temporal resolution of 30 μs, and stored in a computer (IBM compatible) for off-line data analysis.

During the recording session, triggered spikes were visually displayed on a storage oscilloscope to set the criterion of the window discriminator for accepted spikes. Recording electrodes were lowered into the cortex approximately perpendicular to the cortical surface via a remotely controlled hydraulic microdrive to depths of 700 to 1100 microns.

1.1.3 Stimuli

Experiments were conducted in a double-walled sound-treated chamber (IAC). Acoustic stimuli were delivered through earphones (STAX 54) sealed in the ear canals. The stimulus set used was a series of 675 tone bursts varying in intensity and frequency which were used to assess the spectral sensitivity and specificity of auditory neurons. Tones were generated by a microprocessor (TMS 320C30; sampling rate of 110 kHz, 78 dB dynamic range). The intensity of the tones varied in 5 dB steps over a range of 70 dB, while the 45 different frequency values covered a range of 4 to 5 octaves centered around the estimated CF of the unit. Each pure tone signal was 50 ms in duration, ramped with a 3 ms (linear) rise/fall time, and repeated in a pseudorandom order every 350 ms. The pure tone series were summed with or without continuous noise depending upon the stimulus condition.

In these experiments, there were five auditory stimulation conditions: 1) pure tone series without noise delivered to the contralateral ear, or 2) to both ears; 3) pure tone series delivered to the contralateral ear with continuous background noise; 4) pure tone in-phase binaural series with continuous background noise; and 5) pure tone out-of-phase binaural series with continuous background noise. The continuous background noise was white random noise (Random Noise Generator, Type 1390-B) lowpass filtered at 10 kHz. The noise level was adjusted to 20 dB above the audio-visually assessed threshold at the characteristic frequency (CF) of the unit. This was achieved by driving a cell with a pure tone matching the CF of the unit approximately 20 dB above the cell's threshold, then masking that driven response with the level of noise required to achieve a masked threshold at 20 dB above the original threshold.

1.1.4 Data Analysis

Spectral sensitivity differences among the five different stimulus conditions were assessed by using parametric statistical measures. Spectral and temporal response parameters were measured from pure tone tuning curves recorded at each penetration site within the primary auditory cortex. The tuning curves analyzed were depictions of the excitatory receptive field of each recorded unit.

The pure tone response parameters measured included 1) the characteristic frequency (frequency with the lowest response threshold); 2) minimum threshold (the lowest tone intensity whose frequency matches the unit's characteristic frequency which elevated the discharge rate); 3) the lowest, and 4) highest frequency at 10 dB, 20 dB, and 30 dB above threshold to which the neuron responded (the lowest and highest frequencies measured at the three intensity levels were used to calculate the spectral bandwidth and Q= quality factor of sampled neurons); 5) minimum onset latency estimated from the neural discharge of multiple level tones whose frequency matches the unit's characteristic frequency; and 6) monotonicity of the rate-level function measured as the slope post-transition point.

1.2.0 Results

1.2.1 Monaural vs. Binaural Stimulation

Characterization of receptive fields of cortical neurons forms a principal experimental basis of attempts to understand the fundamental organization and information processing mechanisms in the auditory cortex. However, receptive field properties may depend on the conditions and contexts in which signals are embedded. In this study, a specific stimulus context, the effects of continuous background noise on receptive field properties in AI, was studied in detail. Cortical locations in which single unit responses to both monaural and binaural tonal stimulation were recorded and analyzed. Frequency response areas were collected from 28 single units recorded in the primary auditory cortex (AI) of seven adult cats. With such limited numbers of single units per cat, there was no attempt to derive a spatial distribution of the neural response properties for each of the four listening conditions. Instead, the analysis attempts to characterize any inherent differences between monaural and binaural single unit response properties as reflected in spectro-temporal tuning curves and any differences in each condition as a result of the addition of a continuous noise masker. This methodological approach differs from the work done in previous studies, which studied the effects of background noise only by measuring changes in rate-level functions. By analyzing the full excitatory receptive field, one can gain a more complete view and thus make a better comparison of the neural response to tones with and without background noise for either monaural or binaural stimulation.

Examples of a frequency response area (FRA) and the parameters that can be extracted from the excitatory receptive field region is presented in figures 1.1A-D. Figure 1.1A depicts a FRA recorded for monaural tonal stimulation and a FRA-derived histogram of the spike onset time and rate-level function measured around the characteristic frequency (CF).

Within the FRA, the length of each vertical line represents the spike response strength to a tone presented at that given frequency/intensity position. The minimum threshold is the lowest stimulus intensity level that evokes activity in the frequency response area-in this example, 52.5 dB SPL. The CF is measured as the stimulus frequency with the lowest stimulus intensity necessary to evoke neuronal activity-in this example, 1.8 kHz. The bandwidth of the response area that was measured at 10 and 20 dB above threshold is taken as the log of the difference between the highest and lowest stimulus frequency that evoked activity at that stimulus level. The spike time histogram is a temporal representation of the onset activity of all the spikes occurring within the neuronal receptive field space. The minimum onset latency for this example is around 11 ms. The rate-level function measured at CF shows an asymptotic monotonic behavior for this monaural stimulation example. Rate-level profiles were constructed by summing for each signal level the spike counts produced by the CF and the two frequencies nearest CF. The peak firing response occurring at 57.5 dB stimulus intensity marks the "transition point" of the neuron's response firing. The transition point separates the fast-growing, low-level slope from the slower-growing, saturating, or decreasing slope.

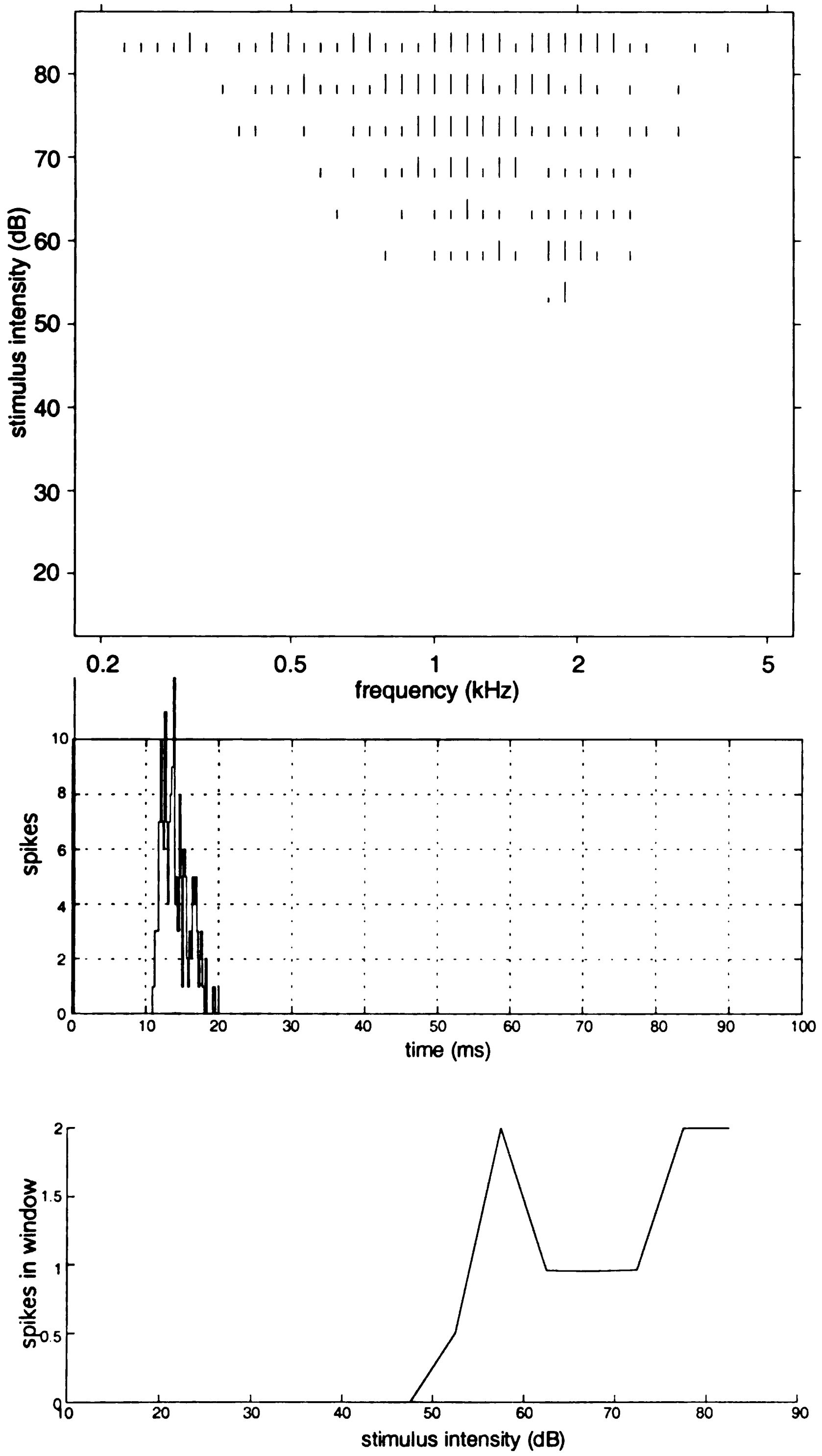
Figure 1.1B depicts the spectral tuning for monaural tonal stimulation in the presence of background noise for the same single neuron shown in figure 1.1A. The FRA depicts two obvious changes that occurred with the addition of background noise in the contralateral stimulation condition: 1) the threshold was elevated to 67.5 dB SPL; and 2) the response strength dramatically decreased. The bandwidth at 10 dB above threshold appears to be narrower with the addition of the continuous noise, although with the reduction in response strength, the patchiness of the receptive field space makes this comparison with the no-noise monaural condition more difficult. The spike time histogram reveals a minimum onset latency of around 15 ms, suggesting that the addition of background noise to the contralaterally-presented tone stimulus acts to prolong stimulus response times. A rate-

Figure 1.1 Frequency response area (FRA) for four different stimulation conditions

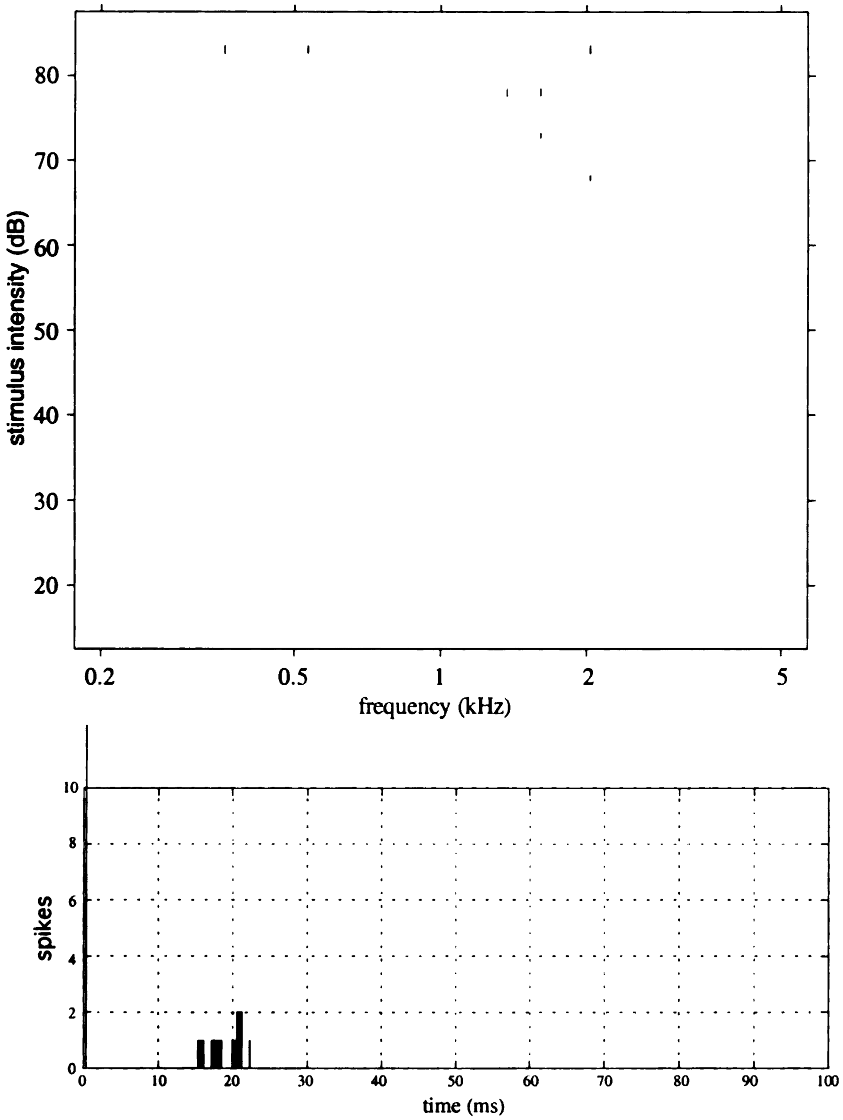
The FRA recorded from a single unit for four different stimulation conditions are displayed in **A** (monaural, no-noise), **B** (monaural, with noise), **C** (binaural, no-noise), **D** (binaural, with noise). Each vertical line within the FRA represents the spike response strength to a tone presented at that given frequency/intensity position. The minimum threshold is the lowest stimulus level that evokes activity in the FRA. The CF is measured as the stimulus frequency with the lowest stimulus intensity necessary to evoke neuronal activity. The bandwidth of the FRA that was measured at 10 and 20 dB above threshold was taken as the log of the difference between the highest and lowest stimulus frequency that evoked activity at that stimulus level. A spike time histogram and rate-level profile, which are displayed for each of the stimulus conditions (except for **B** due to limited responses), were extracted from spike activity recorded at and near the CF of the FRA. In the monaural stimulation conditions (**A** and **B**), the addition of noise resulted in the elevation of threshold, reduced response strength, and delayed onset latency of response at CF. In the binaural stimulation conditions (**C** and **D**), the addition of noise resulted in the elevation of threshold, reduced response strength, and narrowing of bandwidth.

Monaural stimulation condition - 556u12c

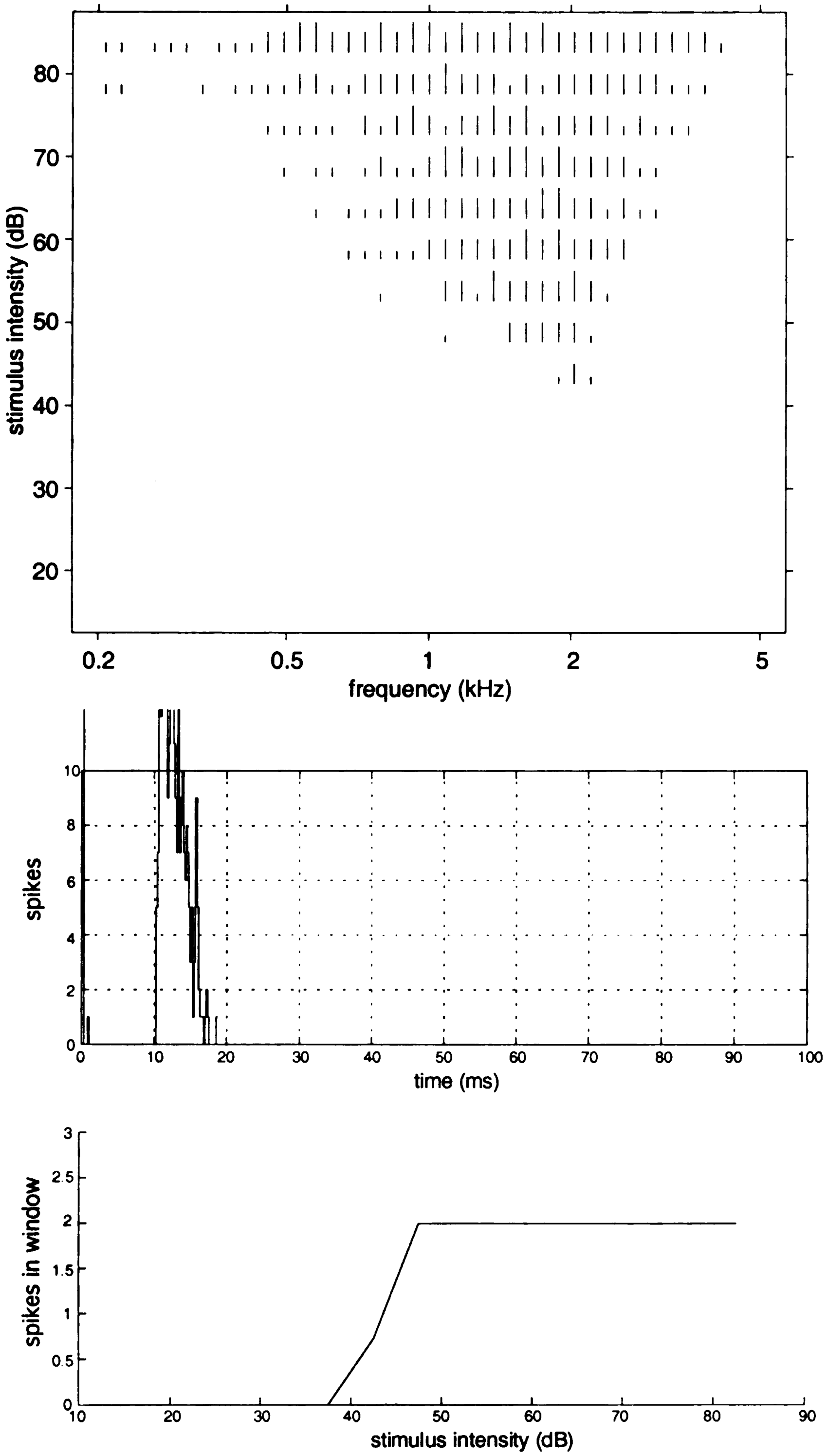
A

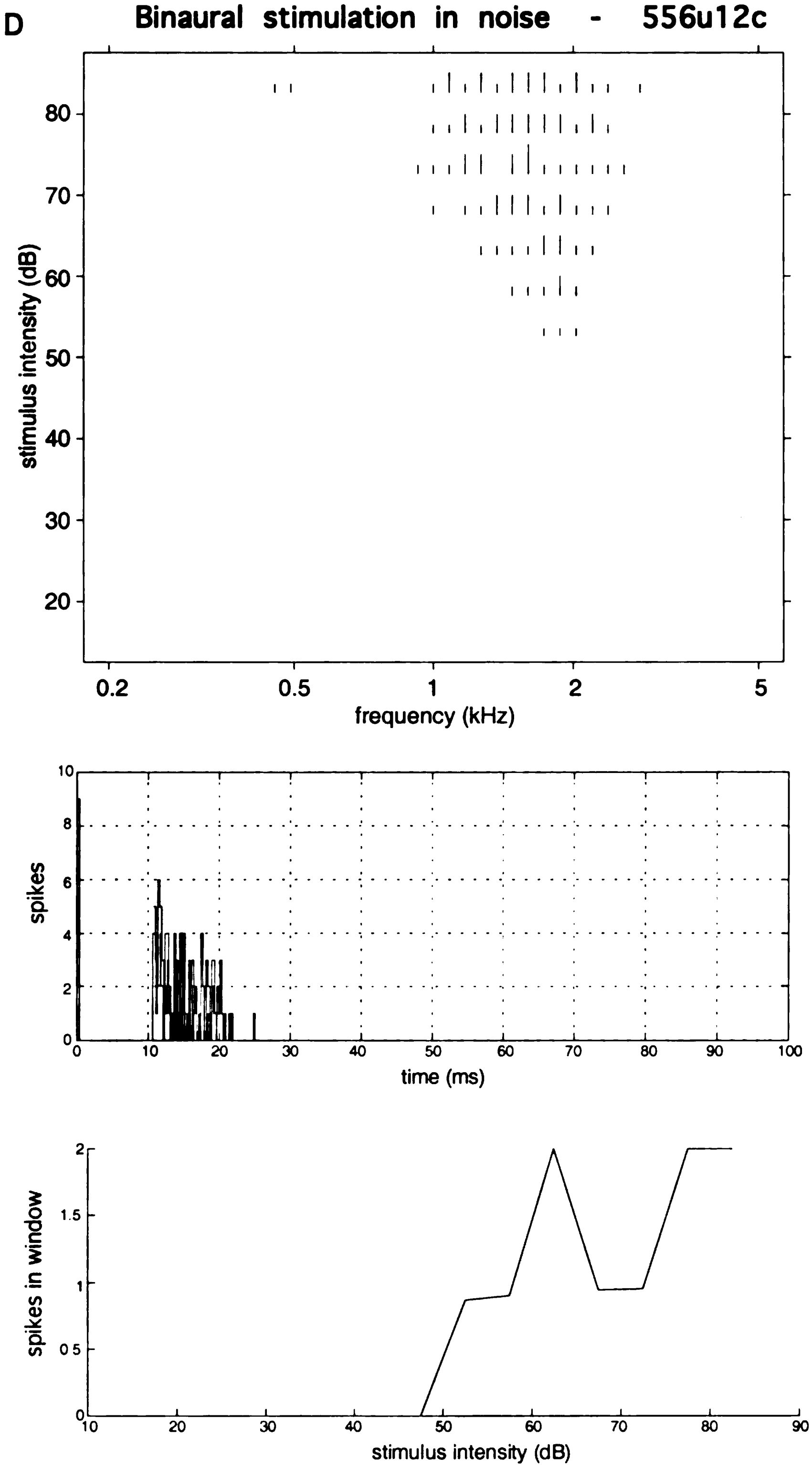


B Monaural stimulation with noise - 556u12c



C Binaural stimulation condition - 556u12c




D
Binaural stimulation in noise - 556u12c
stimulus intensity (dB)
20
30
40
50
60
70
80
frequency (kHz)
0.2
0.5
1
2
5
spikes
0
2
4
6
8
10
time (ms)
0
10
20
30
40
50
60
70
80
90
100
spikes in window
0
0.5
1
1.5
2
stimulus intensity (dB)
10
20
30
40
50
60
70
80
90

level function could not be constructed due to the reduction in spike activity for this stimulus condition.

The effect of binaural stimulation on the spectral-temporal tuning curves can be seen in figure 1.1C, which depicts the response parameters for binaural tone stimulation without background noise. The signals to the two ears were presented at the same intensity and without time-delay or phase difference ("in-phase"). Compared to the response characteristics measured in the monaural condition seen in figure 1.1A, three noticeable differences are observed: 1) the response strength was increased as seen by the more robust spike activity depicted in the frequency response area and the spike histogram; 2) the threshold at CF was lower by 10 dB; and 3) the transition point of firing occurred at a stimulus intensity that was 10 dB lower than in the monaural condition. Although the transition point of the rate-level function shifted down by 10 dB, the monotonicity of the unit remained the same. The minimum onset latency measured in the binaural stimulation condition was also relatively the same as in the monaural condition occurring roughly around 10 ms. As a result of the data inclusion criterion described at the beginning of this section, this cell would be classified as an EE neuron.

The effect of background noise on the spectral tuning of binaural tonal stimulation is seen in figure 1.1D. Compared to the response parameters measured for binaural tonal stimulation without background noise that was depicted in figure 1.1C, three obvious changes are noticed: 1) spike response magnitudes decreased with the addition of binaural noise, as illustrated by the depressed amplitude in the spike histogram; 2) tuning bandwidth measured at either 10 or 20 dB above threshold was narrowed; and 3) the stimulus intensity at which the transition point occurred was shifted up by 15 dB. Despite the shift in transition point seen in the rate-level function, the neuron still behaved monotonically in the presence of binaural background noise. The minimum onset latency was only modestly lengthened (11.3 ms) with the addition of background noise.

To gain an overview of the main effects of monaural vs. binaural receptive field properties with and without background noise, general averages of response parameters for all units were obtained. The following group statistics were compiled from parameters extracted from the frequency response area that were measured for each stimulation condition. The original goal was to study the neural behavior of low-frequency cells with and without background noise via monaural and binaural stimulation. What was actually achieved was a range of CFs spanning 0.5 to 9.5 kHz with 64% of the units being between 0.5 to 2.5 kHz (figure 1.2A). The correlation of CF measured from monaurally- and binaurally-derived frequency response areas matched per unit was high (R=0.93) indicating a basic similarity in the best frequency receptive field parameter (figure 1.2B). Due to the similarity of CF measured in either monaurally- or binaurally-derived FRAs, the use of CF as a variable for comparisons with other receptive field parameters will be accomplished using the CF measured in the monaural condition. The background noise was selected to mask tones presented at 20 dB above unit threshold in the monaural condition using restrictive audiovisual criteria. Figure 1.3 shows the distribution of the noise levels required to achieve measured threshold shifts for both monaural and binaural conditions. Although there was a wide range of noise levels used to elevated thresholds as measured objectively with FRAs, the absolute noise level used to invoke a threshold shift in the monaural stimulation condition was not correlated with the amount of threshold shift actually achieved.

To determine whether significant changes in receptive field response properties within an individual cell occurred due to a change in stimulus condition, a within-subjects ANOVA (repeated measures) was used in the following analysis to test the null hypothesis that there was no difference in response parameter. Table 1.1A compares the mean values for receptive field response properties measured in the monaural and binaural stimulation conditions without background noise. The standard deviations calculated for each mean value in this table and the following three tables are indicated in parentheses. No

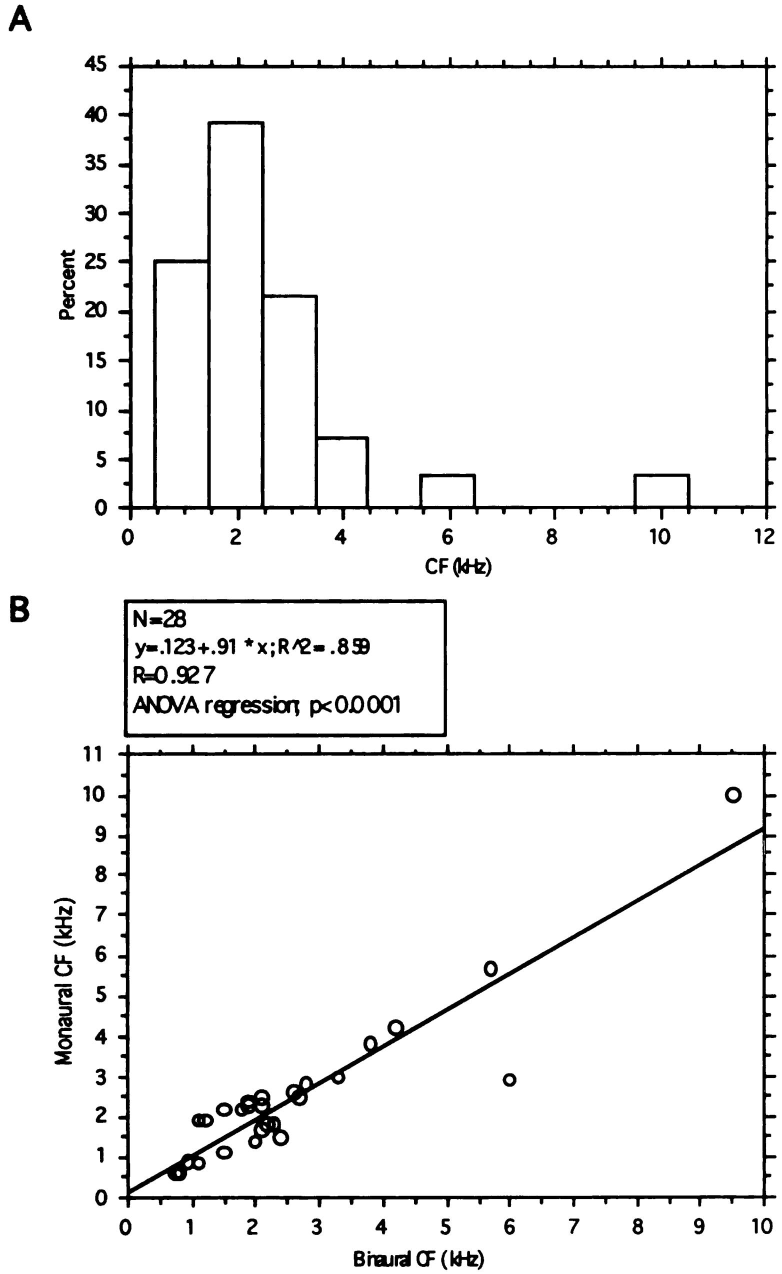


Figure 1.2 Characteristic Frequency (CF) of single units recorded from AI

In **A**, a histogram of the percentage of CFs represented in the sampled population revealed that the majority (64%) of the sampled single units had CFs below 2.5 kHz. In **B**, correlation of CF from FRAs recorded in monaural and binaural stimulation conditions revealed a significantly high correlation (R=.927; p<.0001) in the CFs measured in either stimulus condition.

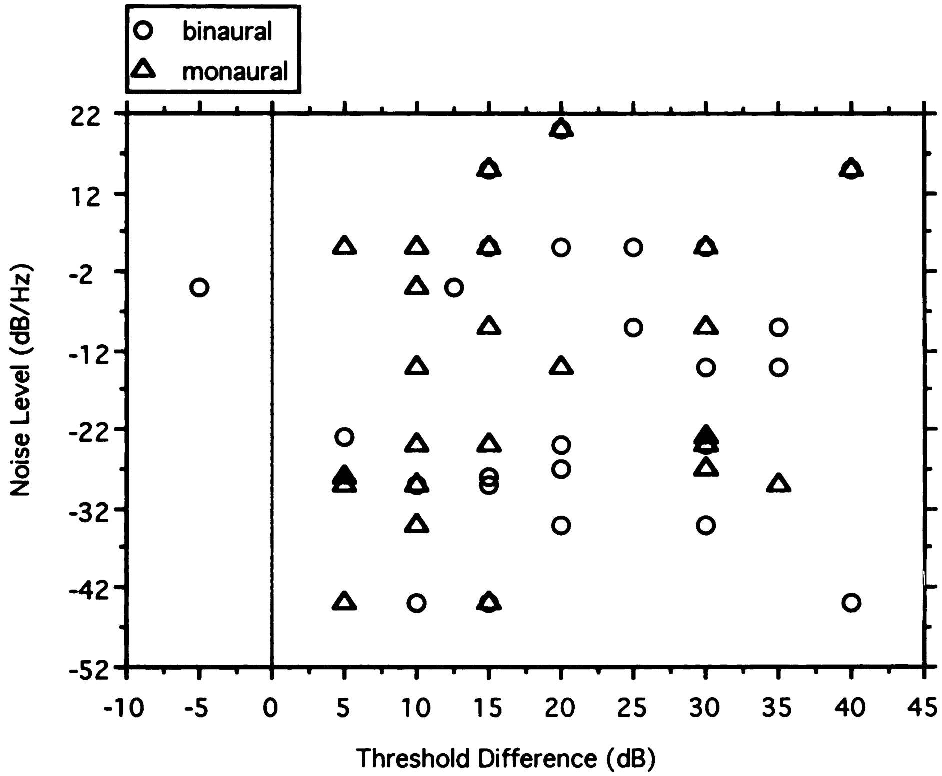


Figure 1.3 Distribution of Masking Noise Levels

The distribution of noise levels, which were used to elicit a 20 dB threshold shift, is plotted as a function of threshold difference (dB) which measures threshold before and after the addition of noise. The absolute noise level used to invoke a threshold shift in the monaural stimulation condition was not correlated with the amount of threshold shift actually achieved.

statistically significant difference in threshold, bandwidth, or onset latency was found in this comparison. Table 1.1B displays the mean values for receptive field response properties measured in the monaural stimulus condition with and without background noise. A significant increase in mean threshold was found with the addition of background noise (ANOVA repeated measures, $p < 0.0001$). In addition, minimum onset latency was found to be significantly prolonged by a mean value of 3 ms with the addition of noise (ANOVA repeated measures, $p < 0.0001$). No significant difference was found for mean bandwidth. Table 1.1C shows the mean values for receptive field response properties measured in the binaural stimulus condition with and without background noise. As in the monaural condition, a significant increase in mean threshold was found with the addition of background noise (ANOVA repeated measures, $p < 0.0001$). Likewise, minimum onset latency was found to be significantly prolonged by a mean value of 3 ms with the addition of noise in the binaural condition (ANOVA repeated measures, $p < 0.5$). Again, no significant difference was observed for mean bandwidth. Table 1.1D shows the mean values for the calculated difference in receptive field response properties for both monaural and binaural stimulation. The difference was calculated as the change in response parameter when noise was added to tonal stimulation. According to table 1.1D, there was no statistical difference in the noise effect on frequency response properties for either monaural or binaural stimulation. While the global analysis shows only some significant latency effects with the addition of noise, a more detailed analysis, presented in the following section, reveals some additional effects that were obscured by averaging over many neurons.

1.2.2 Threshold

The distribution of FRA threshold, which is plotted in figure 1.4A as a function of CF, shows that among the 28 single units sampled, mid- to high-threshold units (> 15 dB SPL) were found for the majority of the sample while a few low- threshold units were

Table 1.1A Means table for receptive field parameters measured from monaural and binaural tonal stimulation.

		Monaural	Binaural
Threshold (dB)	N=28	31.1 (11.7)	27.6 (12)
Bandwidth 10 (octaves)	N=28	0.67 (0.5)	0.61 (0.35)
Bandwidth 20 (octaves)	N=28	1.0 (0.7)	0.89 (0.41)
Bandwidth 30 (octaves)	N=23	1.4 (0.93)	1.1 (0.64)
Onset Latency (ms)	N=28	15.4 (5.3)	14.4 (3.3)

Table 1.1B Means table for receptive parameters measured from monaural tonal stimulation with and without background noise.

		Monaural	Monaural w/ Noise
Threshold (dB)	N=28	31.143 (12)	48.107 (8.5) ***
Bandwidth 10	N=27	0.679 (0.53)	0.669 (0.43)
Bandwidth 20	N=23	1.054 (0.69)	1.052 (0.71)
Onset Latency	N=28	15.442 (5.3)	18.752 (5.1) ***

Table 1.1C Means table for receptive parameters measured from binaural tonal stimulation with and without background noise.

		Binaural	Binaural w/ Noise
Threshold (dB)	N=28	27.571 (12)	48.018 (9.4) ***
Bandwidth 10	N=27	0.588 (0.34)	0.732 (0.48)
Bandwidth 20	N=18	0.856 (0.39)	1.190 (0.82)
Onset Latency	N=25	14.898 (3.2)	17.866 (6.0) *

Table 1.1D Mean difference table for receptive parameters measured from monaural and binaural stimulation conditions.

		Monaural Difference	Binaural Difference
Threshold (dB)	N=28	16.964 (11)	20.446 (11)
Bandwidth 10	N=23	-0.074 (0.63)	0.144 (0.46)
Bandwidth 20	N=14	0.127 (0.84)	0.420 (0.77)
Onset Latency	N=25	3.428 (3.0)	2.967 (5.4)

level of significance

* < 0.05

*** < 0.0001

obtained around 2 to 4 kHz. Labeling of unit types (e.g., high-threshold) are based upon the contralateral tonal stimulation condition, to make comparisons of responses with previous studies consistent. Regression analysis of the mean threshold values, which are plotted in figure 1.4B, showed a positive correlation between monaural and binaural thresholds. As seen by the unity line (dotted line) in figure 1.4B, the majority of the units showed higher thresholds in the monaural condition (above the dotted line) while only 28% of the units showed higher thresholds in the binaural condition (below the dotted line). On average, binaural thresholds were 3.5 dB lower than monaural thresholds (see table 1.1A).

Although the targeted threshold shift for these experiments was 20 dB with the addition of background noise in the monaural condition, the range of threshold difference, which is plotted in figure 1.5A as a function of CF, that was actually achieved spanned 5 to 40 dB. The horizontal dotted line represents the measurement probability error (~8 dB) for threshold shifts per unit. The majority of the units were above this measurement probability error line. The mean threshold increase with the addition of background noise was 17 dB for the monaural condition and 20 dB for the binaural condition (see Table 1.1B & 1.1C). Figure 1.5B shows a scatterplot of the distributions of threshold difference measurements in monaural and binaural stimulation conditions. Regression analysis of the noise-induced threshold shifts revealed a positive correlation between monaural and binaural threshold shift. The diagonal dotted line represents the unity line. More than half of the units showed a tendency for larger threshold shifts with binaural stimulation (below the unity line), while 21% of the units showed no difference in threshold shift between either of the stimulation conditions.

The effect of background noise on FRA threshold, seen in figure 1.6, was observed by plotting the amount of threshold difference as a function of initial FRA threshold measured in the absence of background noise. Regression analysis of the effect of background noise on FRA threshold revealed a significant negative correlation observed in both monaural and binaural stimulation conditions. The amount of variability in the regression was similar

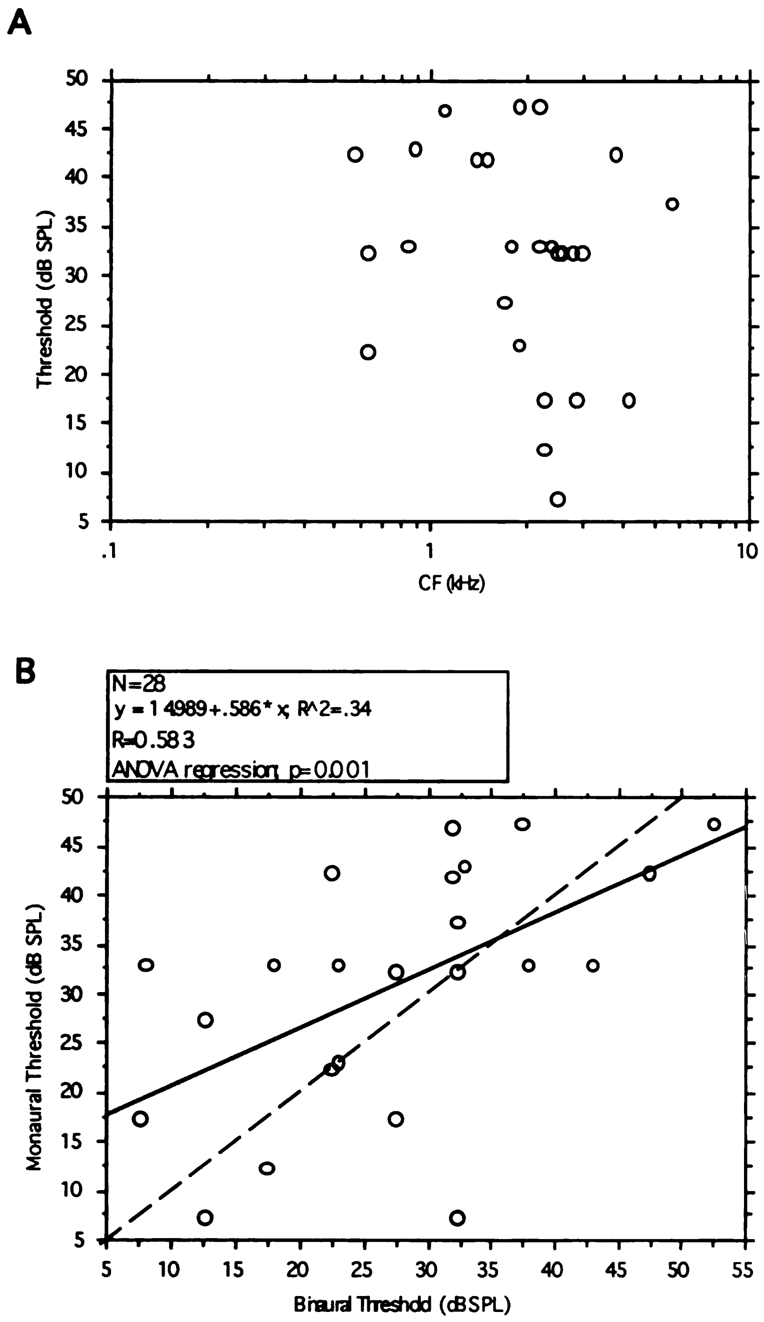


Figure 1.4 FRA thresholds from monaural and binaural stimulation conditions

In **A**, the distribution of FRA thresholds from monaural recordings is plotted as a function of CF. The distribution of monaural thresholds revealed mid-to high-thresholds for the majority of the FRAs recorded, although a few low-thresholds were present in units with CFs between 2 and 4 kHz. In **B**, the distribution of monaural FRA thresholds is plotted as a function of binaural FRA thresholds. Regression correlation of FRA thresholds revealed a significant correlation (R=.58; p=.001) between monaural and binaural conditions.

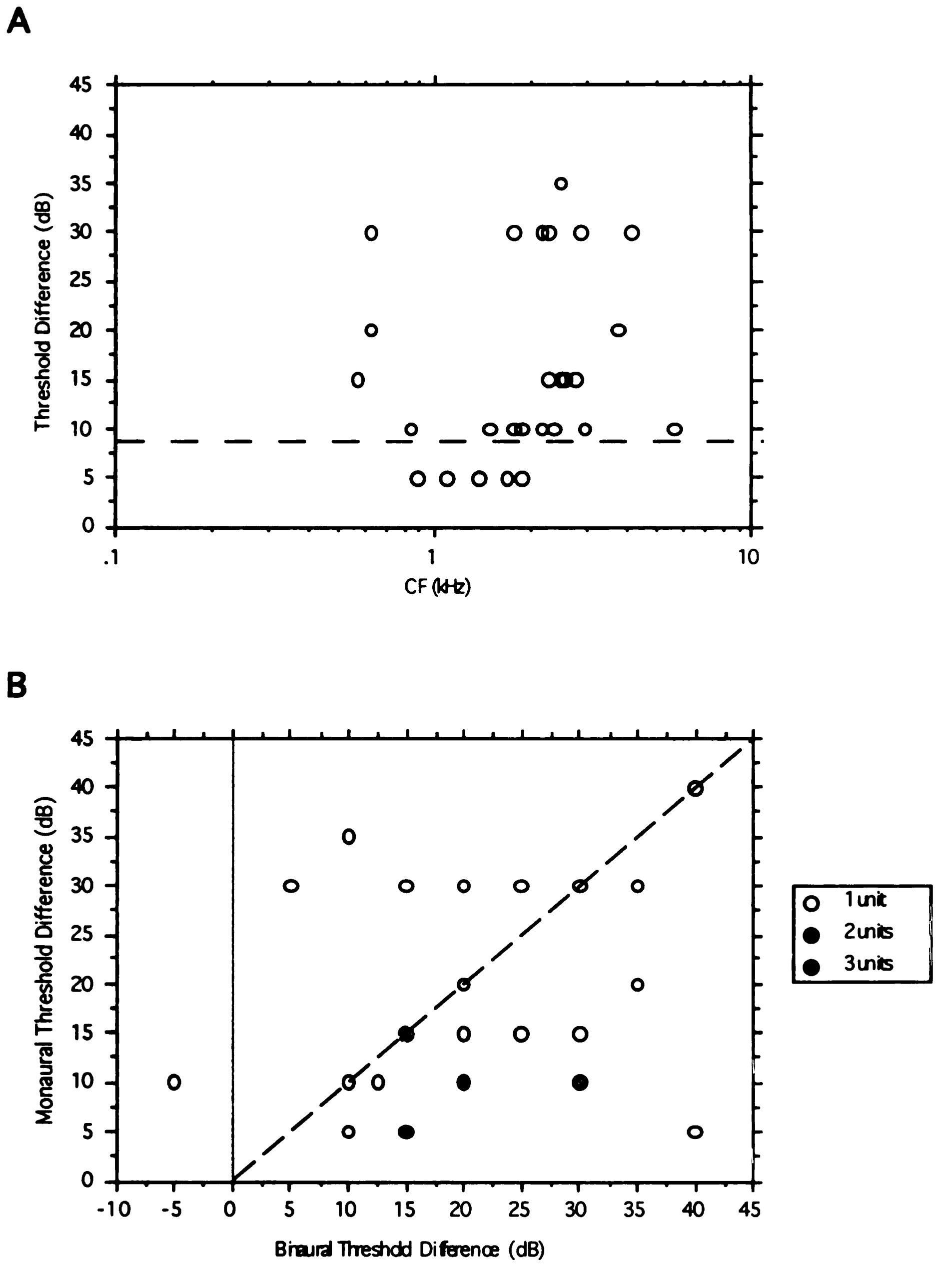


Figure 1.5 FRA threshold difference in monaural and binaural stimulation conditions

In **A**, the distribution of threshold difference, which measures the amount of threshold shift before and after the addition of continuous background noise, is plotted as a function of CF. The majority of the threshold shifts were above the measurement probability error (~ 8 dB) which is represented by the horizontal dotted line. The distribution of threshold differences were spread evenly across CF. In **B**, monaural threshold difference is plotted as a function of binaural threshold difference. The diagonal dotted line represents the unity line between monaural and binaural threshold difference. Monaural threshold difference was not correlated with binaural threshold difference. More than half of the units showed a tendency for larger threshold shifts with binaural stimulation, while 21% of the units showed no difference in threshold shift between either of the stimulation conditions.

between monaural and binaural stimulation conditions. The negative correlation suggests that low-threshold units had larger threshold shifts than high-threshold units. This applies, irrespective of whether stimuli are presented to one or both ears. It is quite possible, though, that this correlation is biased by the data inclusion criterion used in these analyses. That is, due to constraints in the dynamic range of the stimulus generation program (which was 72 dB), large threshold shifts in high-threshold units could have been mistaken as being a 'no response' when in fact the threshold of the unit could have been beyond the limits of the tested stimulus range.

1.2.3 Bandwidth

An important informational parameter of neural processing of signals which would be missed through the recording of rate-level functions is the spectral extent of receptive field parameter expressed as bandwidth. The distribution of bandwidth measurements made at 10 dB and 20 dB above threshold were plotted as a function of CF in figures 1.7A and 1.7C, respectively. The range of bandwidth measured at both intensity levels in the contralateral stimulation condition extended from a quarter of an octave to over two octave bands. According to bandwidth measurements made at 20 dB above threshold, 55% of the units were narrowly tuned (bandwidth of one octave or less) while the other 45% were broadly tuned (bandwidth larger than one octave). In fact, figure 1.7C reveals two clusters of bandwidth, one around 0.5 octaves and another around 1.8 octaves with only a few intermediate values. The scatterplots for monaural and binaural bandwidth 10 dB and 20 dB, seen in figures 1.7B and 1.7D, respectively, show a wide distribution of values. Regression analysis of bandwidth revealed no significant correlation between monaural and binaural stimulation conditions. In figure 1.7D, on the other hand, a positive correlation between monaural and binaural stimulation was found for the bandwidth 20 measure. The variance in the regression revealed that approximately 50% of the units had wider bandwidths in the monaural condition. The mean bandwidth at 10 dB above threshold for

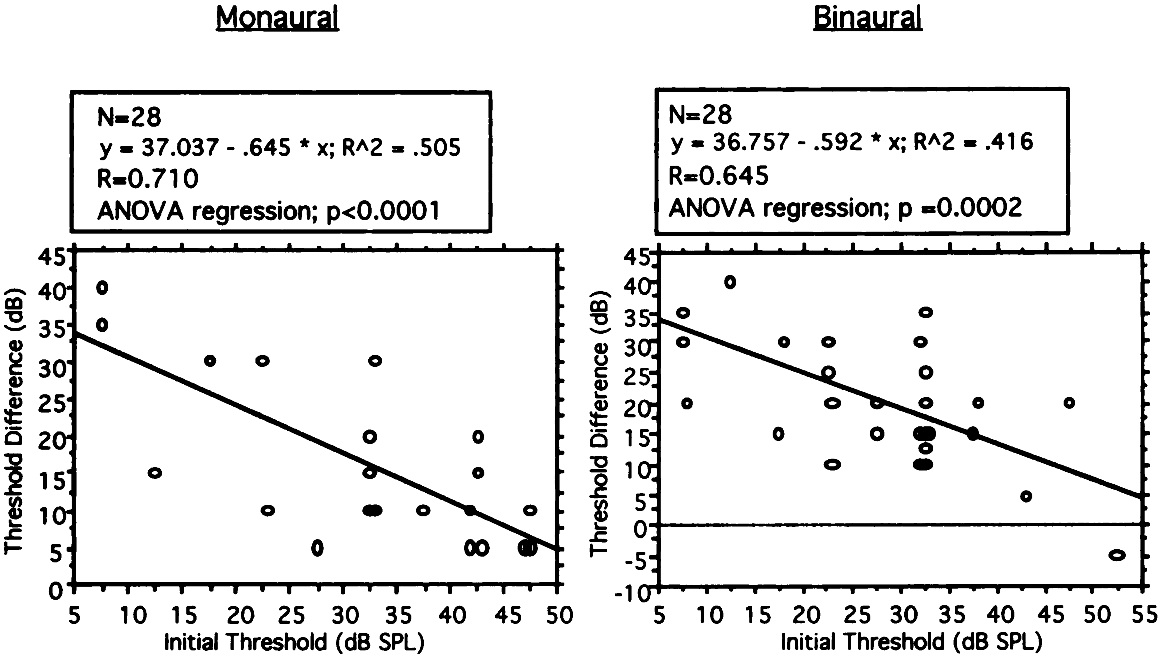


Figure 1.6 Background noise effect on FRA threshold in monaural and binaural conditions

The effect of background noise on FRA threshold is represented by a plot of the threshold difference as a function of initial threshold for both monaural and binaural stimulation conditions. Regression analysis revealed a significant negative correlation between threshold difference and initial threshold for both conditions. The amount of threshold shift decreases as a function of increasing initial threshold.

the monaural condition was 0.67 octaves and at 20 dB above was 1.0 octaves while for the binaural condition, bandwidth at 10 dB above was 0.61 octaves and at 20 dB above was 0.89 octaves (see table 1.1A). Therefore, there was no significant difference in mean bandwidth either between monaural and binaural tonal conditions nor between noise and no-noise conditions in either monaural or binaural conditions.

The effect of background noise on bandwidth is depicted in figures 1.8A-D. The addition of background noise resulted, for the majority of neurons (N=19, 73%) in a change of bandwidth. Interestingly, the change was about equally distributed between expansion and contraction of the receptive field bandwidth. The distributions of bandwidth difference, which measures the change in bandwidth before and after the addition of background noise, are plotted in figures 1.8A (for bandwidth at 10 dB above threshold) and figure 1.8B (for bandwidth at 20 dB above threshold) as a function of CF. Accordingly, the distributions of bandwidth 10 and bandwidth 20 difference were equally spread around zero across the CF axis. As a check to ensure that changes in bandwidth were true experimental findings, a measurement probability error (indicated in the figures by a dotted line) was calculated from bandwidth measurements of frequency response areas that were recorded twice per unit (N=31). The measurement probability error was defined as the mean absolute bandwidth difference between twice-recorded frequency response areas. The excursions from zero that are beyond the measurement probability error (the dotted line at +/- 0.4 octaves) represent bandwidth changes per unit that are undoubtedly attributable to the addition of background noise. Regression analysis of bandwidth difference measures, seen in figure 1.8B, revealed no significant correlation in bandwidth 10 difference (upon the addition of background noise) between monaural and binaural stimulation conditions. Bandwidth 20 difference measures, on the other hand as seen in figure 1.8D, were positively correlated between monaural and binaural stimulation conditions. The percentage of units above and below the unity line were roughly the same. The variance in the data suggests that there is no substantial difference in the effect of noise

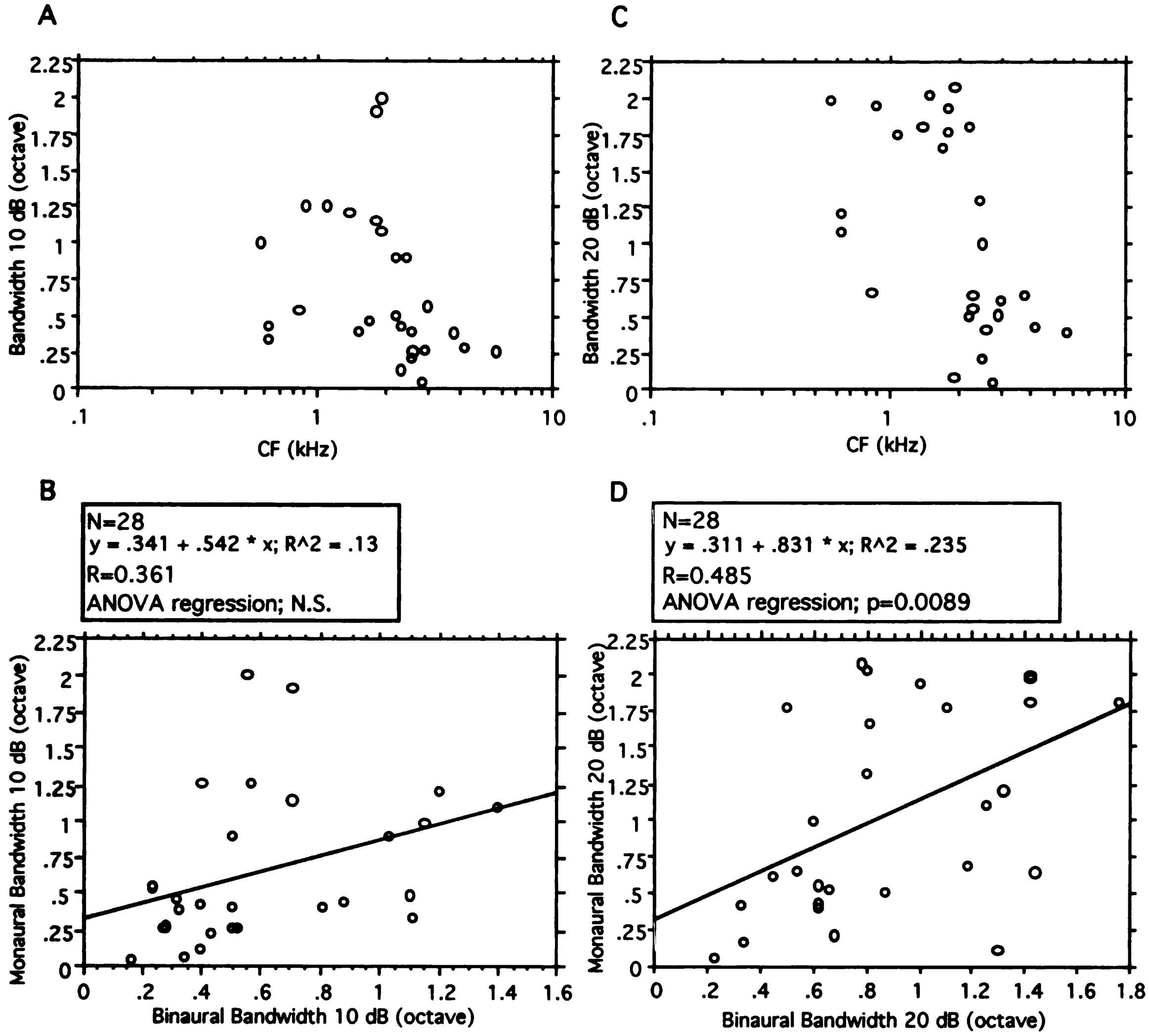


Figure 1.7 Distribution of bandwidth in monaural and binaural stimulation conditions

The distributions of bandwidth (in octaves) measured at 10 dB above threshold are depicted in **A** and **B**, while those for bandwidth measured at 20 dB above threshold are depicted in **C** and **D**. In **A** and **C**, the distributions of monaural FRA bandwidth were evenly distributed across CF. In **B** and **D**, monaural bandwidth is plotted as a function of binaural bandwidth. In **D**, a significant positive correlation (R=.485; p=.009) was found between monaural and binaural bandwidth at 20 dB. Furthermore, binaural bandwidth tended to be wider than monaural bandwidth at 20 dB above threshold.

on bandwidth in frequency response areas recorded in either monaural or binaural tonal conditions. The mean bandwidth difference at 10 dB above threshold was -0.074 octaves for monaural stimulation and 0.144 octaves for binaural stimulation. At 20 dB above threshold, mean bandwidth difference was 0.127 octaves for monaural stimulation and 0.420 for binaural stimulation (see table 1.1D).

The effect of background noise on receptive field bandwidth was further studied by correlating the bandwidth difference with each unit's initial bandwidth prior to introducing the background noise stimuli. The bandwidth difference at 10 dB or 20 dB above threshold was plotted as a function of initial bandwidth for the monaural condition seen in figures 1.9A and 1.9B and for the binaural condition seen in figures 1.9C and 1.9D. In the monaural stimulation condition, regression correlation of bandwidth revealed a significant negative correlation between bandwidth difference and the initial tone-evoked response bandwidth at both 10 dB and 20 dB above threshold. The variance in the bandwidth difference revealed an inverse relationship in bandwidth when background noise was present during tone detection suggesting that neurons whose spectral tuning is narrow near threshold tend to have broader tuning when continuous noise is present in the background while neurons whose spectral tuning is broad near threshold tend to have narrower tuning when continuous noise is present. This trend was only weakly observed in the binaural stimulation condition for bandwidth at 10 dB above threshold (see figure 1.9C) and was not present for binaural bandwidth at 20 dB above threshold (see figure 1.9D).

1.2.4 Onset Latency

By recording the onset latencies of the unit's response to each frequency-level combination, either a spike-time histogram or latency-level function can be produced to display the minimum onset latency response to tonal stimuli. Minimum onset latencies recorded in cat AI are typically on the order of tens of milliseconds. The distribution of minimum onset latency, which was plotted as a function of CF and shown in figure 1.10A,

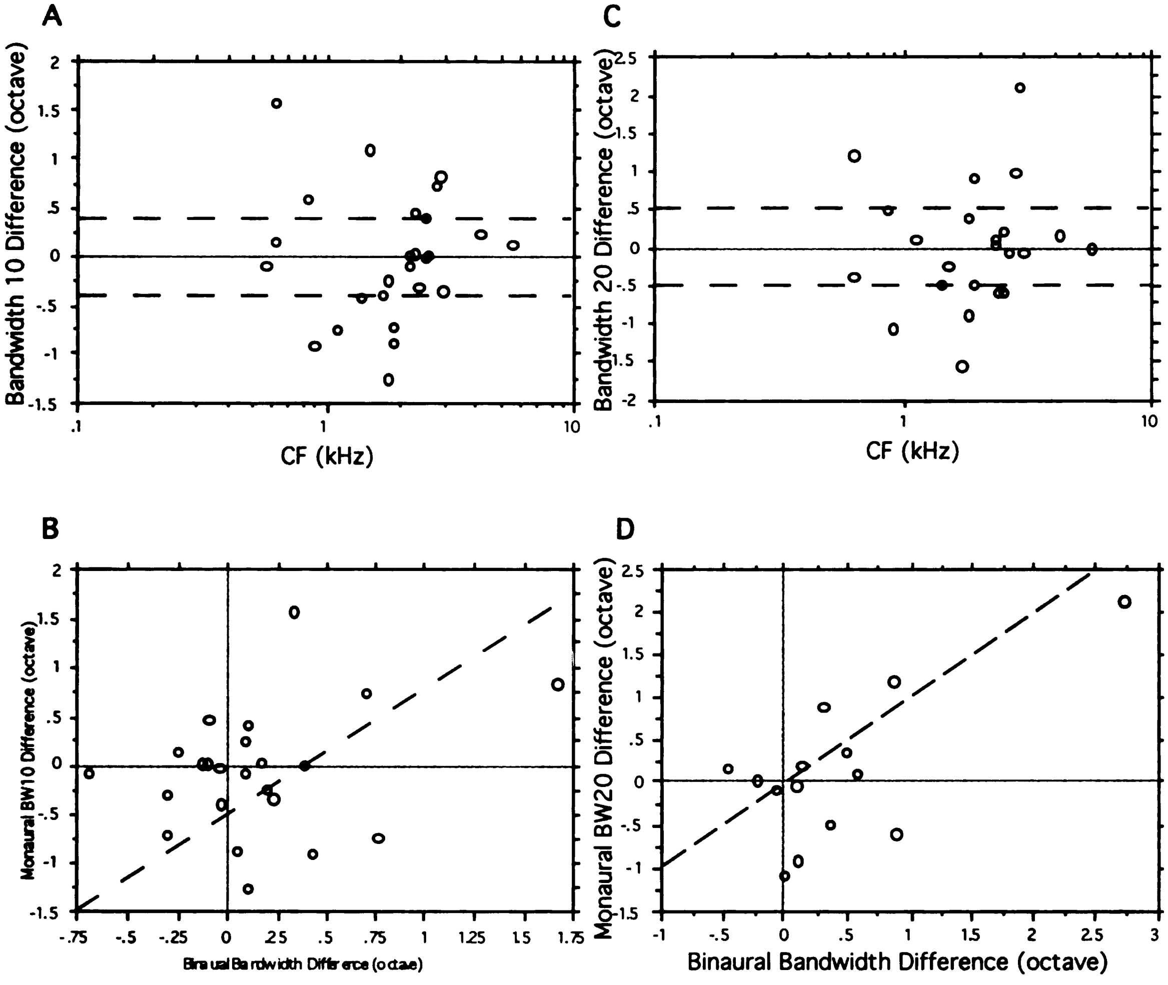


Figure 1.8 Distribution of bandwidth difference in monaural and binaural conditions

The distribution of FRA bandwidth difference (in octaves) measured at 10 dB above threshold are depicted in **A** and **B**, while those for bandwidth difference measured at 20 dB above threshold are depicted in **C** and **D**. In **A** and **C**, the distributions of bandwidth difference are plotted as a function of CF. In both figures, excursions from zero that are beyond the measurement probability error (the dotted line at +/-0.4 octaves) represent bandwidth changes per unit. Bandwidth differences were evenly distributed across CF in either monaural or binaural stimulus conditions. In **B** and **D**, monaural bandwidth is plotted as a function of binaural bandwidth. The diagonal line in either plot represents the unity line.

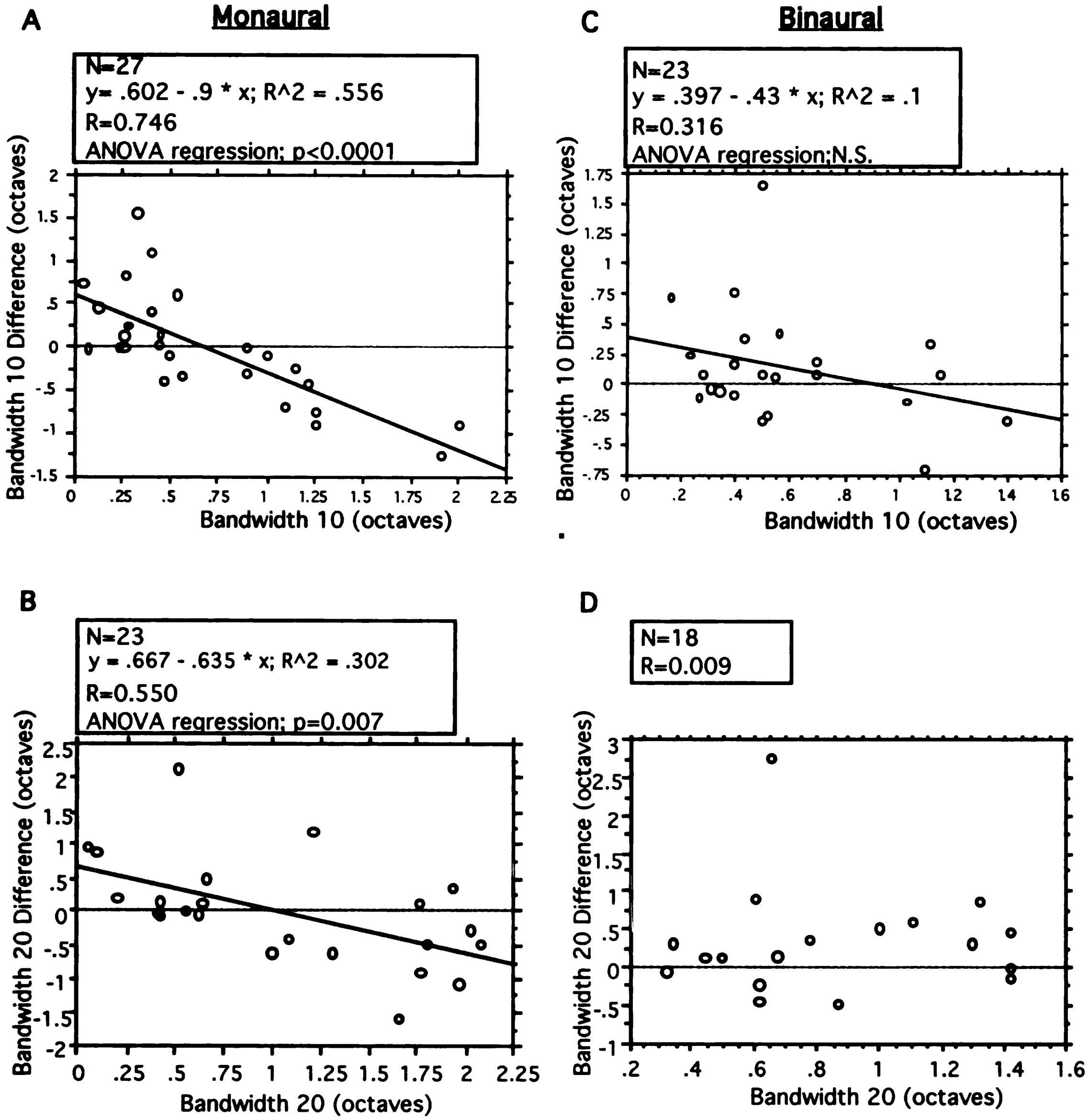


Figure 1.9 Background noise effects on bandwidth for monaural and binaural conditions

The effects of background noise on FRA bandwidth measured at 10 dB and 20 dB above threshold were depicted for the monaural stimulus conditions (**A** and **B**) and for the binaural conditions (**C** and **D**). In **A**, the bandwidth 10 difference was significantly and negatively correlated with the initial bandwidth 10 (R=-.746; p<.0001). In **B**, the bandwidth 20 difference was also significantly and negatively correlation with the initial bandwidth 20 (R=-.550; p=.007). In **C**, although the correlation was not significant, a negative trend was evident. The negative correlations revealed are indicative of an inverse relation of bandwidth with the addition of continuous background noise. The significance of the background noise effect on bandwidth was more evident in the monaural stimulus conditions.

revealed a majority of units responding within the first 20 ms of the tone burst stimulus. The distributions of minimum onset latency for monaural and binaural stimulation conditions plotted in figure 1.10B revealed a positive correlation in minimum onset latency between monaural and binaural tonal stimulus conditions. The variance in minimum onset latency suggested that there were few differences in onset latency between either monaural or binaural conditions. The mean onset latency for either tonal stimulation condition was 14-15 ms (see table 1.1A). Therefore, in the temporal domain, with regards to onset latency, there appears to be no difference between monaural and binaural stimulation.

The effect of background noise on onset latency was measured by calculating the onset latency difference for each unit. According to the distribution of onset latency differences for monaural stimulation, seen in figure 1.11A, the majority of the units, irrespective of CF, displayed positive onset latency differences; the minimum onset latency measured was usually longer when background noise is present. A few units did show a negative onset latency difference, and in fact their minimum onset latencies became shorter in the presence of background noise. In figure 1.11B, which shows the distribution of monaural minimum onset latency difference plotted as a function of latency difference from the binaural stimulus conditions, no significant correlation was found due to two extreme difference values along the binaural dimension. These extreme effects of noise on binaural FRAs, which is demonstrated by the minimum onset latency, could be attributable to characteristics of low CF neurons. The mean latency difference was 3 ms. Since the range of latency was greater than 10 ms, the averaged difference showed no significant background noise effect on onset latency for either monaural or binaural tonal stimulation. To test whether there was any relationship between noise-induced onset latency changes and tone-evoked onset latency, regression analysis between the two was conducted for both the monaural and binaural stimulation conditions which is depicted in figure 1.12. The variance in onset latency difference did not show a significant relationship between

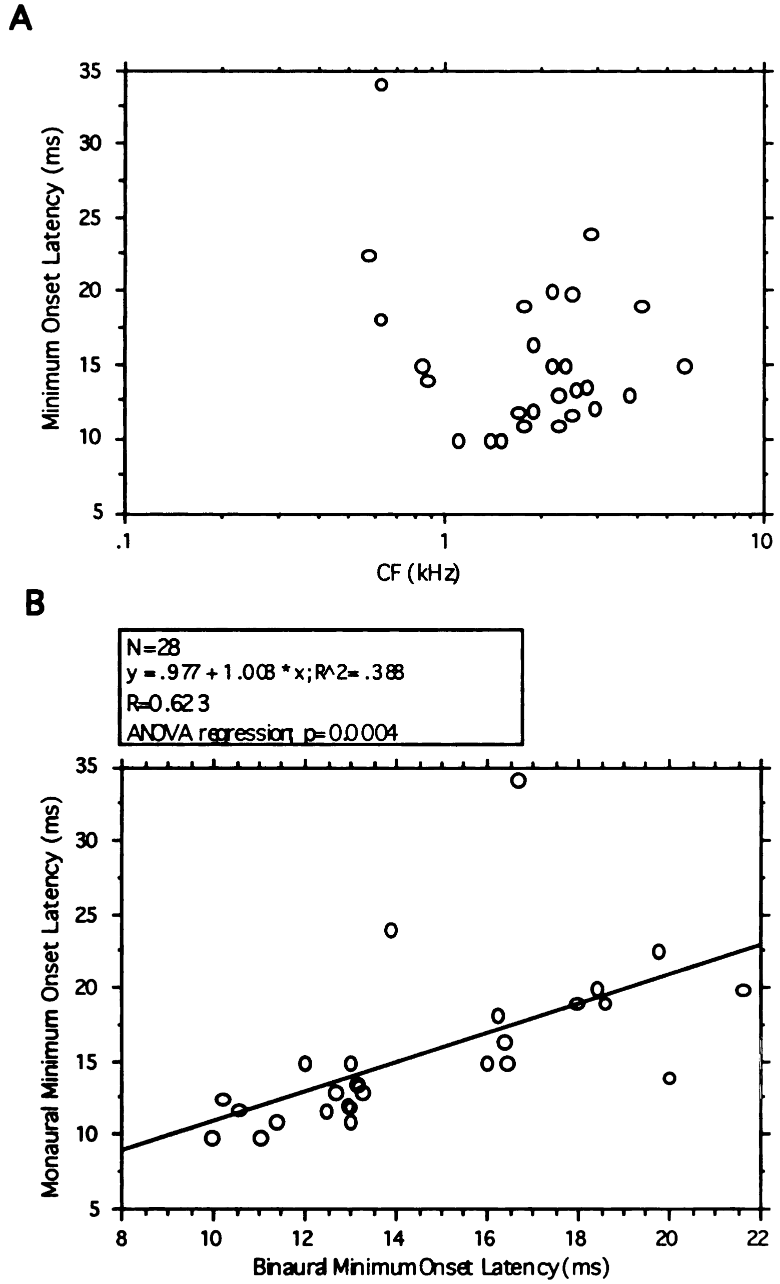


Figure 1.10 Distribution of minimum onset latency for monaural and binaural conditions

In **A**, the distribution of minimum onset latency measured from spike activity around the CF of the monaural FRA is plotted as a function of CF. The minimum onset latencies were widely distributed across CF. In **B**, the distribution of monaural minimum onset latency is plotted as a function of binaural minimum onset latency. Regression correlation of minimum onset latency revealed a significant positive correlation (R=.623; p=.0004) with a slope approximating unity, which indicates equivalency of onset latency irrespective of stimulus condition.

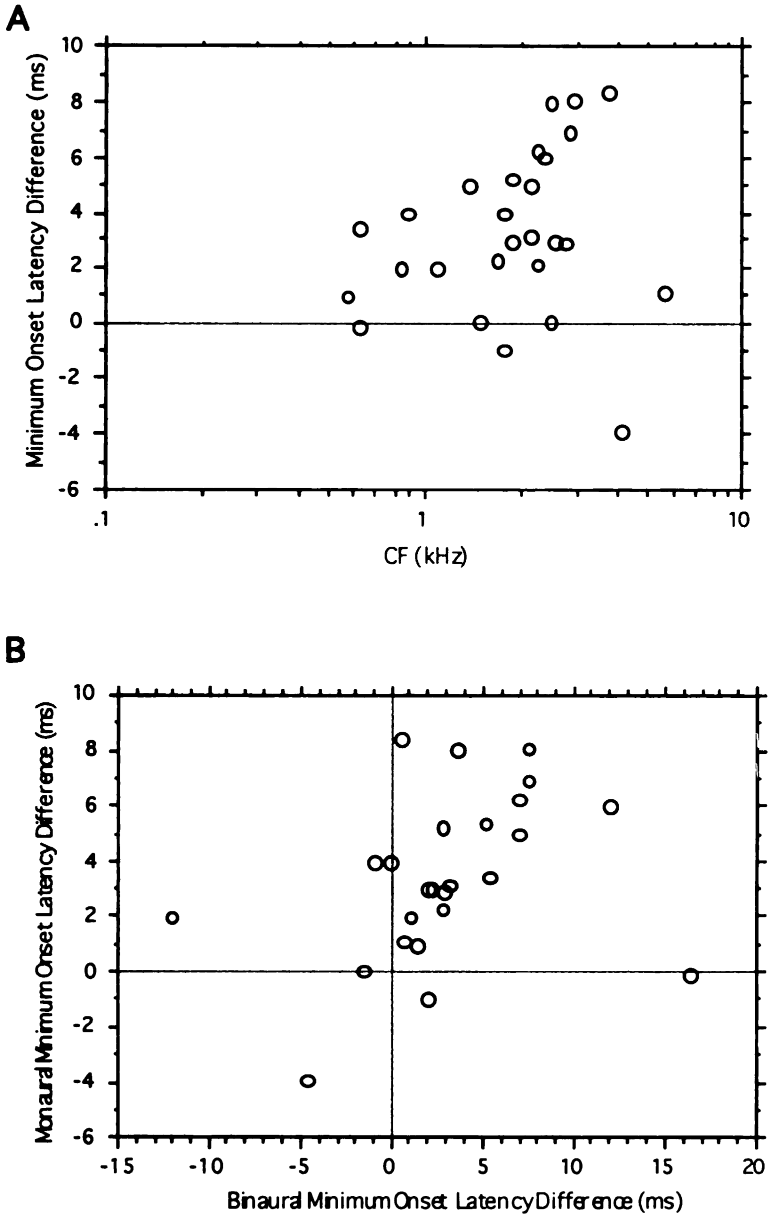


Figure 1.11 Distribution of onset latency differences in monaural and binaural conditions

In **A**, the distribution of minimum onset latency difference, which measures the amount of change in monaural onset latency before and after the addition of continuous background noise, is plotted as a function of CF. The distribution of onset latency differences was widely distributed across CF with the largest shifts in onset latency occurring at higher CF values. A positive onset latency difference indicates longer latencies in the presence of background noise. A few of the units exhibited shorter latencies in the presence of background noise. In **B**, the distribution of monaural minimum onset latency difference is plotted as a function of binaural minimum onset latency difference. Due to two extreme difference values in the binaural stimulation conditions, the regression correlation was not significant. The extreme effect of background noise on low CF (~1 kHz) minimum onset latency could not be ruled out.

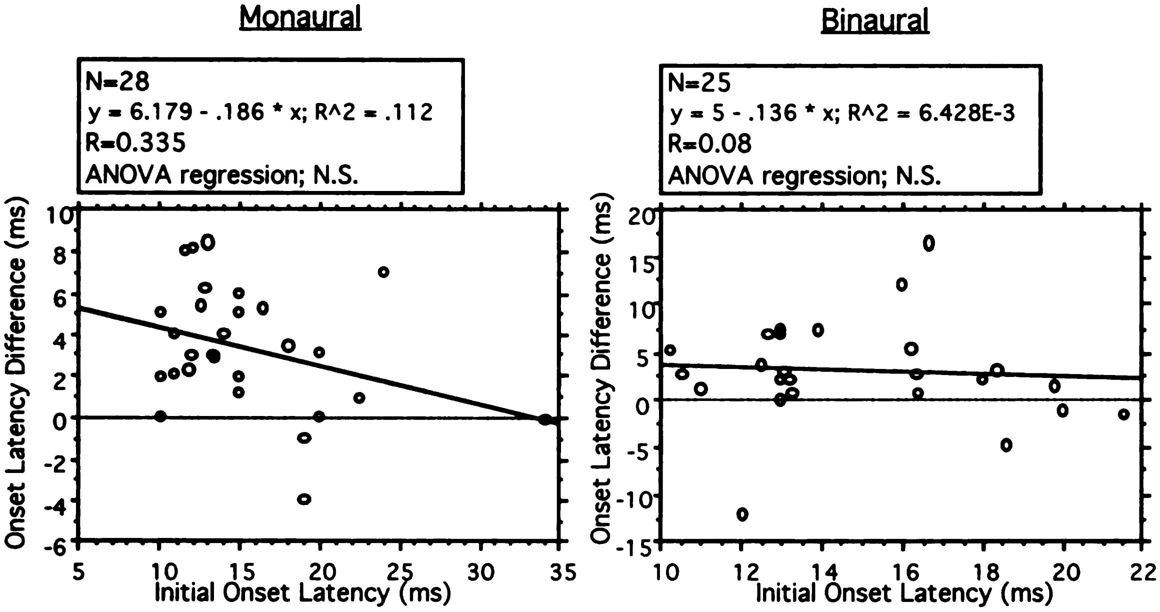


Figure 1.12 Background noise effects on FRA minimum onset latency for monaural and binaural stimulus conditions

The effect of background noise on FRA minimum onset latency is represented by a plot of the onset latency difference as a function of initial onset latency for both monaural and binuaral stimulation conditions. Although there appeared to be an inverse relation between onset latency difference and inital onset latency for the monaural condition, no significant regression correlation was found for either stimulation conditions.

onset latency difference and onset latency to tonal stimulation in either monaural or binaural stimulation conditions.

1.2.5 Binaural Configuration of Signal-in-Noise Stimuli

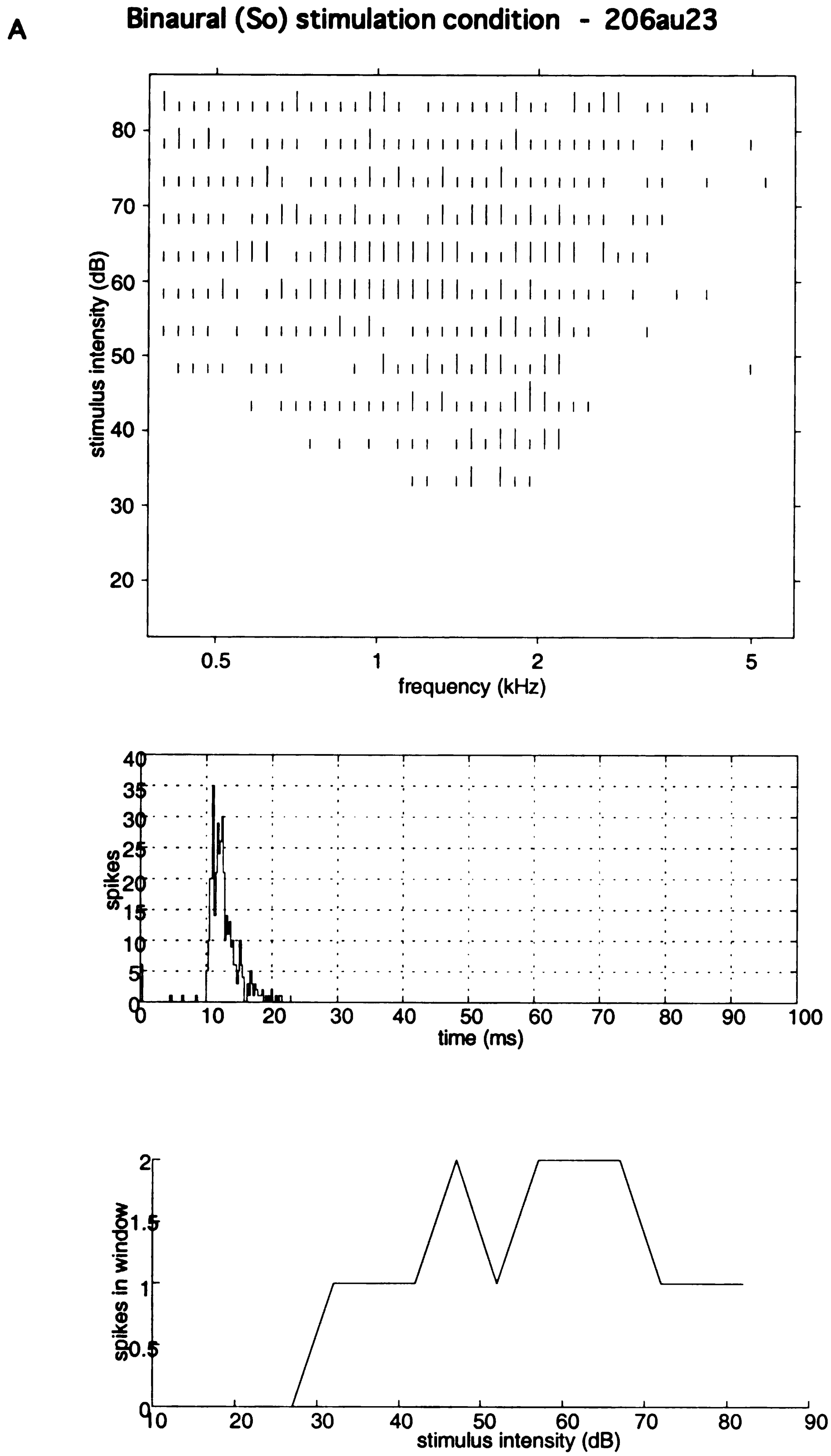
Binaural interaction of acoustic stimuli is known to be inherent in the auditory signal processing by the time neurally-encoded signals reach the cortical level. Neural tuning of different binaurally-configured stimuli should be evident, and should be reflected by the receptive field properties of individual cells. Psychophysical experiments have demonstrated large effects on threshold detection in the presence of noise for different binaurally-configured stimuli. Furthermore, neurophysiological studies in AI low-frequency neurons have found a subset of units that are phase-dependent (IPD cells). In the present study, however, there was no attempt to characterize or even isolate units that may be phase-dependent in the low-frequency region of AI, since the goal of this study was to study the effects of background noise on the spectral tuning of tonal stimuli in either monaural or binaural stimulation conditions. Therefore, to maximize the likelihood of seeing the largest effects physiologically in the receptive field properties of AI neurons, only the two most extreme (in terms of interaural phase difference) binaural configurations were selected as stimuli in the current series of experiments: 1) homophasic (0° in-phase) tones in continuous background noise and 2) antiphasic (180° out-of-phase) tones in continuous background noise. Cortical locations where responses to homophasic (So) tones with and without noise and to antiphasic (Spi) tones with noise were recorded are included in the subsequent analyses. Frequency response areas were collected from 61 single units recorded in the primary auditory cortex of nine adult cats. The response parameters of the spectral tuning curves were analyzed to study the effects of the binaural configuration of tonal stimuli in continuous noise on spectral tuning properties. Again, the methodological approach here was to study single unit responses to these binaural stimuli and not to derive a spatial distribution of the neural response properties ascribed to each of the three binaural listening conditions.

Receptive field responses to the binaurally-configured stimuli are illustrated in the following examples from two AI neurons. Figure 1.13A depicts a frequency response area recorded for binaural (So) tonal stimulation, a histogram of the spike onset time at CF, and a rate-level function at CF. The broadly-tuned unit described in this figure had a threshold of 32.5 dB SPL and an estimated CF of 1.8 kHz. The minimum onset latency shown in the histogram was around 10 ms. The non-monotonic rate-level function had a transition point at around 47.5 dB SPL. Figure 1.13B shows two obvious changes in the frequency response area when background noise was added to the homophasic tonal stimulation condition: 1) the threshold was elevated by 15 dB to 47.5 dB SPL and 2) the spike response decreased with the addition of background noise, as was evidenced by the patchiness of the response area. The histogram shows a slightly longer minimum onset latency with the addition of the background noise, as was seen in earlier examples (in Section 1). The rate-level function shows an upward shift of the transition point by 15 dB with the addition of binaural background noise. The effect of the out-of-phase tonal stimuli (Spi) upon background noise masking as reflected in the spectral tuning of primary auditory neurons is seen in figure 1.13C. Compared to the homophasic tones in noise condition, the antiphasic tones in noise condition yielded less of a threshold elevation. The reduction in evoked spike responses was still apparent in the patchiness of the frequency response area. The histogram reveals a similar time course for spike activity and the transition point shown in the rate-level function was the same as in the homophasic tones in noise condition.

The variability in tuning curve response differences between the two binaurally-configured tones in noise stimuli can be viewed in the next example. The single unit shown in figure 1.14A had a threshold of 42.5 dB at a CF around 0.8 kHz and was sharply tuned. According to the spike histogram, the minimum spike onset latency was

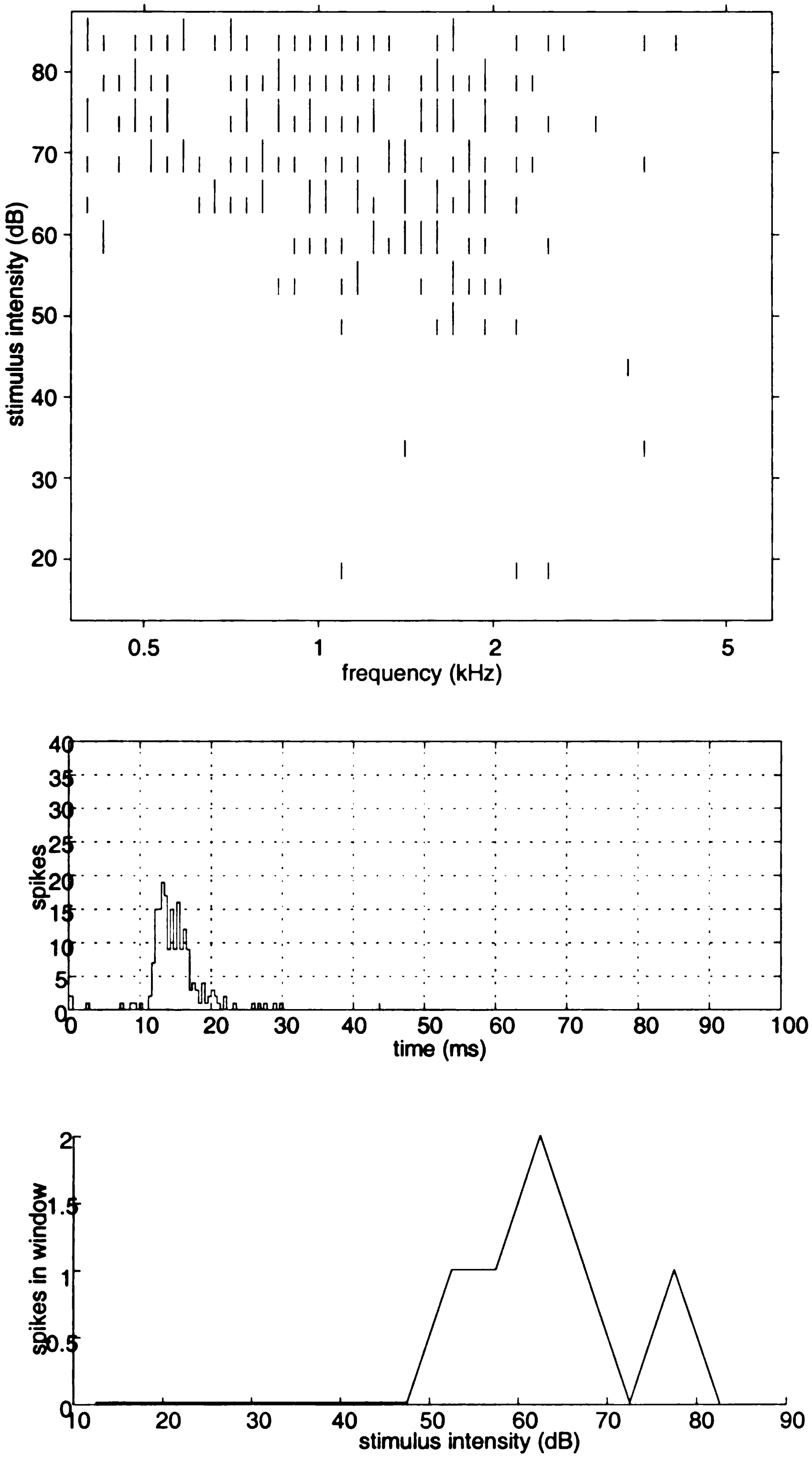
Figure 1.13 Binaural FRA for three different stimulation conditions

The binaural FRA recorded from a single unit for three different stimulation conditions are displayed in **A** (tones in-phase, no-noise), **B** (tones in-phase, in noise), and **C** (tones out-of-phase, in noise). The FRA-derived spike time histogram and rate-level profile are displayed for each corresponding stimulus condition. In **A**, the FRA depicts a broadly-tuned single unit with an estimated CF of 1.8 kHz and response latency of ~ 10 ms which demonstated non-monotonic behavior. The addition of continuous background noise, which is seen in **B**, resulted in the elevation of threshold, a decrease in spike activity, a delayed minimum onset latency, and an elevation of the transition point. Compared to the homophasic stimulus-in-noise condition, the inversion of interaural tonal phase in continuous background noise resulted in less of a threshold elevation, which is illustrated in **C**. All other measured properties were the same between the two signal phase conditions.


A
Binaural (So) stimulation condition - 206au23
stimulus intensity (dB)
20
30
40
50
60
70
80
0.5
1
2
5
frequency (kHz)
spikes
0
5
10
15
20
25
30
35
40
0
10
20
30
40
50
60
70
80
90
100
time (ms)
spikes in window
0
0.5
1
1.5
2
10
20
30
40
50
60
70
80
90
stimulus intensity (dB)

B

Binaural (So) with Noise Stimulation - 206au23



Binaural (Spi) with Noise Stimulation - 206au23

C

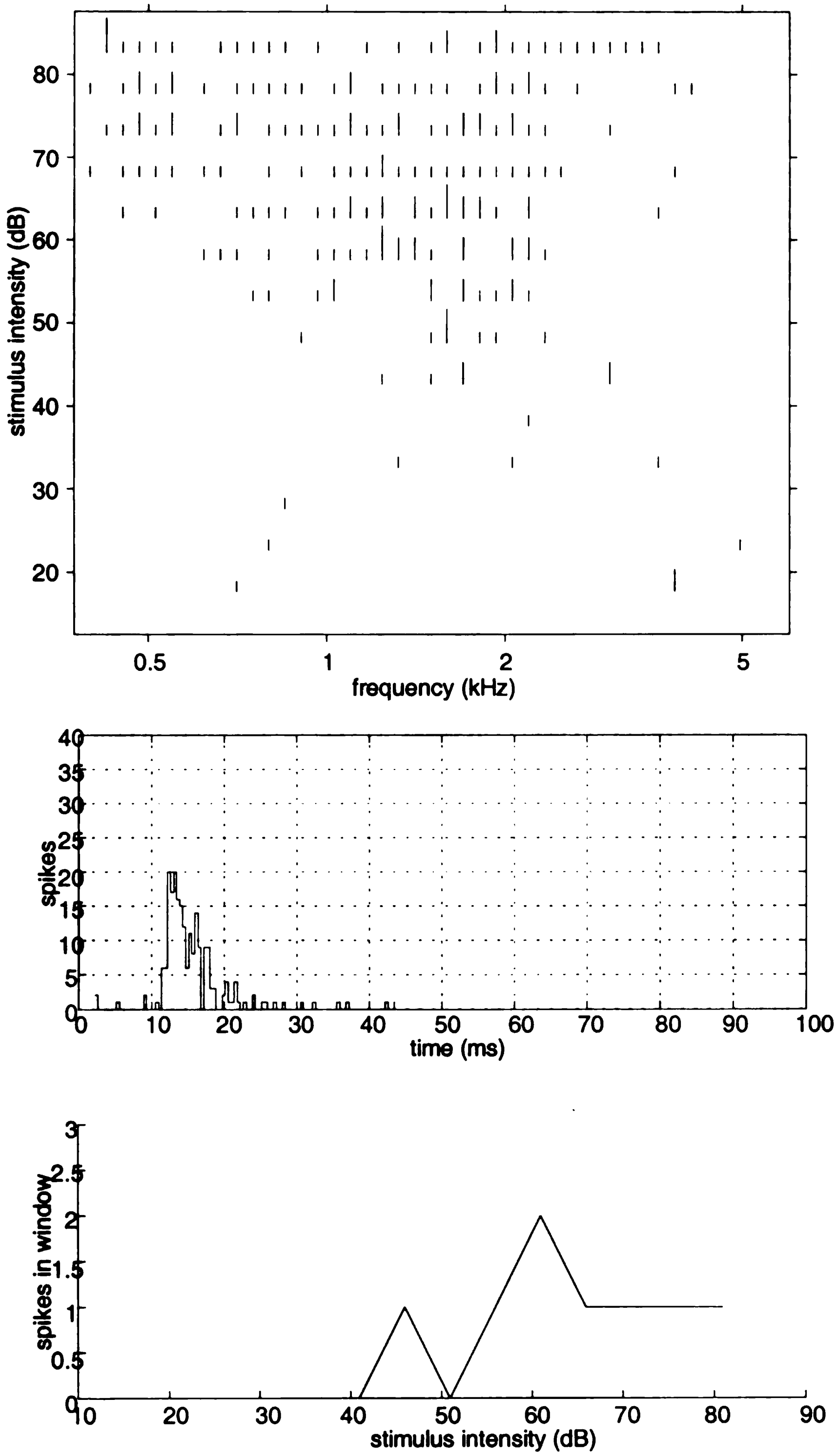
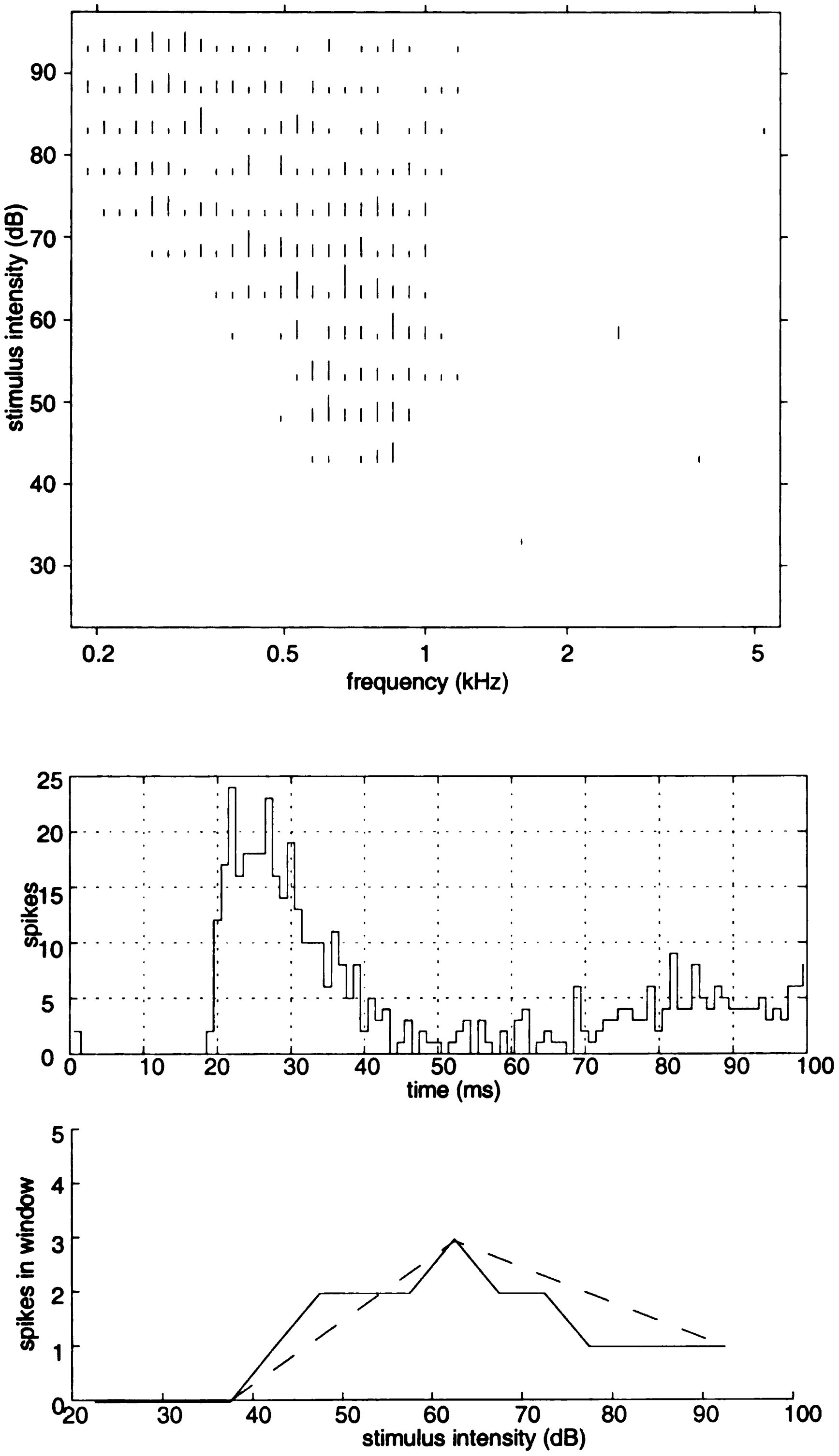


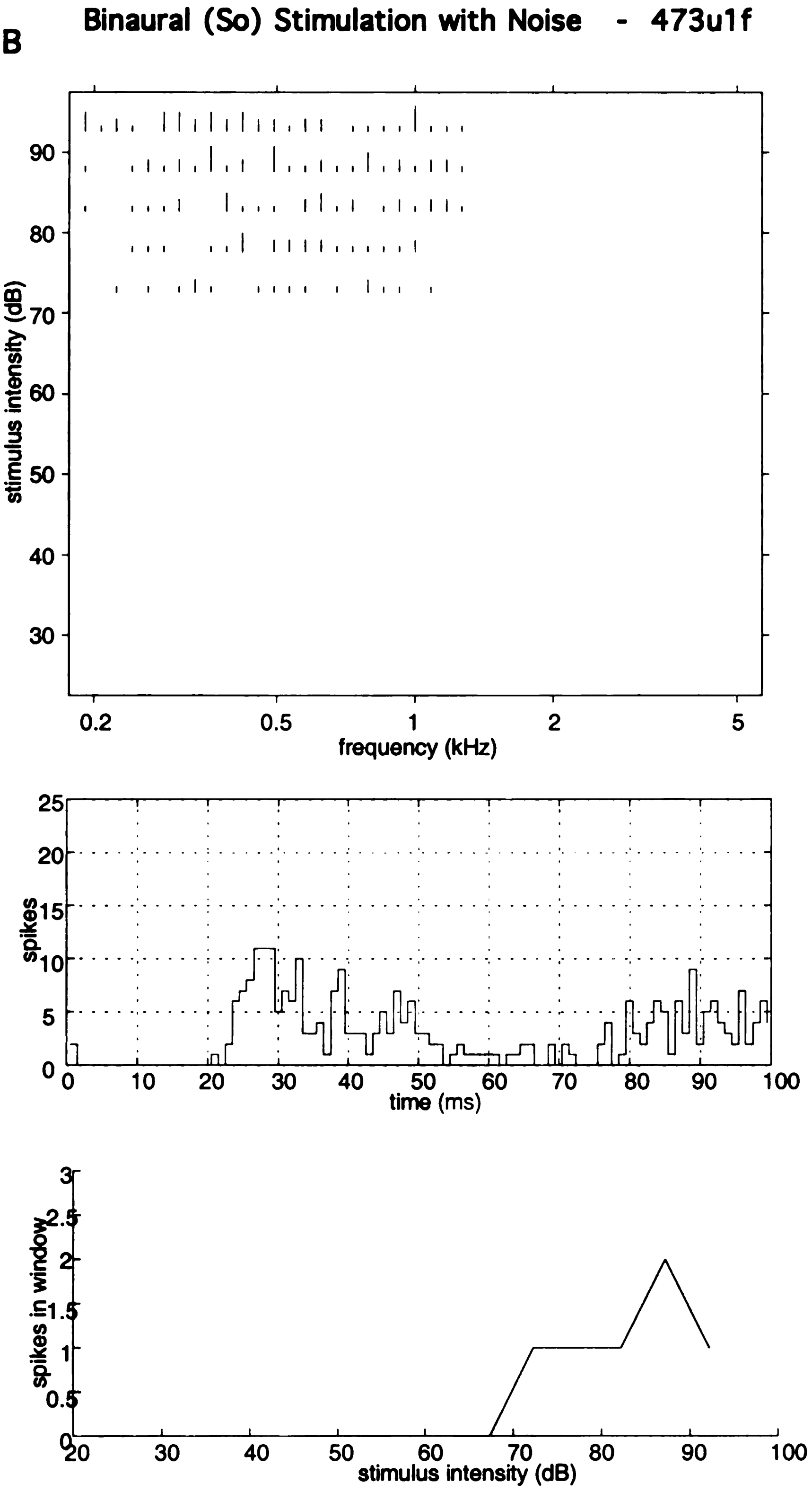
Figure 1.14 Another example of binaural FRA for three different stimulation conditions

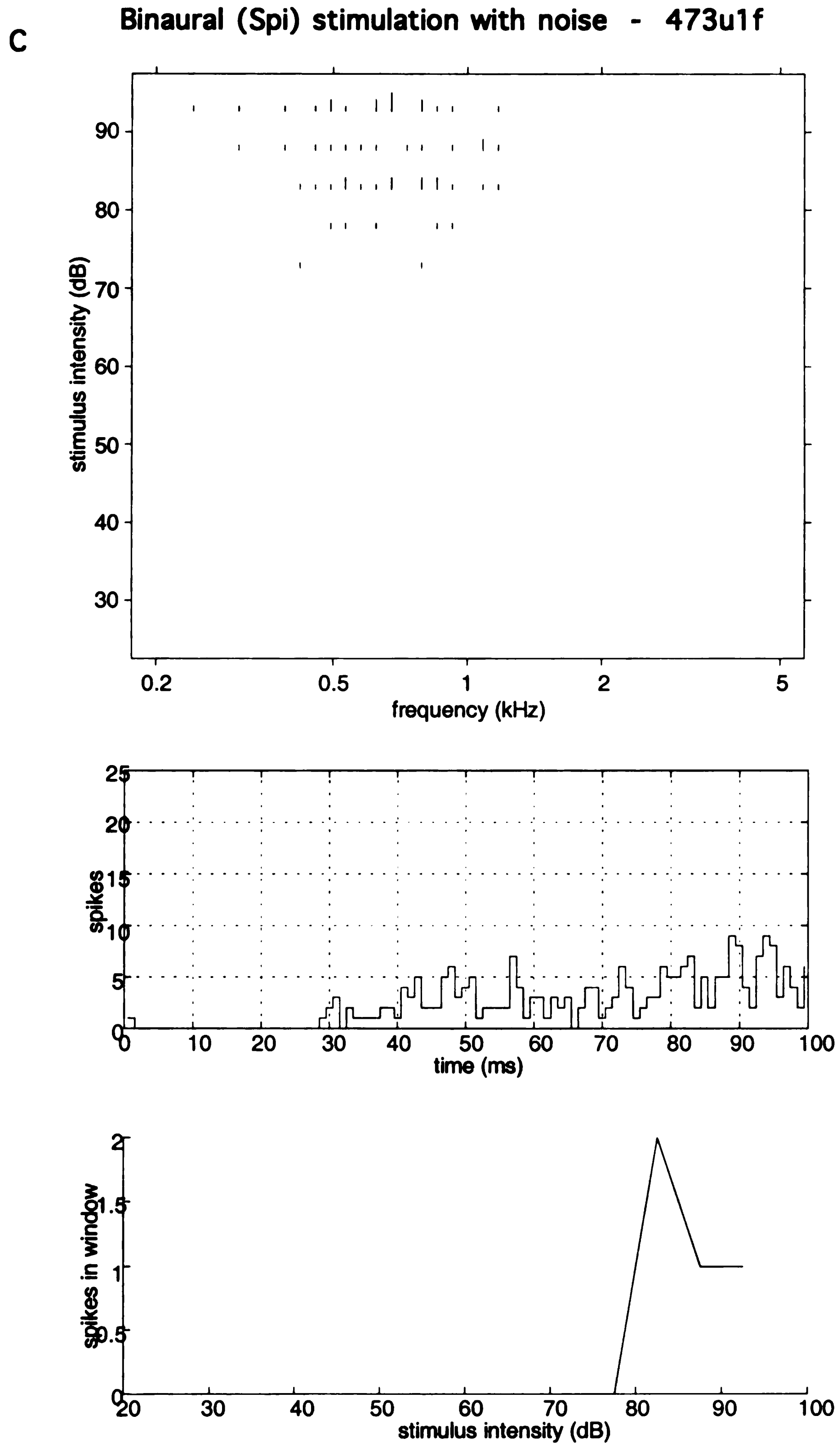
The binaural FRA recorded from another single unit for three different stimulation conditions are displayed in **A** (tones in-phase, no-noise), **B** (tones in-phase, in noise), and **C** (tones out-of-phase, in noise). The FRA-derived spike time histogram and rate-level profile are displayed for each corresponding stimulus condition. In **A**, the FRA manifests a sharply-tuned single unit with an estimated CF of 800 Hz and an onset latency of ~18 ms displaying non-monotonic behavior. The addition of continuous background noise, seen in **B**, resulted in an elevation of threshold, a broadened bandwidth near threshold, a delayed onset latency, and an upward shift in the transition point. When interaural phase of the tones were inverted in the presence of continuous background noise, bandwidth broadening was not apparent and onset latency was delayed even further, which was evident in **C**.

A **Binaural (So) Stimulation Condition - 473u1f**



B


Binaural (So) Stimulation with Noise - 473u1f
stimulus intensity (dB)
90
80
70
60
50
40
30
0.2
0.5
1
2
5
frequency (kHz)
spikes
25
20
15
10
5
0
0
10
20
30
40
50
60
70
80
90
100
time (ms)
spikes in window
3
2.5
2
1.5
1
0.5
0
20
30
40
50
60
70
80
90
100
stimulus intensity (dB)


C
Binaural (Spi) stimulation with noise - 473u1f
stimulus intensity (dB)
30
40
50
60
70
80
90
frequency (kHz)
0.2
0.5
1
2
5
spikes
0
5
10
15
20
25
time (ms)
0
10
20
30
40
50
60
70
80
90
100
spikes in window
0
0.5
1
1.5
2
stimulus intensity (dB)
20
30
40
50
60
70
80
90
100

about 18 ms. The rate-level function was non-monotonic, with a transition point at 62.5 dB SPL and a slope decreasing to more than half of the peak response amplitude. The effect of background noise on the homophasic tonal stimulus condition as shown in figure 1.14B resulted in elevated response threshold and broadened FRA bandwidth near threshold. The addition of background noise delayed the minimum onset latency by at least 2-3 ms. The transition point shifted upward by 25 dB. The effect of flipping the interaural phase of the tone by 180° while maintaining the same level of binaural background noise is shown in figure 1.14C. Although the threshold response remained the same as in the homophasic tones-in-noise example, the bandwidth of the frequency response area near threshold was not as broad, due to a reduction in response at frequencies below 0.5 kHz. Compared to the homophasic condition, the phase inversion of the tones also resulted in a delay in the minimum onset latency by at least 5 ms as manifested in the spike-time histogram. As in the previous unit's example, the transition point of the firing rate did not shift with the phase inversion of the tones.

From these examples of two different CF single units, binaurally-configured stimuli appear to exert their influence differentially over the spectral tuning characteristics of cells in the primary auditory cortex. The degree of difference in threshold, bandwidth near threshold, and minimum onset latency between the two extreme cases of interaurally phase different stimuli appears to be dependent upon the CF of the unit. However, whether these observed differences covary with CF or are even statistically significant will be examined more completely in a later section.

The following group statistics were compiled from the frequency response area extracted parameters measured for each stimulation condition. Of the 61 single units studied, the CF of 57% of the units were below 2 kHz. For the purpose of classifying these units, cells with a CF of 2 kHz or less were considered low-frequency while those with a CF greater than 2 kHz were considered high-frequency cells. Due to the nature of this study, a within-subjects ANOVA (repeated measures) was used to test the null hypothesis that no changes occurred with the addition of noise. If the null hypothesis was rejected, then a paired t-test was used to test the null hypothesis that no changes occurred due to the change of tonal stimulus interaural phase. According to table 1.2, which displays the mean values for frequency response parameters measured in three binaural stimulus conditions (no-noise, homophasic, and antiphasic), parametric statistical analysis using ANOVA repeated measures revealed a significant effect of binaural stimulus configuration in noise upon response threshold and onset latency (see table 1.2). The presence of continuous background noise effectively elevated thresholds and prolonged minimum onset latencies (ANOVA repeated measures, $p < 0.0001$) compared to the no-noise tonal condition. In addition, the difference in mean threshold and mean minimum onset latency between the two binaurally-configured tone-in-noise stimulus conditions revealed lower mean threshold and longer minimum onset latency in the antiphasic condition (paired t-test, $p < 0.05$). There appeared to be no significant difference in bandwidth attributable to either the addition of background noise or the change in tonal stimulus interaural phase.

As a second step to viewing global changes in receptive field response properties due to binaural stimulus configuration, correlation of the response difference between the two binaurally-configured stimulus in noise conditions and the response measured in the tonal stimulus condition without noise was performed for each parameter measured. Table 1.3 reflects the resulting correlation coefficients. Three parameters were found to be significantly correlated with differences measured between the two binaurally-configured stimulus condition: 1) characteristic frequency was found to be significantly correlated with the difference in threshold ($R=-0.263$, $p < 0.05$) and minimum onset latency ($R=0.265$, $p < 0.05$), 2) onset latency was found to be significantly correlated with onset latency difference ($R=-0.403$, $p < 0.01$), and 3) monotonicity was found to be significantly correlated with bandwidth 10 difference ($R=-0.334$, $p < 0.01$).

Table 1.2
Mean parameter values and corresponding standard deviations (in parantheses) for three different stimulus conditions.

		No-noise	Homophasic	Antiphasic
Threshold (dB) ***	N=61	23.5 (8.9)	39 (9.5)	36.3 (10.7)
Bandwidth 10 (octaves)	N=61	0.697 (0..482)	0.845 (0.601)	0.726 (0.496)
Bandwidth 20 (octaves)	N=54	1.199(0.670)	1.343(0.951)	1.162 (0.654)
Onset Latency (ms) ***	N=57	14 (5)	16 (5)	17 (7)

level of significance

*** < 0.0001

Table 1.3
R values from regression analysis of spectral parameter difference between the two noise conditions (homophasic and antiphasic) and the parameter from the no noise condition

	Threshold Δ	BW10 Δ	BW20 Δ	Onset Latency Δ
	N=61	N=61	N=54	N=57
CF	-0.263*	-0.066	-0.129	0.265*
Threshold	0.082	0.094	-0.104	-0.147
BW10	0.019	0.072	0.239	-0.036
BW20	0.167	-0.085	0.108	0.050
Onset Latency	0.076	0.009	-0.013	-0.403**
Monotonicity	-0.014	-0.334**	-0.227	0.101

level of significance

 * < 0.05
** < 0.01

In summary, the overall group statistics showed significant effects of binaural stimulus configuration on response threshold and minimum onset latency, but no effect on frequency response bandwidth. Correlation of the difference in response properties of receptive fields obtained from the binaural tone-in-noise conditions with the binaural response property without noise finds that there are certain response properties that covary with some of the differences observed in the binaurally-configured stimulus conditions. The analysis in the following sections will cover the effect of binaural stimulus configuration on receptive field response properties in further detail.

1.2.6 Threshold

A distribution of FRA thresholds plotted as a function of CF, as seen in figure 1.15, revealed that of the 61 single units analyzed, 85% of the units had mid- to high-thresholds (> 15 dB) while the remaining 15% had low-thresholds. In this section, labeling of unit types (e.g., high-threshold) are based upon the homophasic tonal stimulation condition. Figure 1.15 also shows that the range of binaural thresholds measured were spread evenly across the range of CF, although at very high frequencies (which only represents three units) only low-threshold units were sampled. The addition of binaural background noise created a mean upward shift of at least 13 dB in threshold (see table 1.2). Again, the level of noise used in this part of the study was identical to the previous section. The goal was to apply enough masking noise to achieve 20 dB masked thresholds in the contralateral stimulation condition. The same noise level was then applied in the binaural stimulation condition. In comparing thresholds between the two binaural configured noise conditions, the mean threshold of the homophasic condition was 3 dB higher than in the antiphasic condition ($p < 0.05$). Regression analysis of threshold whose distributions for either binaural condition were plotted in figure 1.16, was done to show any effect of interaural phase difference on tuning threshold. The frequency response thresholds measured in the

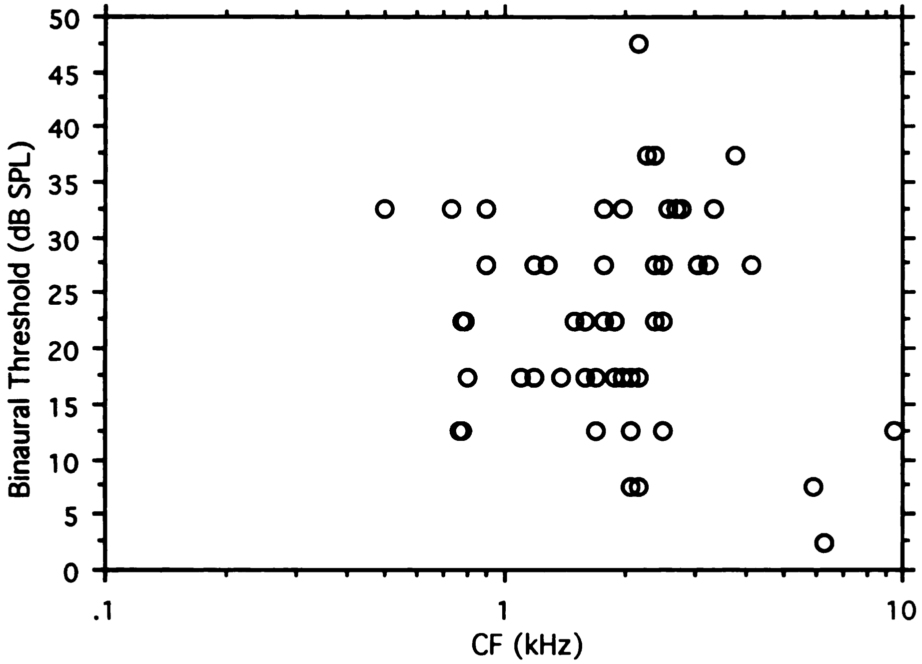


Figure 1.15 Distribution of binaural FRA threshold

The distribution of binaural FRA threshold is plotted as a function of CF. Thresholds were widely and evenly distributed across the frequency axis. Of the 61 single units, 85% of the units had mid- to high- thresholds (>15 dB) while the remaining 15% had low thresholds.

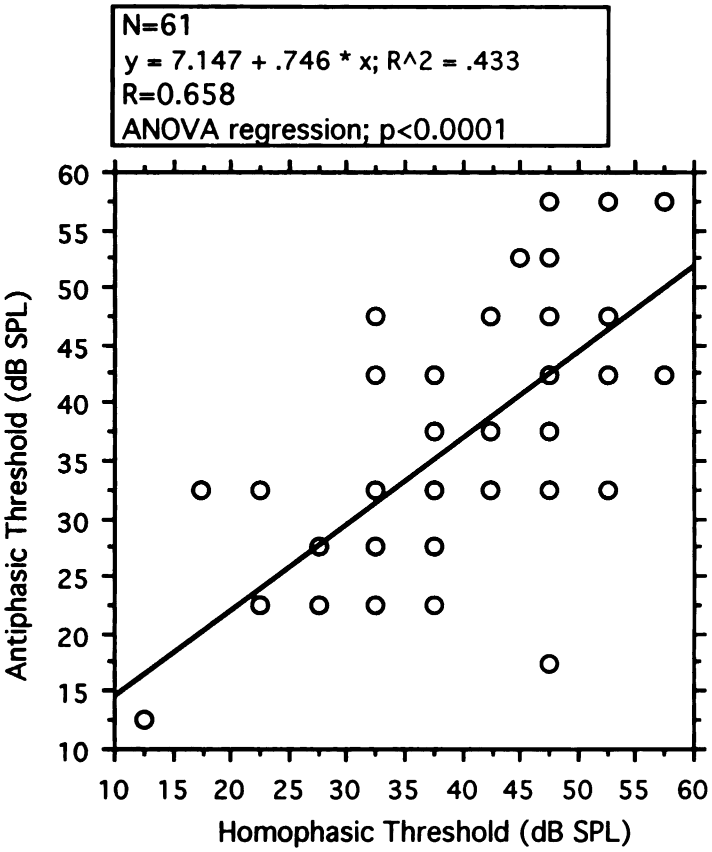


Figure 1.16 Distribution of FRA threshold measured in continuous background noise

The distribution of FRA threshold measured in continuous background noise is plotted to show the relationship in terms of threshold between two different binaural stimulus configurations. Thresholds from the antiphasic stimulus condition are plotted as a function of thresholds from a homophasic stimulus condition. Regression correlation of thresholds from two binaural stimulus conditions revealed a significant positive correlation (R=.658; p<.0001) and a regression slope slightly less than the unity line.

homophasic and antiphasic signal-in-noise conditions showed a strong positive correlation (R=0.66) indicating a similar variance in threshold between the two binaural conditions.

Further analyses of the effect that interaural phase has on threshold response were done by calculating the threshold difference that results from the two stimulus conditions. As seen in figure 1.17, threshold differences between homophasic and antiphasic signal-in-noise conditions were widely distributed across CF. Of the 61 single units, 25% of the units had higher thresholds in the antiphasic signal-in-noise condition, 20% showed no difference in threshold level between the two signal-in-noise conditions, and the remaining 55% of the units had higher thresholds in the homophasic signal-in-noise condition. Threshold differences between the So and Spi condition were not significantly correlated with response threshold without noise or any of the other measured spectral tuning parameters other than CF (see table 1.3).

1.2.7 Bandwidth

Receptive field bandwidths were measured at 10 dB, 20 dB, and 30 dB above threshold to compare the effect of interaural phase differences on excitation bandwidth in the FRA. The distribution of bandwidth plotted as a function of CF in figure 1.18 revealed a range of bandwidths measured at 10 dB above threshold extending from less than a quarter of an octave to three octaves. The percent distribution of bandwidths measured at 10 dB and 20 dB above threshold shown in figure 1.18B and 1.18C revealed sharply-tuned units in slightly less than 45% of the population (bandwidth of 0.5 octave or less at 10 dB above threshold and bandwidth of one octave or less at 20 dB above threshold). The remaining 65% of the units were broadly tuned (bandwidth larger than 0.5 octave at 10 dB above threshold and bandwidth larger than one octave at 20 dB above threshold). Regression analysis of bandwidth as seen in figure 1.19A revealed a positive correlation between bandwidths measured at 10 dB and 20 dB for each binaural stimulation

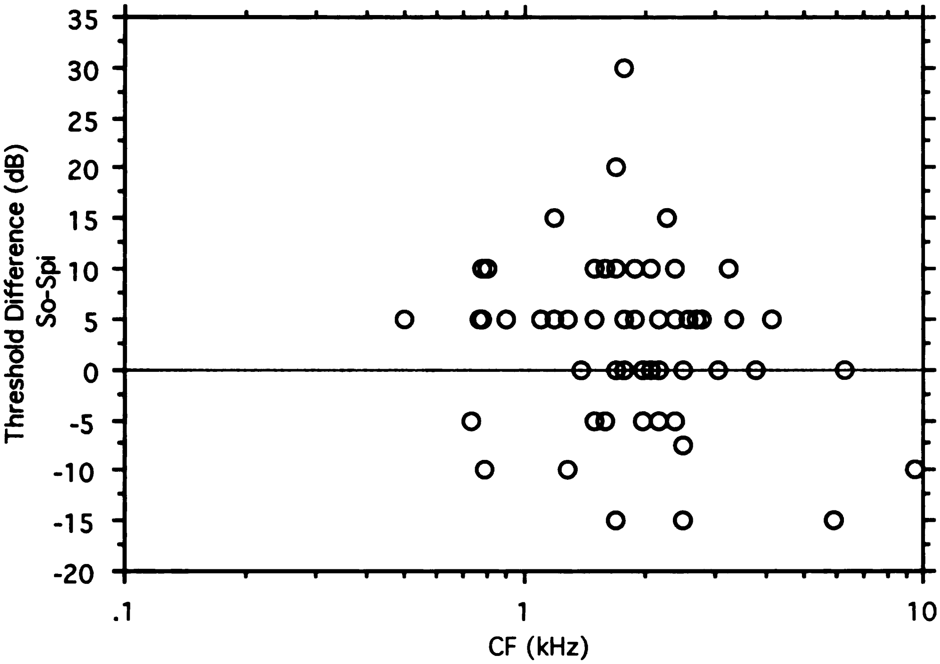


Figure 1.17 Distribution of FRA threshold difference

The distribution of FRA threshold difference, which measures the amount of change between interaural in-phase and out-of-phase tones presented in noisy backgrounds, is plotted as a function of CF. Threshold differences were widely distributed across CF.

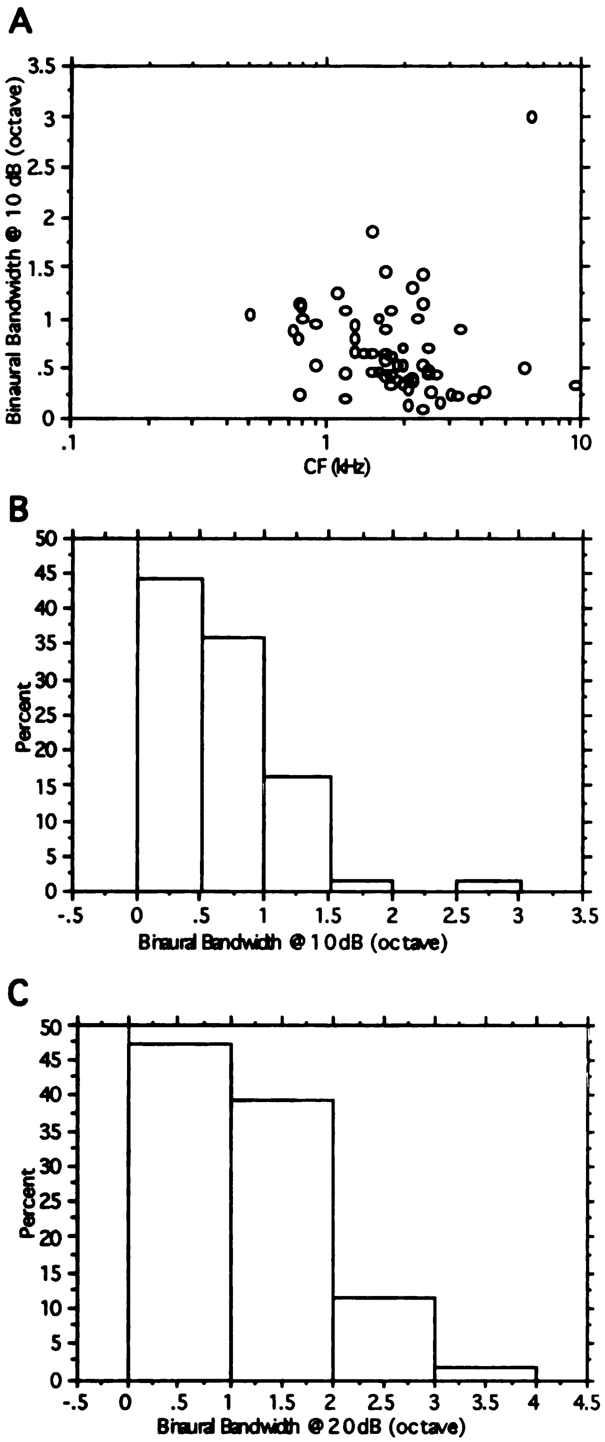


Figure 1.18 The distribution of binaural FRA bandwidth

In **A**, the distribution of FRA bandwidth, which is plotted as a function of CF, revealed a wide distribution of bandwidth 10 dB across the frequency axis. In **B**, the distribution of the percentage of bandwidth 10 measured revealed that the majority of the units had bandwidths of less than one octave. In **C**, the distribution of the percentage of bandwidth 20 measured revealed that the majority of the units had bandwidths less than two octaves.

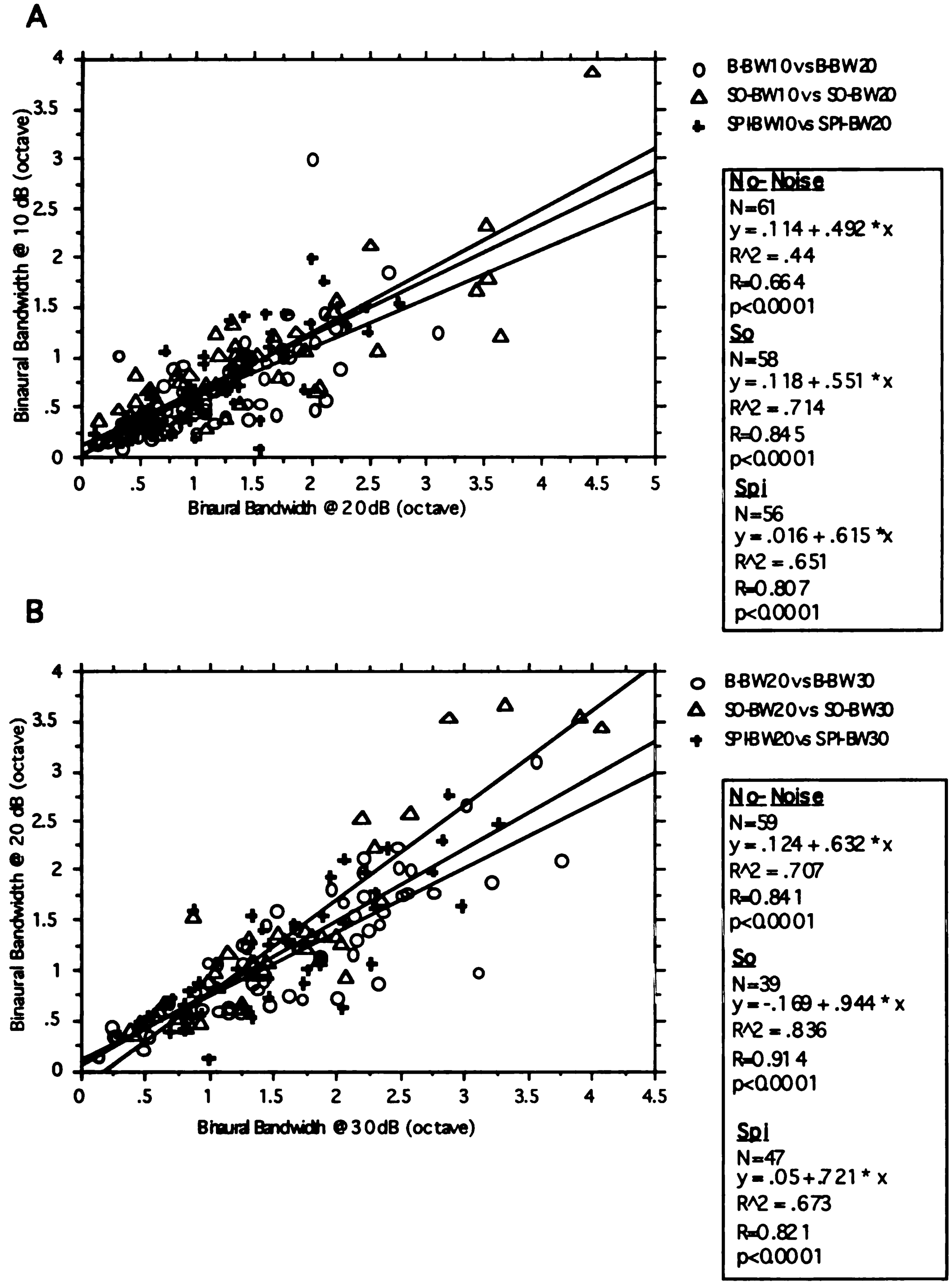


Figure 1.19 Distribution of binaural bandwidths at three stimulus levels

In **A**, binaural bandwidth measured at 10 dB (re: threshold) is plotted as a function of bandwidth at 20 dB (re: threshold) for each stimulus condition . In **B**, binaural bandwidth measured at 20 dB (re: threshold) is plotted as a function of bandwidth at 30 dB(re: threshold for each binaural stimulus condition. Regression correlation of the bandwidths in **A** and **B** showed significant positive correlations.

condition. The variance calculated for each stimulation was similar regardless of the addition of noise or interaural stimulus condition. This is an indication that the rate at which tuning broadens as stimulus level increases by 10 dB near threshold may not differ when background noise is present nor when tones have different interaural phase. Figure 1.19B depicts three positive regression correlations of bandwidth measured at 20 dB and 30 dB above threshold, one for each stimulation condition. The variability in the ratio of bandwidth for the three stimulation conditions is greater for this 10 dB increase in stimulus level, but according to ANOVA repeated measures, the differences noted among the three slopes were not statistically significant.

The effect of background noise on bandwidth was again examined by plotting the noise-induced change in bandwidth (So with noise - No noise) as a function of the tone-evoked excitation bandwidth of the frequency response area. Figure 1.20 shows a scatterplot for each of the three bandwidth level measured. Regression analysis of the three functions revealed significant negative correlation for bandwidth levels 10 dB and 30 dB above threshold and their respective bandwidth differences. This bandwidth finding in the binaural stimulus condition further supports the earlier negative correlation result depicted in figure 1.9A for the monaural stimulation condition. That is, at least for bandwidth levels 10 dB and 30 dB above threshold, the direction of bandwidth change appears to be dependent upon the sharpness of tuning of the unit as described by tonal stimulation. Sharply-tuned neurons respond to tonal stimulation in the presence of background noise by broadening bandwidth while broadly-tuned neurons respond by sharpening bandwidth.

Scatterplots of the distribution of bandwidths at 10 dB, plotted in figure 1.21A, and 20 dB, plotted in figure 1.21B, above threshold show a positive correlation between homophasic and antiphasic stimulation conditions. Sharply- and broadly-tuned units were divided according to the aforementioned criterion in these figures by the dotted line. In figure 1.21B, 68% of the sharply-tuned units were located above the unity line, indicating that their bandwidth was wider in the antiphasic condition. In contrast, 69% of the

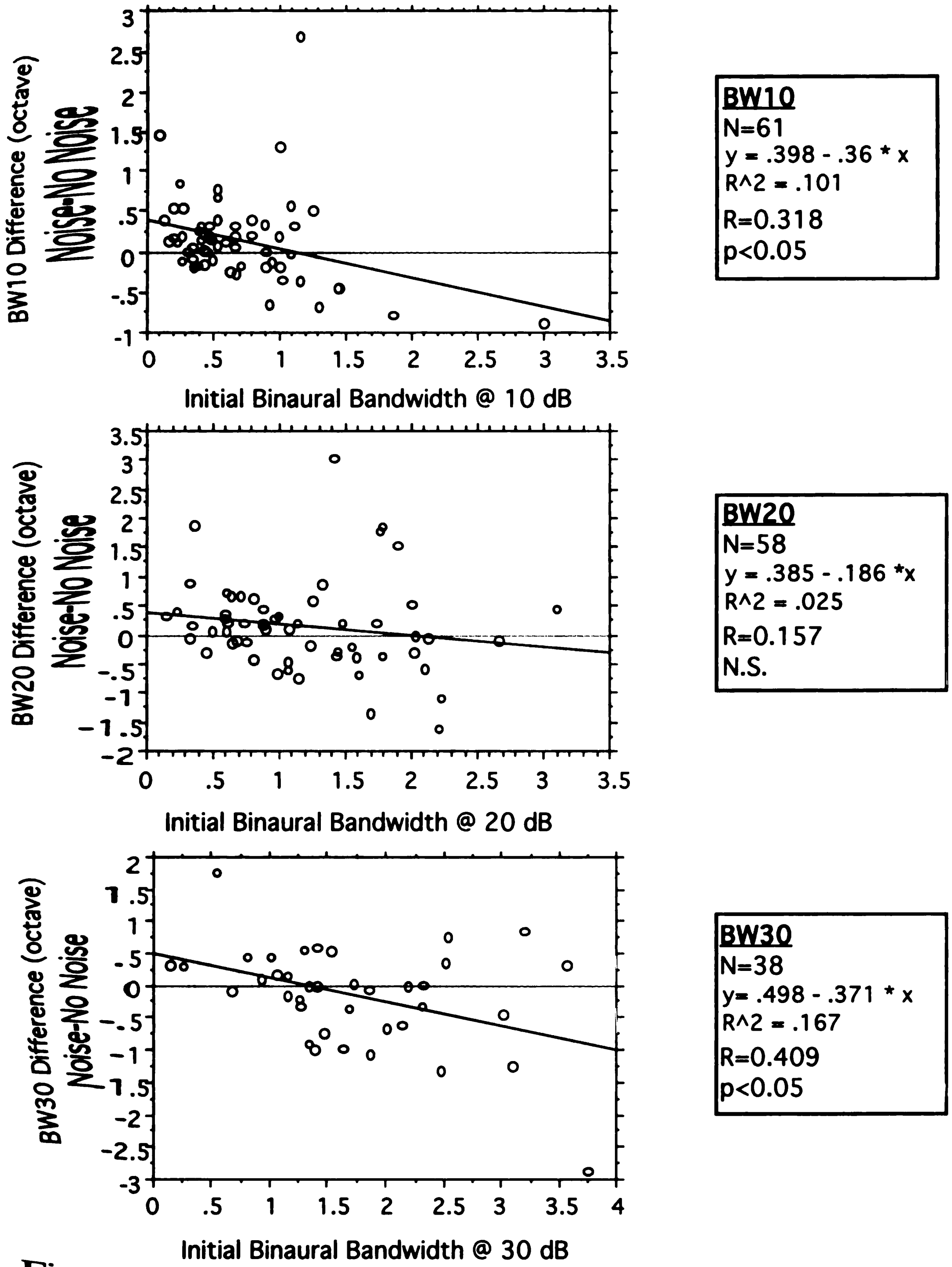


Figure 1.20 Background noise effects on bandwidth at three levels above FRA threshold

The effect of background noise on bandwidth was depicted by plotting bandwidth difference (in-phase condition only) before and after the addition of background noise as a function of initial bandwidth at three different levels above the FRA threshold for each single unit. Regression correlation of bandwidths revealed a significant inverse relationship between bandwidth difference and initial bandwidth at 10 dB and 30 dB above threshold.

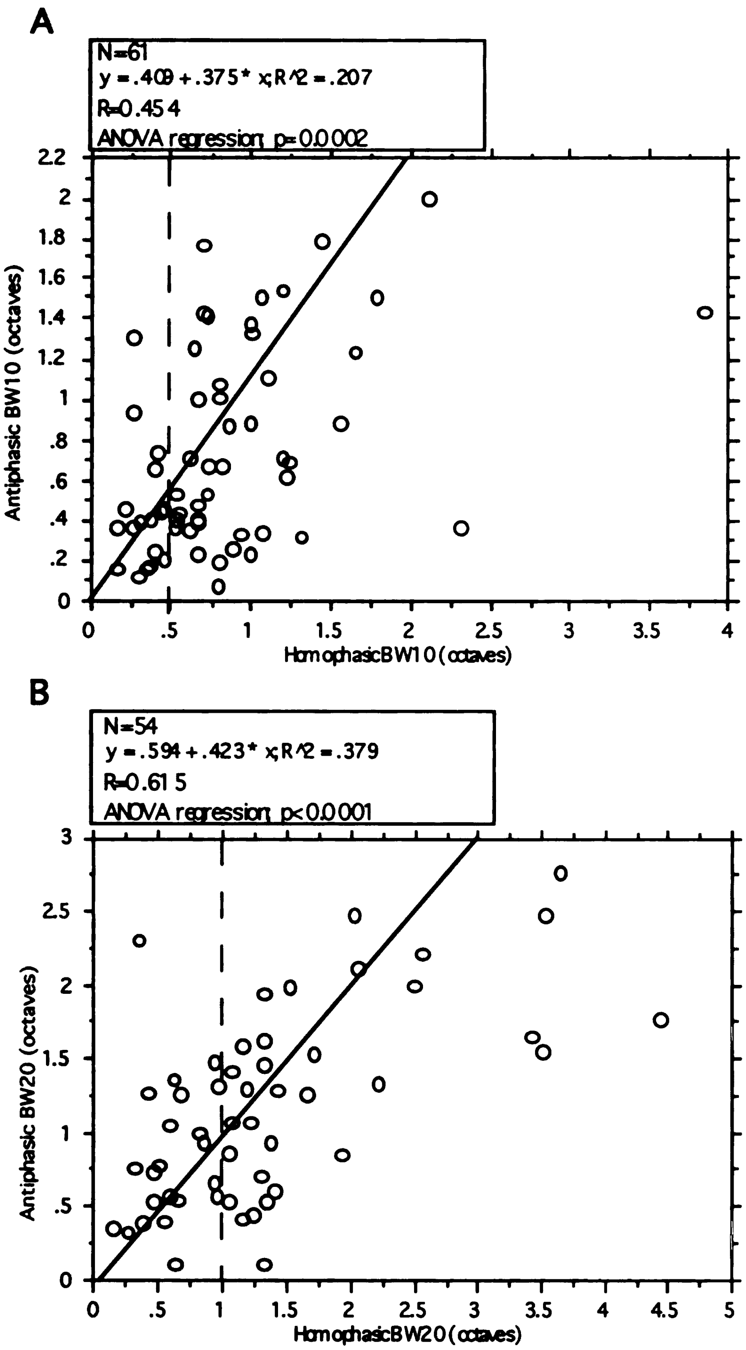


Figure 1.21 Distribution of bandwidth for antiphasic and homophasic signal-in-noise conditions

The distribution of bandwidth for antiphasic and homophasic signal-in-noise conditions are depicted in A for bandwidth 10 and in B for bandwidth 20. In both A and B, bandwidths in the antiphasic condition were plotted as a function of homophasic condition. Regression correlation revealed significant positive correlations for both bandwidths 10 dB and 20 dB.

broadly-tuned units were located below the unity line, indicating that their bandwidth was wider in the homophasic condition.

To further study the effect of interaural phase difference on binaural signal-in-noise stimulation, bandwidth differences between homophasic and antiphasic tones in continuous background noise were calculated. According to the distribution of bandwidth differences which were plotted in figures 1.22A and 1.22B as a function of CF, bandwidth 10 differences and bandwidth 20 differences between homophasic and antiphasic signal-in-noise conditions were widely distributed across CF and centered around zero. The distribution of bandwidth 10 or bandwidth 20 difference across CFs greater than 1 kHz did not appear to be segregated by sharpness of tuning (as classified in the binaural, no-noise condition). In the presence of background noise, units below 1 kHz which also tended to be broadly tuned showed greater tendency to have narrower bandwidths in the antiphasic signal-in-noise condition than in the homophasic. Of the 61 single units, 50% of the units had broader tuning near threshold in the homophasic signal-in-noise condition, 18% showed no difference in bandwidth 10, and the remaining 32% had broader tuning near threshold in the antiphasic signal-in-noise condition.

Correlation of bandwidth difference with other parameters extracted from FRAs revealed a significant negative correlation between bandwidth 10 difference and monotonicity (see table 1.3). The classification of monotonicity depends upon an individual unit's firing capacity as sound level increases. In figure 1.23A, binaural bandwidth difference measured at 10 dB above threshold is plotted as a function of monotonicity. The figure shows that as the classification of monotonicity increases towards monotonic behavior (monotonicity > 0), the bandwidth 10 difference decreases such that frequency response tuning broadens in the antiphasic condition with respect to the homophasic signal-in-noise condition. Further classification of neurons by sharpness of tuning, as illustrated in figure 1.23B, reveals a significant negative correlation between monotonicity and bandwidth difference at 10 dB above threshold that is present in the

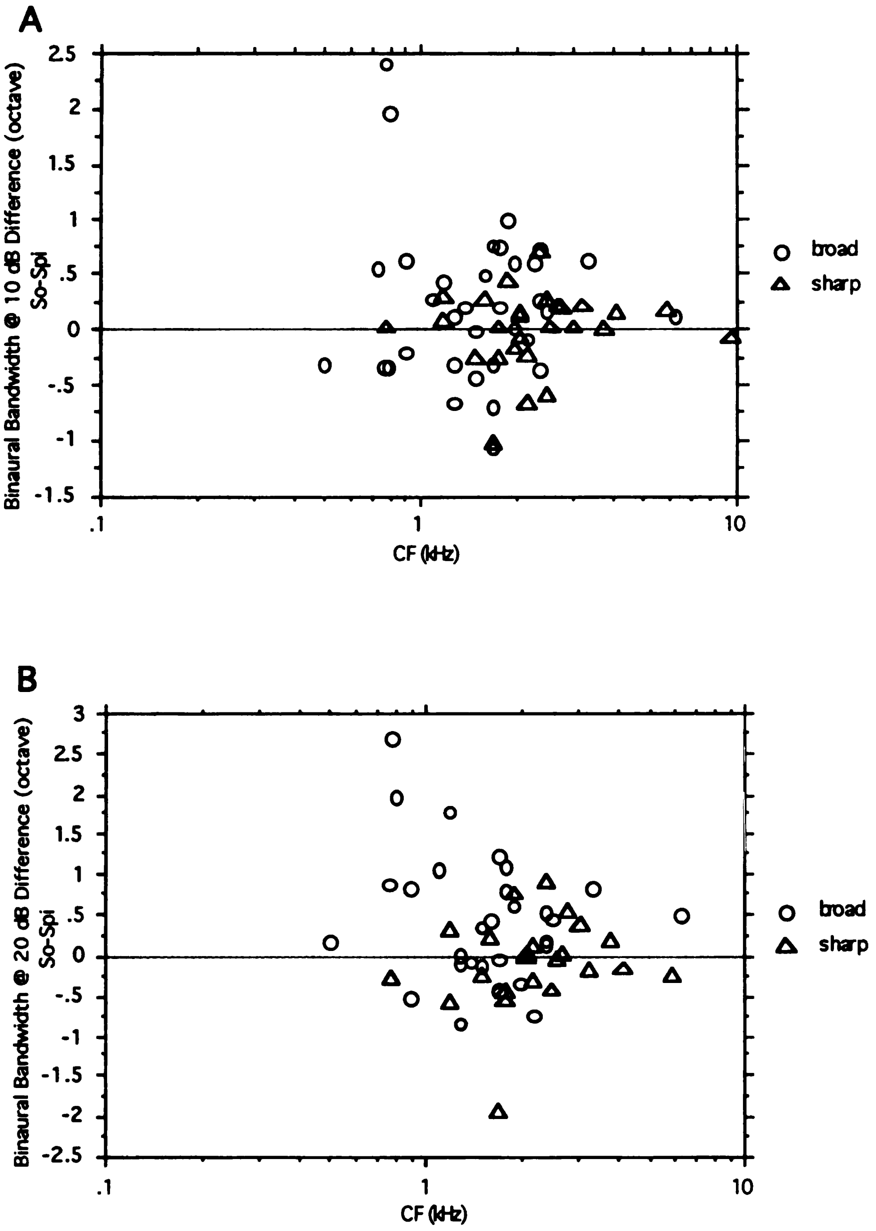


Figure 1.22 The effect of interaural phase difference on bandwidth

The distributions of bandwidth difference at 10 dB (A) and at 20 dB above threshold (B) measured between the two binaural phase conditions in the presence of continuous background noise are plotted as a function of CF. Negative bandwidth differences indicated broader bandwidth tuning in the antiphasic condition, while positive differences reflected narrower bandwidth tuning. The distribution of values were subcategorized on the basis of their sharpness of tuning in the absence of noise. Units with CFs less than 1 kHz, which were also mostly broadly-tuned, tended to exhibit narrower bandwidths in the antiphasic signal-in-noise condition than homophasic condition.

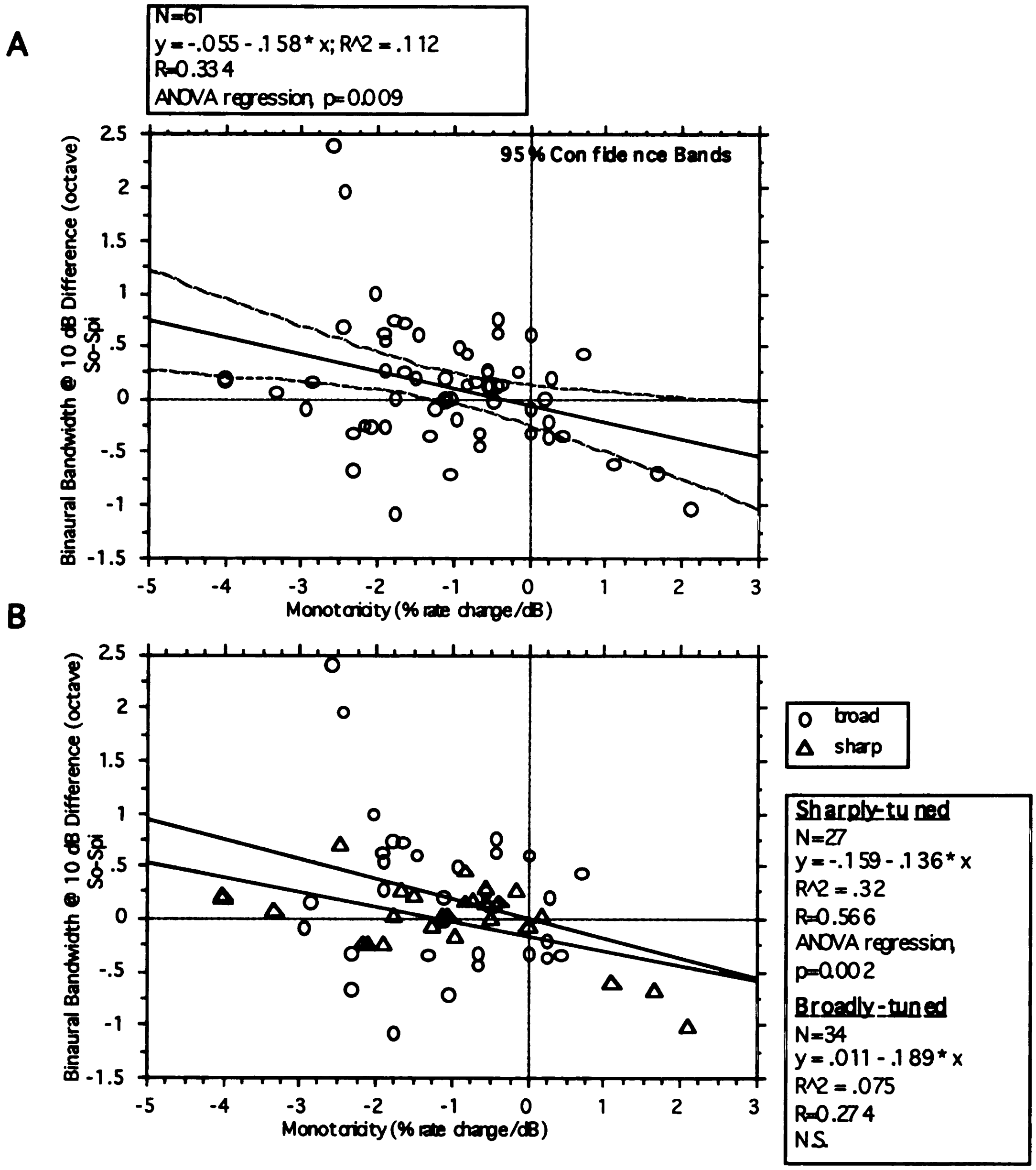


Figure 1.23 The effect of monotonicity on bandwidth difference between two interaural phase difference conditions

The distribution of bandwidth 10 dB differences, which measures the amount of change in bandwidth attributable to a change in interaural phase difference, is plotted as a function of monotonicity. In **A**, a significant negative correlation was found such that as monotonicity increases towards more monotonic behavior, frequency response tuning was broader in the antiphasic condition relative to the homophasic one. In **B**, when bandwidth difference values were subcategorized on the basis of the sharpness of tuning in the absence of noise, a significant negative correlation was found only for the sharply-tuned units.

sharply-tuned units and not for the broadly-tuned ones. This suggests that sharply-tuned, non-monotonic units have a greater tendency of being more narrowly-tuned in the antiphasic than the homophasic signal-in-noise condition; and that sharply-tuned units which are increasingly more monotonic have narrower bandwidths in the homophasic than antiphasic signal-in-noise condition.

1.2.8 Onset Latency

The distribution of minimum onset latency across CF, as seen in figure 1.24A, shows the majority of the units responding to binaural tones within the first 20 ms post-stimulus. The largest cluster of onset latencies resided in the 10 to 15 ms range which is consistent for AI neurons. Figure 1.24B shows a strong positive correlation in onset latency between antiphasic and homophasic signal-in-noise condition. The regression slope approximated the unity line. Statistically significant prolongation of mean minimum onset latency of at least 2 ms was observed in both signal-in-noise conditions (see table 1.2). Between the two binaural signal-in-noise stimulus conditions, mean onset latency was 1 ms longer in the antiphasic condition than the homophasic condition (paired t-test, $p < 0.05$).

The distribution of minimum onset latency differences between homophasic and antiphasic conditions, plotted in figures 1.25A and 1.25B (in logarithmic frequency units) were found to be weakly correlated with CF. At CF less than 1.5 kHz (or 0.18 in log frequency units), latencies measured in the homophasic condition tend to be shorter, thus resulting in negative difference values. As unit CF increases onset latency difference becomes increasingly positive such that at higher CFs the minimum onset latency in the antiphasic signal-in-noise condition first becomes no different and then even shorter than in the homophasic condition. The effect of stimulus interaural phase and background noise on minimum onset latency is represented by the distribution of onset latency differences plotted as a function of binaural onset latency in figure 1.26. Correlation of onset latency

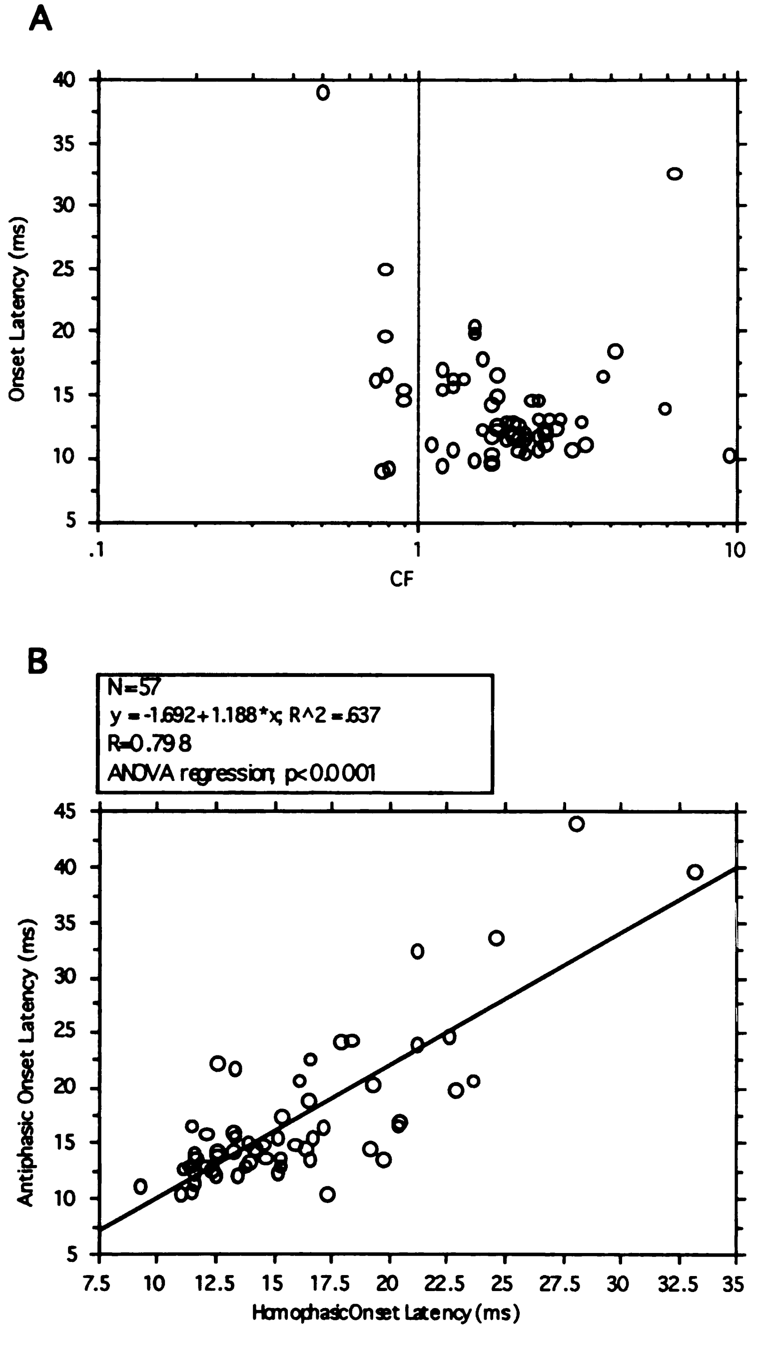


Figure 1.24 Distribution of binaural onset latency

In **A**, the distribution of minimum onset latency is plotted as a function of CF. The majority of the units responded to binaural tones within the first 20 ms after stimulus presentation, which is typical of AI neurons. In **B**, the distribution of onset latencies measured in both stimulus phase conditions revealed a significant positive correlation.

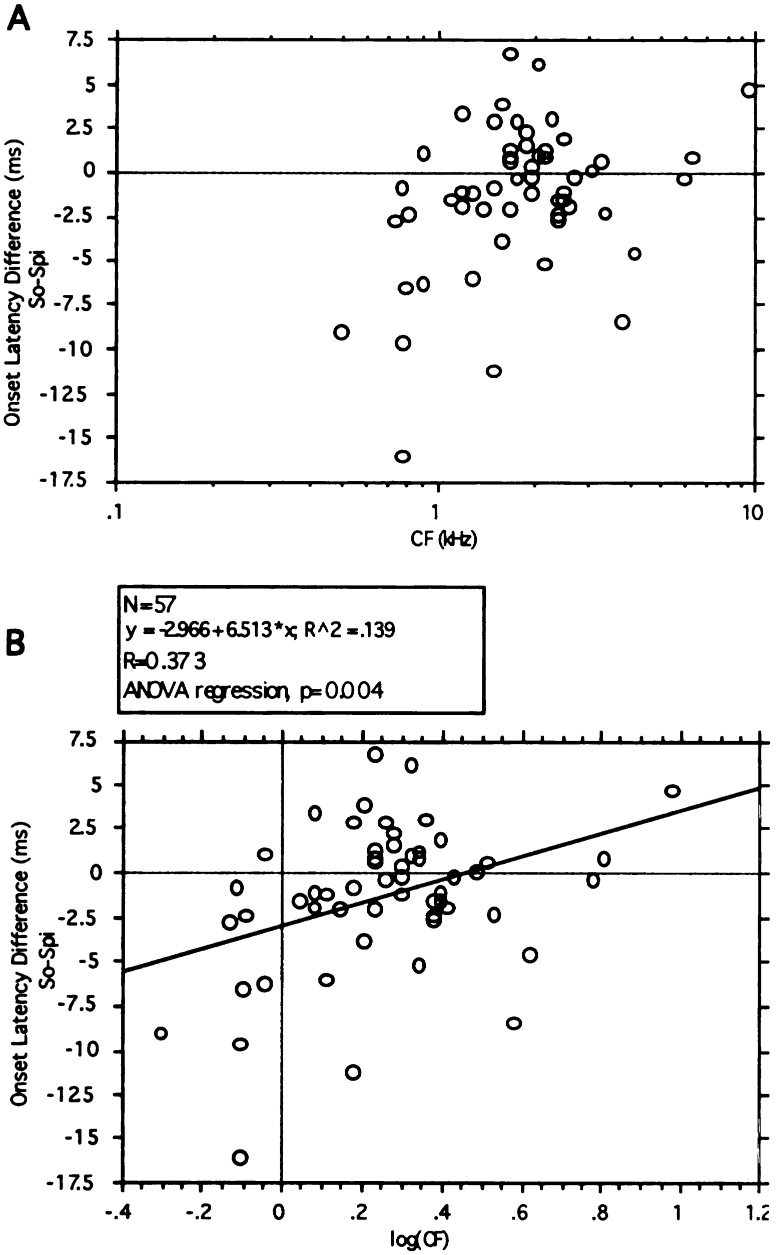


Figure 1.25 Interaural phase difference effect on onset latency

The distribution of onset latency difference is plotted in **A** as a function of CF (in kHz) and in **B** as a function of log frequency. Units with CF below 1.5 kHz or .18 in log frequency tended to exhibit shorter onset latencies in the homophasic condition.

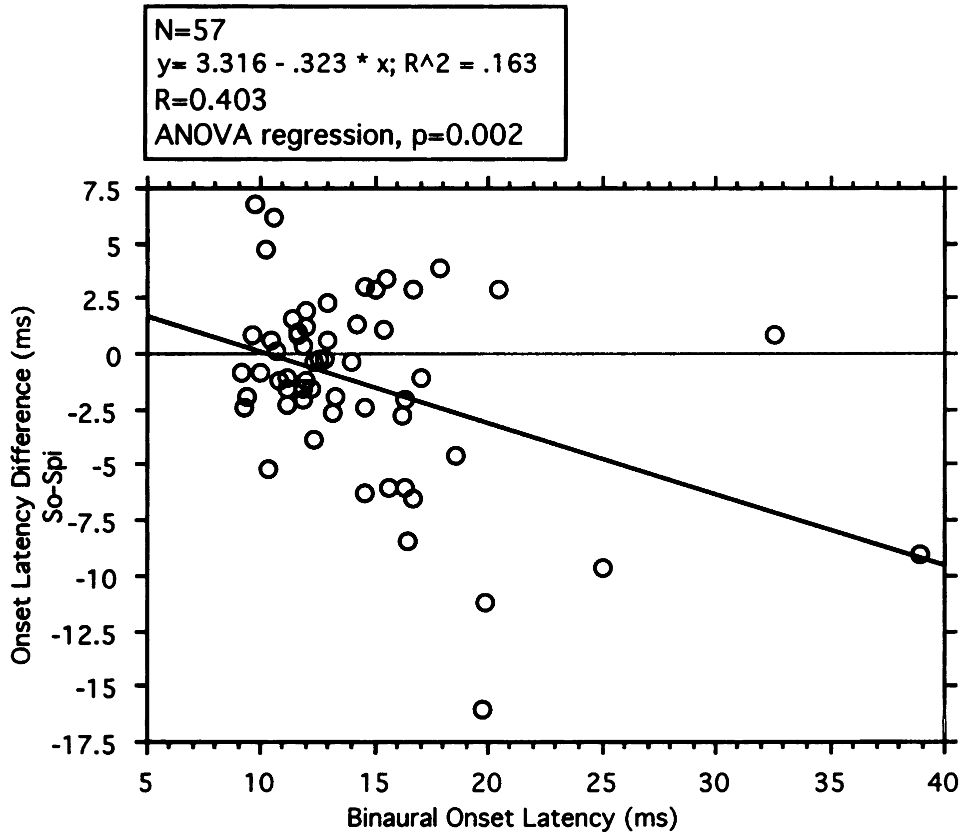


Figure 1.26 Effect of interaural phase difference and background noise on onset latency

The distribution of onset latency difference between two interaural phase different stimulus conditions in the presence of background noise is plotted as a function of initial onset latency in the absence of background noise. Regression correlation shows a significant negative correlation suggesting that units with longer onset latency in the absence of noise tend to manifest longer onset latencies in the antiphasic condition in the presence of noise.

differences between the two signal-in-noise conditions with onset response latency to tonal stimuli showed a significant inverse relationship. Longer latency units demonstrated even longer latencies in the antiphasic than the homophasic condition in the presence of continuous background noise.

1.3.0 Discussion

The goal in this series of experiments was three-fold, 1) to characterize differences in spectro-temporal tuning properties of low-frequency cortical units in AI that may arise from monaural (Sm) versus binaural (So) tonal stimulation; 2) to determine the effects of background noise on monaural and binaural receptive fields of cortical low-frequency neurons in AI, and 3) to examine the effect of binaurally-configured tones-in-noise on spectral tuning properties of low-frequency cortical neurons in AI. By investigating any inherent differences in receptive field properties of cortical neurons as a result of monaural and binaural tonal stimulation in at least one variant contextual listening environment (i.e., continuous background noise versus no background noise), one can attempt to achieve a basic understanding of how real life acoustic demands imposed by our dynamic listening environment are organized and processed in the auditory cortex.

Results include: 1) Overall, response thresholds were lower in binaural than in monaural receptive fields. 2) The amount of noise-induced threshold shifting was similar between monaural and binaural conditions. However, within the binaural condition, the amount of threshold shift was slightly greater in the homophasic than antiphasic tone-in-noise condition. 3) Low to moderate levels of continuous background noise induced significant onset latency delays in both monaural and binaural tonal stimulation conditions. 4) The amount of noise-induced onset latency delay was similar between monaural and binaural conditions. However, within the binaural condition, the amount of onset latency delay was greater in the antiphasic than homophasic tone-in-noise condition. 5) The difference in minimum onset latency derived from two different binaurally-configured stimuli was negatively correlated with binaural onset latency and positively correlated with CF. That is, high-CF, short latency neurons tended to show a prolonged latent period in the presence of continuous background noise, while low-CF, longer latency neurons tended to show a shortened latent period in the presence of continuous background noise.

6) Continuous background noise differentially changed receptive field bandwidths near threshold. Sharply-tuned neurons usually broadened in the presence of continuous background noise, while tuning of broadly-tuned neurons usually narrowed in the presence of continuous background noise. 7) Differences in receptive field bandwidths derived from two different binaurally-configured stimuli for sharply-tuned neurons were negatively correlated with neuronal response monotonicity. Sharply-tuned, non-monotonic neurons manifested sharper tuning in the antiphasic condition, while sharply-tuned, monotonic neurons manifested broader tuning in the antiphasic condition. Broadly-tuned neurons, irrespective of monotonicity, manifested broader tuning in the homophasic condition.

Given these findings, the conclusions can be drawn: 1) Continuous background noise differentially influences receptive field parameters. Induced changes correlate with basic subclassifications of neurons on the basis of their inherent neuronal response properties, such as bandwidth and monotonicity. 2) Differences in binaural receptive field properties can be elicited by the binaural configuration of signals-in-noise. 3) Continuous background noise exerts similar influences on monaural and binaural receptive fields.

1.3.1 Technical Considerations

EFFECTIVENESS OF MASKER. With regards to the level of the noise masker used in this series of experiments, the absolute level of the noise was adjusted by using an audiovisual criterion of response decrement to the masker, to yield 20 dB masked tone thresholds in the contralateral stimulation condition. This method of setting masking levels only allowed for an approximation of response masking and, as manifest in the actual response threshold shifts recorded in detailed analysis, often underestimated or overestimated the effectiveness of the masker. Although the targeted 20 dB shift was only realized as an averaged difference in the binaural condition and only slightly underestimated in the monaural condition, the data reflecting correlative trends between specific receptive field property can be interpreted as reflecting relative masking effects. Intracellular comparisons between monaural and binaural receptive field parameters including masked differences should not be biased by the range of effective masking. Likewise, intracellular comparisons of receptive field properties between the two binaurally-configured signal-in-noise condition should not be biased by the range of effective masking.

The nonlinear relationship between threshold shift and the absolute level of the noise masker, as seen in figure 1.3, merely reflects the variance in the tone-evoked neural sensitivity of the single units sampled. Once the absolute masker level was set to bring about a sensitivity shift, then further increments in the masker level should linearly match adjustments in the cell's sensitivity to tones as reported in other studies (Phillips and Cynader, 1985; Phillips, 1987). In the present study, this linear relationship was not reconstructed, since only one masker level was employed.

BARBITURATE ANESTHESIA. The use of barbiturate anesthesia in neurophysiological recordings has been documented for its reduction in the magnitude of evoked activity (Erulkar, Rose, and Davies, 1956; Kiang, Neame, and Clark, 1961, Teas and Kiang, 1964) and suppression of spontaneous activity in the cortex (Calford and Semple, 1995). Barbiturates act upon GABAa receptors which are the major inhibitory receptors in AI (Houser et al., 1984). Recordings of the excitatory FRAs in AI could, thus, be affected by the increase in inhibition which would be reflected in the spectro-temporal parameters reported in this study. However, though anesthetic effects on the overall response cannot be ruled out, the effect would have been consistent across stimulus conditions per unit; and thus, would not strongly influence the comparisons made in this study.

1.3.2 Monaural vs. Binaural Receptive Fields

Comparison of cortical response to monaural and binaural stimulation has been previously studied by varying tone stimulus intensity and recording spike activity as a function of stimulus level (e.g., Brugge, et al., 1969; Brugge and Merzenich, 1973; Phillips and Cynader, 1985; Phillips, 1995; Phillips and Hall, 1986). Analyzing spike activity as a function of stimulus level for a given frequency tone, commonly referred to as the rate-level function, allows a quick assessment of a neuron's response behavior, but is limited in describing other receptive field properties of a neuron such as frequency resolution and spectral sensitivity. In the current study, measurements of spectral response properties from FRAs provided more detail of similarities and differences in cortical neural responses between monaural and binaural stimulation by analyzing parameters of the entire excitatory FRA.

The results reported in this study indicate only small difference in receptive fields between monaural and binaural stimulation. The only binaural advantage found in the data was the lowered response threshold for low-threshold units. At the cortical level, the FRA, which is composed of multiple frequency inputs of varying sensitivity, may derive its inputs from one ear but more commonly from two ears. Convergence of multiple inputs with differing thresholds to cortical cells' CF is greatly expanded in the binaural condition as a result of heterotypic and homotypic corticocallosal projections (Rouiller et al., 1991) and thalamocortical projections from binaural interactive cells (Jones, 1984). Thus, the summation of excitatory post-synaptic potentials (epsps) could be greater in the binaural condition due to more convergence leading to a slightly more effective synaptic coupling and lower threshold response for ordinarily low threshold units.

1.3.3 Effects of stimulus interaural phase on the cortical representation of tonal stimuli

The effect of stimulus-interaural-phase on frequency selectivity in cortical low-frequency cells is not a subject to be overlooked when studying the effect of background noise on binaural stimulation. It is known from psychophysical studies on binaural masking that threshold level differences (called the BMLD) up to 15 dB can occur for low-frequency tones (2 kHz or less) (Hirsch, 1948) or for speech (Licklider, 1948). These threshold differences are brought about by the interaural phase inversion of a low-frequency tone masked by a simultaneously-gated broadband noise. Greater detectability of the signal occurs in the antiphasic signal-in-noise condition. In the current study, the physiological threshold difference found in AI neurons were on the average much less than those reported in psychophysical studies. The average response threshold difference was 3 dB lower in the antiphasic condition. Nonetheless, the dispersion in the data revealed that 55% of the units showed release from masking in the antiphasic condition. In addition, the dispersion in the data reveals some units having as much as 20 to 30 dB release from masking in the antiphasic condition. This indicates that low-frequency cortical cells do have the capability of demonstrating large sensitivity differences based upon interaural stimulus phase differences. It is possible that by increasing the level of effective masking more neurons would have had larger threshold differences as reported in psychophysical BMLD studies (Goldstein and Stephens, 1975; Zwicker and Zwicker, 1984). However, the goal of this study was not to measure BMLD thresholds in the cortex, but to determine what degree of impact interaural stimulus phase difference may have on binaural signal processing in continuous background noise masking at the cortical level.

A physiological basis for frequency sensitivity differences can perhaps be explained by AI neurons whose sensitivity is dependent upon interaural disparities in stimulus phase. Interaural phase dependent (IPD) cells, which have been described in AI (Benson and Teas, 1976; Brugge et al, 1969; Brugge and Merzenich, 1973; Orman and Phillips, 1984), are most common in cells with a CF less than 2-3 kHz. Cortical IPD sensitivity is generally displayed by cells whose binaural response types are EE (Reale and Brugge, 1990), although EO and OE response cells have also demonstrated maximal firing at a specific interaural time delay or "characteristic delay" (CD) (Rose et al., 1966; Yin and Kuwada, 1983b). Although there was no attempt to isolate IPD-sensitive cells in this study, it is likely that this subpopulation of cells was part of the sample and that differences in response threshold observed in the two binaurally-configured stimulus conditions may have been a reflection of CD-dependent response strength.

Another difference in frequency selectivity found for the two binaurally-configured stimuli was the differential change in excitatory bandwidth of only the sharply-tuned neurons that appears to be dependent on the monotonicity of the cell. The differential changes to bandwidth that are observed with two binaurally-configured stimuli may possibly be explained by convergence patterns onto AI neurons. Two different inhibitory mechanisms have previously been postulated to affect response area tuning curves in AI (Calford and Semple, 1995): 1) level-dependent inhibition resulting in reduction of spike firing at higher stimulus levels and 2) sideband inhibition affecting bandwidth tuning at lower stimulus levels. Presumably then, a sharply-tuned, non-monotonic neuron is being shaped by both of these inhibitory mechanisms while a sharply-tuned, monotonic neuron is only affected by the first of these two mechanisms. Therefore, the differential bandwidth finding that sharply-tuned neurons decrease or increase bandwidth tuning in one specific binaurally-configured stimulus condition depending on the monotonicity suggests an interaction of these two inhibitory mechanisms that might be tied into the phase dependency of the neuron. Subcortical projections to AI neurons may contribute to sideband inhibition specific to interaural phase disparities. In the inferior colliculus of guinea pigs, Jiang et al. (1997) reported that different populations of neurons responded differentially to So and Spi tones. Neural responses were measured by their increase or decrease in discharge rate to the binaurally-configured stimuli. Thus, the identification of interaural phase selective units in subcortical areas whose projections influence cortical frequency selectivity may provide some explanation to this difference in tuning bandwidth for sharply-tuned AI neurons. Broadband masking affecting phase-sensitive, intracortical convergence selectively upon the sharply-tuned, non-monotonic neurons could reduce sideband inhibition in the homophasic condition resulting in broader bandwidth. Broadband masking has previously been shown not to affect nonmonotonic and monotonic rate-level functions other than shifting the dynamic range of the response (Phillips, 1985, 1990). In this study, masking effect upon the excitatory and inhibitory processes related to sound pressure level was not analyzed. However, the previous work done by Phillips was limited to contralateral stimulation of the system and does not take into account possible binaural excitatory and inhibitory interactions that could affect processes related to sound pressure level.

1.3.4 Cortical receptive fields: the effect of continuous noise masking

Previous studies which looked at the effects of continuous noise masking on neural responses in AI have primarily analyzed rate-level behavior at CF and not the response area of the neuron to many frequencies of varying level. From rate-level data, it has been reported that the effects of continuous noise were 1) independent of other neural properties of the cell such as monotonicity and 2) remarkably homogeneous across cells (Phillips, 1985; Phillips and Cynader, 1985; Phillips and Hall, 1986; Philips, 1990). Phillips and his colleagues have found that neurons in AI do not exhibit sustained activity in the presence of continuous noise. The main effect of the continuous noise was to elevate tone thresholds and prolong latent periods in a linearly dependent manner with SPL of the masking stimulus.

Despite methodological differences, similar findings in threshold and minimum onset latency are reported in the current study for both monaural and binaural stimulation conditions. However, it was observed that the presence of continuous background noise appears to evoke profound changes in the excitatory bandwidth of AI receptive fields. While global noise-induced changes in bandwidth were not detected when averaging across all studied cells, continuous noise masking revealed an inversion effect on bandwidth tuning when cells were subdivided based upon sharpness of tuning. This inversion effect, which was observed at stimulus levels that were 10 dB and 30 dB above threshold, was present in both monaural and binaural stimulation conditions.

The significance of this observation is underscored by three points. First, the fact that bandwidth inversion with continuous background noise is observed at both 10 dB and 30 dB above threshold is evidence for a noise effect on excitatory and inhibitory processes affecting sharpness of tuning and not just random fluctuations of near- threshold responses. Second, despite small differences in bandwidth reported for binaural stimuli with interaural phase differences, the effect of continuous noise on bandwidth outweighs the effect stimulus interaural phase. Third, since the bandwidth inversion effect is observed not only in the binaural conditions but also robustly in the contralateral stimulation condition, any mechanisms involving binaural interactions are secondary to the mechanism influenced by continuous background noise.

One possible explanation for this differential masking effect in excitatory bandwidth may be found by considering the multiple convergent inputs upon AI neurons and perhaps the selective elimination of convergent inputs with continuous background noise. A previous tone-on-tone masking study (Phillips and Hall, 1992) described, in a small portion of their data, a differential masking effect in response firing to continuous tonal maskers dependent upon frequency location of the tonal masker with respect to the cell's response area. Monotonic units showed selective suppression of responses to tone signals on the same side as the masker. Non-monotonic units, on the other hand, showed enhanced responses to tonal signals on the same side of the response area as the masker when tonal maskers were placed at the high-frequency edge of the excitatory response area. They interpreted this differential sensitivity to the masker as evidence for independent inputs to the high- and low-frequency sides of the excitatory response area.

In the present study, the increase or decrease in excitation bandwidth in the presence of a continuous broadband masker also suggests selective elimination of convergent inputs to the cell. The FRA of a central neuron is composed of multiple frequency inputs with varying sensitivity. Although these neural responses were recorded in layer 3 which is principally composed of pyramidal cells with incoming thalamic projections (Imig and

Morel, 1984), intrinsic cortical connections within isofrequency borders (Read, per communication) or extrinsic corticocortical connections (Imig and Reale, 1980; Rouiller et al., 1991) may act to locally modulate the neural response via inhibitory inputs. In broadly-tuned neurons, a low-level broadband noise masking neural activity in the lower portion of a cell's response area could selectively eliminate less sensitive sideband excitatory inputs that would result in a narrowing of frequency selectivity. In sharply-tuned neurons, which is presumably a result of sideband inhibition, a broadband masker could selectively eliminate those inhibitory inputs which would cause a broadening in excitation bandwidth of the response area.

1.3.5 Critical Band Effect in AI?

As indicated by the results of this study, there was no significant difference in the noise-induced threshold shift, the bandwidth change, or onset latency shift between the two binaural stimulation conditions (from table 1.1D). Therefore, it appears that the masking effects of continuous noise backgrounds under contralateral tone-in-noise and bilateral tone-in-noise stimulation exert the same amount of change in the spectral sensitivity of AI neurons. A behaviorally relevant measure of spectral integration capacity is the critical band. Previous work (Ehret and Schreiner, 1997) found critical bandwidth behavior that matched psychophysical behavior in some AI neurons, located only in the central and ventral regions of the isofrequency lamina. Only those neurons showed properties consistent with neurophysiological measurements in the inferior colliculus (Schreiner and Langner, 1997; Ehret and Merzenich, 1988) and with behavioral studies. Since critical bandwidth as measured in AI is a direct measure of frequency resolution and spectral filtering, one can extrapolate the current findings with those from psychophysical studies done in humans. The similarity in spectral sensitivity between monaural (Sm) and binaural (So) stimulation conditions in the presence of continuous background noise appears to be consistent with psychophysical data which found critical bandwidth estimates of SmNm and SoNo to be equivalent (Sever and Small, 1979).

The observation of latency shifts due to noise masking has been seen in many stages of the auditory system (in AI: Phillips, 1985). In this study, the minimum onset latency of spike response appears to be dependent upon the binaural configuration of the tone in continuous noise. Homophasic signals in noise manifest shorter latencies on average than antiphasic signals in noise. The difference in onset latency between the two binaural conditions appears to be the greatest in units with CF less than 1 kHz which coincides with the frequency range where timing difference effects between the two ears are the greatest.

Minimum onset latency in neurons can be considered an indicator of neural integration time and can also be viewed as an indicator of the lower boundary of temporal integration as represented in AI. Psychophysical data from van den Brink and Houtgast (1990) revealed that efficient temporal integration in signal detection is confined to a narrow frequency window which corresponds to the critical bandwidth. The critical bandwidth measured in the antiphasic signal-in-noise condition is broader than in the homophasic signal-in-noise condition (Sever and Small, 1979). Therefore, the difference in integration time could be attributed to the difference in critical bandwidth. The longer integration time measured in the antiphasic condition could be a result of a broader critical band.

1.3.6 Implications of binaural interactions on the masking of complex sounds

Mechanisms for binaural interactions exhibited by single auditory neurons of the brain stem have been postulated to serve an important role in mammalian sound-localization behavior (Yin and Chan, 1990; Yin and Kuwada, 1983). How binaural interactions, originating in the brain stem and then transmitted upstream, affect cortical receptive fields in AI are still at best hypothetical. Coincident detection models formulated for neurons in the inferior colliculus (IC) have proposed possible methods in which auditory neurons can physiologically process monaural and binaural signal interactions (Jeffress, 1948; Yin, Chan, and Carney, 1987) that are relevant to many theories of binaural hearing (Colburn, 1977; Jeffress, et al., 1956; Osman, 1971; Sayers and Cherry, 1957; Webster, 1951). Psychoacoustical studies of BMLD using decorrelated signals and/or noise to the two ears have shown that the magnitude of the signal or noise correlation is the dominant factor in determining BMLD threshold (Durlach et al., 1986; Hirsch, 1948; Jeffress, Blodgett, and Deatherage, 1952; Langford and Jeffress, 1964; Licklider, 1948). Onset latency differences reported in the current results between binaurally-configured stimuli may also change as a function of timing delays introduced by a cross-correlator in the auditory brain stem (Yin, Chan, and Carney, 1987). Therefore, it is plausible that sound-localizing capabilities in AI neurons are contributing to spectral sensitivity changes invoked by variation in background listening environments.

It would be hard to believe though that the spectral sensitivity changes observed in the binaural stimulation cases can entirely be accounted for by a subcortical cross-correlator, especially given physiological evidence for other forms of binaural interactions that occur in cortex. Tracer-injection studies have provided anatomical evidence for cortical convergence from cortico-cortical connections (Code and Winer, 1985; Code and Winer, 1986; Imig and Reale, 1981; Imig and Brugge, 1978) that may play a role in the modulation of excitatory and inhibitory processes forming the binaural receptive field areas in AI neurons. Although callosal innervation has been interpreted as not being essential in shaping the binaural properties observed in AI since many of the cells that are callosally connected are EE cells (Imig and Brugge, 1978), intracortical connections between classes of neurons categorized by their sharpness of tuning (Read, per communication) may be instrumental in the modulation of receptive field properties. Since only 15% of the synaptic terminals in layer 4 of primary auditory cortex originate from thalamic inputs (LeVay and Gilbert, 1976), it is conceivable that cortico-cortical connections exert an influence in shaping spectral sensitivity in a variety of contextual backgrounds.

1.3.7 Closing remarks

The effects of masking on the spectral sensitivity of AI neurons can be considered to be a result of complex interactions of thalamocortical projections from cells that may be performing binaural computations of sensory inputs, and cortico-cortical intrinsic and extrinsic connections, whose role may be in sharpening or broadening frequency selectivity and sensitivity. Psychophysical explanations of signal detection and recognition contemplating critical band-like behavior in AI neurons can perhaps account for some physiological behavior observed in this study; however, knowledge of spectral sensitivity differences based upon sharpness of tuning characteristics may limit the utility of modeling cortical neural responses to binaural complex sounds based upon theory of critical bands. Further understanding of spectral sensitivity to binaural complex sounds may be gained, instead, by searching for patterns of similar neural behavior among acknowledged functional organizations in the primary auditory cortex.

Chapter 2: A multi-unit study of the effects of continuous background noise on the cortical representation of tonal stimuli and its topography

2.0.0 Introduction

Auditory neurophysiological research over the last fifty years has uncovered several organizational features in the primary auditory cortex (AI) that provide a basis for understanding how sounds are represented in the early stages of cortical processing. The earliest and most prominent organizational feature that has been observed in cat AI and at each processing level of auditory system beginning at the cochlea is tonotopicity (Merzenich et al., 1975). Neurons with similar characteristic frequencies (CF) are organized in isofrequency lines whose low-to-high frequency gradient spans across the posterior-to-anterior axis of the cortex.

Since then, mapping studies in cat AI have found gradients of other response parameters (e.g., binaural interaction, excitatory bandwidth, monotonicity) along the isofrequency dimension confirming an earlier hypothesis (Tunturi, 1950; Tunturi, 1952) that a secondary organization may exist orthogonal to the frequency representation. The gradient of sharpness of tuning (Schreiner and Mendelson, 1990) for excitatory bandwidth near threshold level (Q10dB) reveals broadly-tuned neurons located along the extremes of the dorsal and ventral extent of the isofrequency dimension and sharply-tuned neurons located in the dorsal-ventral center. Inter-related with sharpness of tuning is monotonicity. The spatial segregation of neurons based upon monotonicity (Schreiner, Mendelson, and Sutter, 1992; Sutter and Schreiner, 1995) reveals non-monotonic neurons located in the dorsal-ventral center (similarly located with sharply-tuned neurons) and in the dorsal third of AI. Related to monotonicity is the evidence for topographic representation of tone intensity along the isofrequency lamina (Heil, Rajan, and Irvine, 1994; Phillips, et al., 1994). Spatial distribution of response latency orthogonal to the frequency dimension has also been found such that shorter latency responses were typically found in the central, more sharply-tuned region while longer latency responses were frequently seen in the dorsal and ventral ends of AI (Mendelson, Schreiner, and Sutter, 1997). In addition to these functional topographies discovered primarily through monaural stimulation techniques, binaural stimulation has revealed the existence of binaural interaction bands (Imig and Adrian, 1977; Middlebrooks, Dykes, and Merzenich, 1980) and a gradient of aural dominance along the isofrequency domain (Reale and Kettner, 1986).

Although these multiple functional topographies are known to exist in AI, there is little knowledge of how robust these organizations are to perturbations in the acoustic context such as changes in the signal-to-noise ratio from the addition of a constant level of continuous background noise. Single unit studies in AI have revealed continuous noise-induced changes in frequency selectivity affecting threshold, onset latency, and dynamic range (Phillips and Cynader, 1985; Phillips and Hall, 1986). Furthermore, they report that these changes are no different for either monotonic or non-monotonic neurons (Phillips and Cynader, 1985; Phillips, 1990) whose functional classes are known to be spatially segregated in AI (Schreiner, Mendelson, and Sutter, 1992; Sutter and Schreiner, 1995).

An auditory evoked potentials (EP) study measuring the effects of continuous background noise on response amplitude and latency in AI (Phillips and Kelly, 1992) found a wide range of response variability in the shapes of EP input-output functions for tones of similar frequencies recorded in different animals. The variability in response was attributed to biased averaging of activity in the general vicinity of the electrode. Given that spatial segregation of functional attributes exist within isofrequency lines, Phillips concluded that the location of the recording electrode is a significant factor contributing to their observed response variability. However, these previous studies do not address whether the functional topography along each isofrequency line changes with continuous background noise.

The significance of having spatially segregated, functionally-distinct classes of neurons along isofrequency lines would be achieved if this organization formed some sort of physiological basis allowing for the perceptual invariances and generalizations made of complex sounds such as speech. Moreover, the study of how perturbations in the acoustic context affect the spatial segregation of subpopulations of neurons could reveal how AI may react physiologically to compensate for these changes in acoustic environment. From psychophysics, the critical band theory has been referred to in the explanation of noise masking effects (Fletcher, 1940) as well as in understanding the auditory system's ability to filter and shape the perception of speech sounds (Sharf, 1970; Zwicker, 1961). In chapter 1, the concept of AI neurons demonstrating critical band-like behavior was brought up in relation to single unit response property changes with continuous background noise. However, it was thought that spectral sensitivity differences based upon sharpness of tuning characteristics may limit the utility of modeling cortical neural responses to binaural complex sounds based upon the theory of critical bands. It was proposed that preservation of elements essential for perceptual invariance may be found in the aggregate response as opposed to the individual response. Therefore, the aim in the current series of experiments is to study the effect of a constant level of continuous background noise on the spatial distribution of binaural receptive field properties of neurons in AI.

2.1.0 Methods

2.1.1 Surgery

The surgical procedures followed for the multi-unit recordings were the same as for the single-unit recording study. A total of seven cats were used in this study. The primary auditory cortex from the right cortical hemisphere was studied.

Initially, cats were anesthetized by an intramuscular injection of ketamine HCL (10 mg/kg) in combination with the tranquilizer acepromazine maleate (0.28 mg/kg). Following cannulation of the brachial vein in the right foreleg, sodium pentobarbital (initially 25-29 mg/kg; titrated to effect) was administered with an IV injection to begin the surgical procedures. An injection of dexamethasone (0.14 mg/kg) to prevent brain edema together with atropine (1 mg/kg) to reduce salivation was administered subcutaneously. After attaching the EKG to monitor the heart rate and respiration and inserting the rectal temperature probe to monitor the animal's body temperature (which was maintained around 37.5 °C by wrapping a temperature-controlled heating pad around the animal's lower body), a tracheotomy was performed and a semisoft tracheal tube introduced to facilitate unobstructed breathing. The tracheotomy was also important for reducing sounds from breathing through an open mouth while the head of the animal was held by a standard orbital-maxillar-palatal holder.

For the craniotomy, the skin over the medial/posterior half of the skull was reflected and the bone over the primary auditory cortex was removed. Following the resection of the dura, a film of silicone oil was applied to protect the cortical surface, and for the multi-unit recording experiments only, a photographic image of the cortical vasculature and sulci was recorded as a reference for the topographical location of the electrode penetration sites. In single unit recording studies, agarose was used to cover the exposed cortical surface to assist in maintaining the stability of the recording site.

After surgery, subsequent delivery of the anesthesizing agents during the recording study were infused intravenously. The anesthetic agent used in the single unit recordings was sodium pentobarbital (titrated to effect); for the multi-unit recordings, ketamine (1-2 mg/kg) with diazepam (1-2 mg/kg) was used. During surgery and stimulus-response recordings, cats were maintained at a surgical level of anesthesia evaluated by on-line monitoring of vital signs including heart rate, body temperature, and flexion and corneal reflexes.

2.1.2 Neural Recordings

Extracellular recordings from each sampled location in the auditory cortex were obtained through the insertion of tungsten microelectrodes (1 to 2 Mohms) mounted on a single hydraulic microdrive (Kopf) while presenting acoustical stimuli to the ears. Triggered spike recordings were pre-amplified (Differential Amplifier DAM-50), bandpass-filtered (1-3 kHz) (Dual Variable Filter, Type VBF 8), windowed through a spike discriminator (BAK Dis-1), amplified, sampled at a temporal resolution of 30 μs, and stored in a computer (IBM compatible) for off-line data analysis.

During the recording session, triggered spikes were visually displayed on a storage oscilloscope to set the criterion of the window discriminator for accepted spikes. Recording electrodes were lowered into the cortex approximately perpendicular to the cortical surface via a remotely controlled hydraulic microdrive to depths of 700 to 1100 microns.

2.1.3 Stimuli

Experiments were conducted in a double-walled sound-treated chamber (IAC). Acoustic stimuli were delivered through earphones (STAX 54) sealed in the ear canals. The stimulus set used was a series of 675 tone bursts varying in intensity and frequency which were used to assess the spectral sensitivity and specificity of auditory neurons. Tones were generated by a microprocessor (TMS 320C30; sampling rate of 110 kHz, 78 dB dynamic range). The intensity of the tones varied in 5 dB steps over a range of 70 dB, while the 45 different frequency values covered a range of 4 to 5 octaves centered around the estimated CF of the unit. Each pure tone signal was 50 ms in duration, ramped with a 3 ms (linear) rise/fall time, and repeated in a pseudorandom order every 350 ms. The pure tone series were summed with or without continuous noise depending upon the stimulus condition.

In these experiments, there were two auditory stimulation conditions: 1) pure tone series delivered in phase to both ears in the absence of continuous background noise and 2) and pure tone series delivered in phase to both ears in the presence of continuous background noise. The continuous background noise was white random noise (Random Noise Generator, Type 1390-B) lowpass filtered at 10 kHz. The absolute level of the continuous noise was fixed at 65 dB SPL (-27 dB/Hz); the absolute level of the pure tones in the series ranged from 50 to 120 dB SPL.

2.1.4 Data Analysis

Spectral sensitivity differences between the two different stimulus conditions were assessed by using parametric statistical measures. Spectral and temporal response parameters were measured from pure tone tuning curves recorded at each penetration site within the primary auditory cortex. The tuning curves analyzed were depictions of the excitatory receptive field of each recorded unit.

The pure tone response parameters measured included: 1) characteristic frequency (frequency with the lowest response threshold); 2) minimum threshold (the lowest tone intensity whose frequency matches the unit's characteristic frequency which elevated the discharge rate); 3) lowest and 4) highest frequency at 10 dB, 20 dB, and 30 dB above threshold to which the neuron responded (the lowest and highest frequencies measured at the three intensity levels were used to calculate the spectral bandwidth and

Q=quality factor of neurons.); 5) minimum onset latency estimated from the neural discharge of multiple level tones whose frequency matches the unit's characteristic frequency; 6) and monotonicity of the rate-level function measured as the slope post-transition point.

2.1.5 Functional maps

Receptive field parameters including threshold, bandwidth, and minimum onset latency were reconstructed spatially across the frequency gradient and within each isofrequency contour across cortical area AI. The spatial distribution of response parameters across the cortical surface was generated by using gridding software (Surfer, Golden Software) that takes into account the actual spatial locations of the recording sites and makes a two-dimensional grid of the data. The third dimension represents the magnitude of the response feature measured. An inverse-distance, squared algorithm was used to interpolate among the four closest neighboring points within a quadrant. The interpolated map of the gridded data for each receptive field parameter were then displayed using graphics software (Spyglass Transform).

2.2.0 Results

2.2.1 Binaural Multi-unit Receptive Fields and Topographies

By studying the receptive field of an aggregate of cells and the relationship of each aggregate's receptive field property to one another, one can begin to examine some aspects of the properties of the functional organization in the primary auditory cortex. In chapter 1, the effect of background noise set upon tonal selectivity was studied in an attempt to understand how frequency selectivity changes when stimuli are presented in a low-level continuous noise listening context. Although the methods that were applied allow for a survey of the frequency response area (FRA) or receptive fields under the influence of background noise, it only mirrored the individual unit's response attributes to the stimuli and fell short of providing a reconstructable basis for defining distributed neural assembly responses to tones embedded in continuous background noise. By delivering tonal stimuli embedded in a fixed-level of continuous background noise and recording multi-unit responses within and across the iso-frequency axis, one can begin to study the robust functional topographies of the primary auditory cortex. In this set of experiments, the effect of a fixed-level of continuous background noise was studied by recording binaural multi-unit FRAs (N=256) from the primary auditory cortex of five adult cats. In addition, the cortical coordinates of the binaural FRAs and the corresponding tuning curve data from seven adult cats were recorded to reconstruct topographies of the measured parameters. Spectral parameter topographies reconstructed from tonal stimulation tuning curves were then compared with those from tonal stimulation with background noise tuning curves to determine whether or not functional topographies are affected by their contextual environment. The level of the continuous background noise was chosen to be -27 dB/Hz, which has been shown in our single unit studies to produce an effect in the majority of units.

An overview of the general trends found for the homophasic tonal stimulation with and without background noise may be gained by examining the general averages of response parameters for all the units sampled. Again, the original goal was to study the effect of background noise on neural behavior in low-frequency cells in the primary auditory cortex. The frequency range of units obtained was primarily between 600 and 2500 Hertz (20% of the unit CFs were above 2500 Hertz). The effect of a fixed-level of background noise on CF can be seen in figure 2.1, which is a scatterplot of CF measured in both no-noise (tonal stimuli are homophasic in both stimulus conditions) and noise conditions. The regression coefficient (r=.74) approximates the unity line for the two variables. Since CF appears to be the same with or without the background noise, future reference to CF will be for the measurement made in the no-noise condition. Table 2.1 represents the mean and standard deviation (shown in parentheses) for the following parameters measured in either quiet or noisy backgrounds: threshold, Q10 (quality factor for bandwidth measured at 10 dB above threshold), Q40 (quality factor for bandwidth measured at 40 dB above threshold), Q65 (quality factor for bandwidth measured at 65 dB above the lowest tonal stimulation level), and minimum onset latency. Since the purpose of the group statistics was to search for significant changes in receptive field property with the application of background noise for each multiple-unit sample, a within-subjects ANOVA (repeated measures) was used to test the null hypothesis that there were no differences within the cell's frequency selectivity in the presence of background noise. Three receptive field parameters showed a difference that was statistically significant ($p < 0.0001$). These three receptive field parameters were threshold, Q65, and minimum onset latency. On average, thresholds were elevated by 11.6 dB, the Q-factor at 65 dB above the lowest tone stimulation level was 0.3 larger, and minimum onset latency was 1.7 ms longer in the presence of background noise. When all the unit locations were averaged together, Q10 and Q40 showed no statistical difference between stimulation conditions.

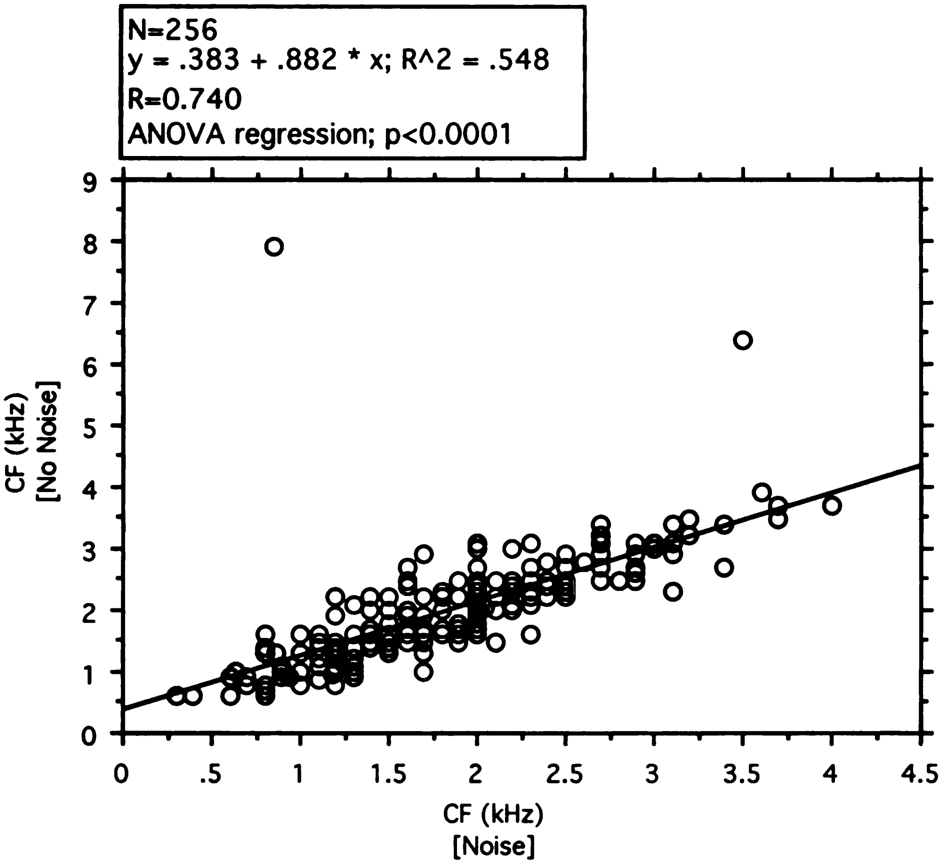


Figure 2.1 Distribution of CF

The distribution of CF measured from FRAs recorded in the presence and absence of continuous background noise was similar under both stimulus recording conditions. The sampled multiple unit responses ranged in CF from 600 Hz to 8 kHz in the absence of noise.

Table 2.1 Means Table for Receptive Field Parameters

All units are represented.

		No-Noise	Noise
threshold ***	N=256	24.0 (10.5)	35.6 (8.3)
Q10	N=254	2.9 (2.8)	2.7 (2.6)
Q40	N=133	1.0 (1.3)	0.9 (0.4)
Q65 ***	N=245	0.9 (0.4)	1.2 (0.7)
onset latency ***	N=245	12.3 (2.8)	14.0 (3.4)

Levels of Significance

*** < 0.0001

Table 2.2 Means table split by broadly- and sharply-tuned units

		Q10<2			Q10 =/> 2	
		No-Noise	Noise		No-Noise	Noise
threshold **	N=118	26.7 (10.9)	36.1 (8.8)	N=138	21.7 (9.7)	35.1 (7.9)
Q10 ***	N=118	1.4 (0.4)	2.2 (2.4)	N=136	4.3 (3.3)	3.2 (2.8)
Q40	N=60	1.0 (1.7)	0.9 (0.5)	N=73	0.9 (1.0)	1.0 (0.4)
Q65	N=112	0.8 (0.4)	1.1 (0.8)	N=133	1.0 (0.4)	1.3 (.6)
onset latency	N=110	12.7 (3.6)	14.3 (3.8)	N=135	12.0 (1.8)	13.7 (3.0)

Levels of Significance

** < 0.01
*** < 0.0001

When the cells were grouped by sharpness of tuning based upon the cell's Q10, a significant interaction effect was revealed for receptive field parameters expressed as threshold and Q10. Table 2.2 shows the results for the averaged parameter values obtained for sharply- and broadly-tuned units in primary auditory cortex. Although the average calculated threshold difference between stimulation conditions already rejects the null hypothesis, the threshold difference was significantly ($p < 0.001$) larger by 4 dB for the sharply-tuned neurons (Q =/> 2) than for the broadly-tuned neurons. As for Q10, whereas there was no significant difference for this receptive field parameter when all the units were averaged together, the Q10 difference measurement changed differentially ($p < 0.0001$) for the two groups characterized by their sharpness of tuning to tonal stimuli. In response to continuous background noise, broadly-tuned units became narrow while sharply-tuned units became broad in their tuning near threshold. There was no observed significant difference in Q40, Q65, or minimum onset latency when units were split by their sharpness of tuning.

When the entire sample of multi-units was subdivided into two groups based upon their CF characteristic, a significant interaction effect, as reported in table 2.3, was again revealed for the receptive field parameters expressed as threshold and Q10. A significant interaction effect was also found for minimum onset latency when cells were subdivided by their CF. The subdivision based upon CF yielded two groups of cells classified as low-frequency (CF =/< 2 kHz) and high-frequency (CF > 2 kHz). Although all the cells showed an elevation in threshold with the addition of continuous background noise, the high-frequency cells showed a larger increase (by 5.4 dB) in threshold compared to the low-frequency cells. As for Q10, the low-frequency cells, which tend to be broader in bandwidth tuning than high-frequency cells to begin with, showed an increase in Q10 with the addition of background noise while the high-frequency cells showed a decrease in Q10 ($p < 0.05$). With regards to the onset timing of spike response to tonal stimulation, minimum onset latencies were more prolonged

Table 2.3 Means table split by low- and high- CF units

		CF =/< 2			CF > 2	
		No-Noise	Noise		No-Noise	Noise
threshold ***	N=126	26.2 (10)	35.0 (7.4)	N=130	21.9 (10.6)	36.1 (9.1)
Q10 *	N=124	2.1 (1.5)	2.4 (2.5)	N=130	3.7 (2.5)	3.0 (2.6)
Q40	N=70	0.9 (0.3)	0.9 (0.4)	N=63	1.1 (1.9)	1.0 (0.4)
Q65	N=121	0.9 (0.4)	1.1 (0.8)	N=124	1.0 (0.4)	1.3 (0.6)
onset latency **	N=119	12.8 (3.5)	13.9 (3.4)	N=126	11.8 (1.8)	14.1 (3.3)

Levels of Significance

* < 0.05
** < 0.01
*** < 0.0001

in the high-frequency cells by 1.2 ms than the averaged minimum onset latency difference observed in the low-frequency cells. No significant interaction effect was observed for Q40 and Q65.

To summarize the general group findings, it appears that threshold and minimum onset latency do not merely change with the addition of background noise, but that the amount of change may be reflected differentially based upon their neural subclassification (i.e., sharpness of tuning and CF). Sharpness of tuning as reflected in Q10 did not appear to be affected by background noise when the multi-units were averaged as an entire group; however, subpopulations of the neurons showed significant differential effects to the presence of background noise.

Given this finding that changes in receptive field parameters due to a change in contextual environment may be observed in differing degrees relative to the neural subclassification of the cell, it was hypothesized that these differences in neural assembly would be reflected in the spatial topography of these receptive field parameters. The spatial topographies of threshold, sharpness of tuning (Q10), and onset latency were measured and reconstructed for seven adult cats. Comparison of the spectral parameter maps for binaural tonal stimulation in quiet and background noise conditions revealed a non-homogeneous change across and within iosfrequency boundaries which will be discussed further in the following sections.

2.2.2 Threshold

The threshold response of the FRAs was defined as the lowest tone stimulation level yielding a neural spike response. The range of thresholds measured for the 256 multi-units in the tonal stimulation condition without background noise extended from 2.5 to 62.5 dB SPL with a peak number of units having a threshold at 22.5 dB SPL as depicted in figure 2.2. The lower end of the distribution of threshold measurements shifted up with the peak increase of 10 dB when a constant -27 dB/Hz level of

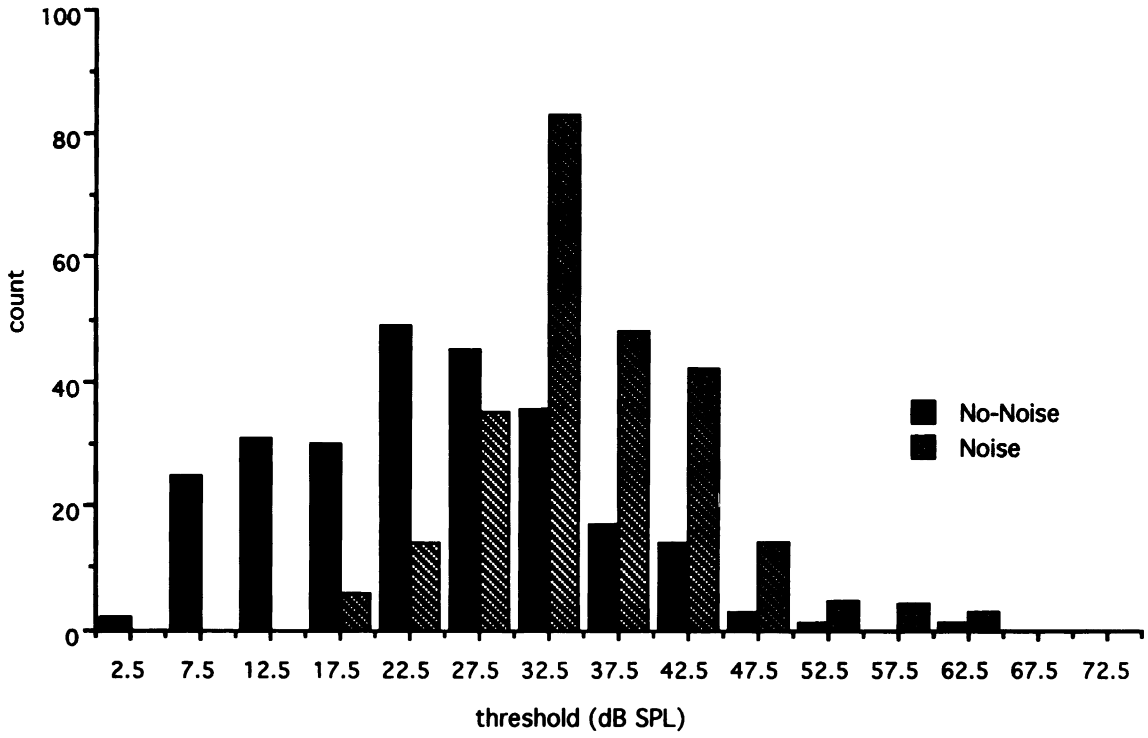


Figure 2.2 Distribution of thresholds

The distribution of thresholds measured in FRAs recorded in the presence and absence of continuous background noise revealed an upward shift in thresholds with the addition of noise, as was the objective.

background noise was present during tonal stimulation. A scatterplot comparison of thresholds, depicted by figure 2.3, measured in both stimulus conditions showed a regression line with a slope of 0.3 ($R=0.4$, $p < 0.0001$). Figure 2.3 shows that thresholds measured in the noise condition tended to be higher than those in the no-noise condition, however, there was a wide scatter in the noise-induced threshold shift. The effect of background noise on threshold was represented by the threshold difference which measures the amount of change attributed to the presence of background noise. The distribution of threshold difference shown in figure 2.4 illustrated a predominantly positive shift in threshold, and was similar for broadly and sharply-tuned neurons, although there tended to be more broadly-tuned units at CF levels lower than 1 kHz. The maximum threshold difference at each CF appeared to increase as CF level increased.

Inspection of the spatial topography of threshold found that five out of the seven experimental cases showed no change in the landscape of neural sensitivity relative to neighboring neural sensitivities. The changes seen in threshold measured in the presence of a low level of background noise merely reflected the differing degrees of threshold shift exerted by the constant noise level, but did not appear to affect the relative positions of areas with high or low neural thresholds. Figures 2.5A-G display a grid of cortical penetrations illustrating the CF gradient for each cat and the spatial topographies for response threshold in quiet and noisy backgrounds and the threshold difference between noise and no-noise condition. In figure 2.5A (cd394), the CF grid shows 138 penetrations in a 6 x 5 mm region of auditory cortex whose frequency representation extended from about 1 to 4 kHz. The tone-evoked neural thresholds found in the no-noise condition ranged from mid- to high-thresholds with a localized region of the lowest thresholds (in purple) in the central region crossing several CF contours; higher threshold units (in blue-green) flanked this central region along each iso-frequency boundary. With background noise, the thresholds were elevated in the

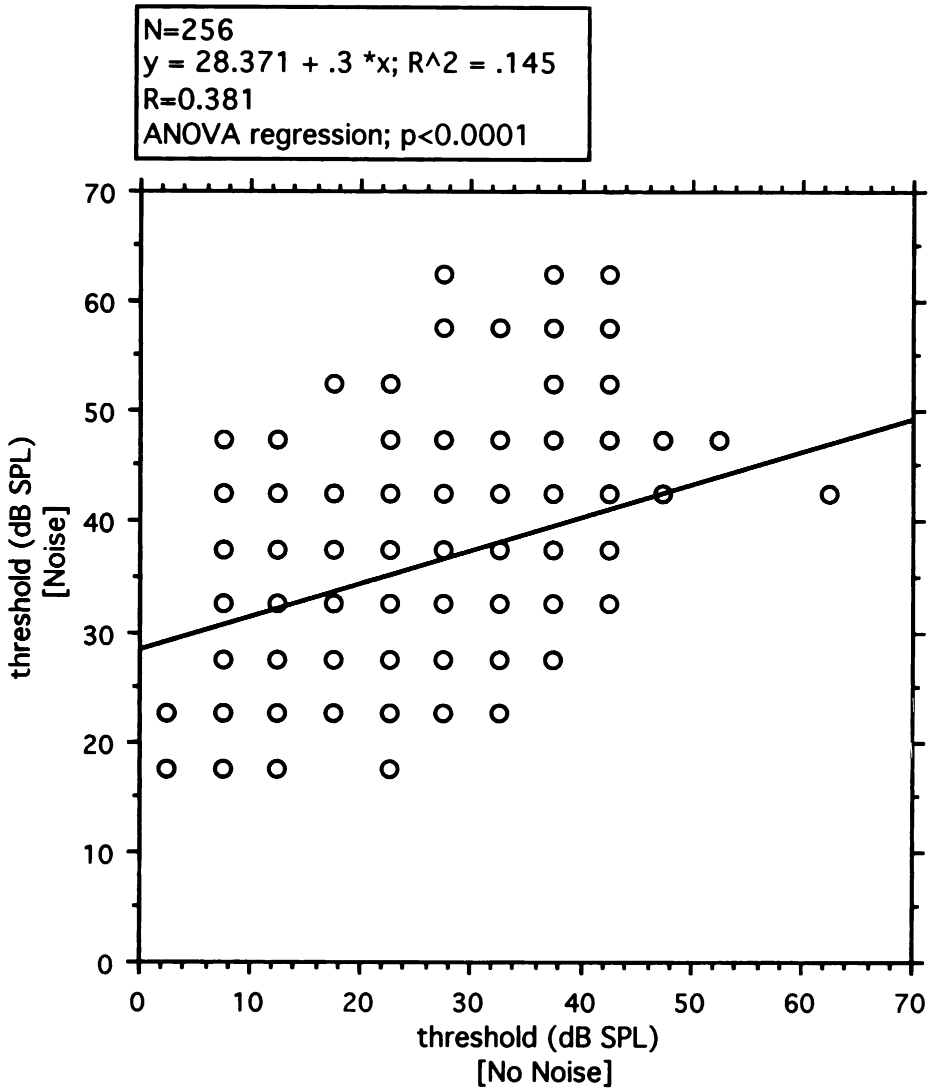


Figure 2.3 Scatterplot of thresholds measured in quiet and noisy backgrounds

Multiple-unit FRA thresholds obtained in the presence of background noise are plotted as a function of thresholds in the no noise condition. Linear regression analysis of the thresholds revealed a significant positive correlation (R=.38; p<.0001) between the two stimulus conditions.

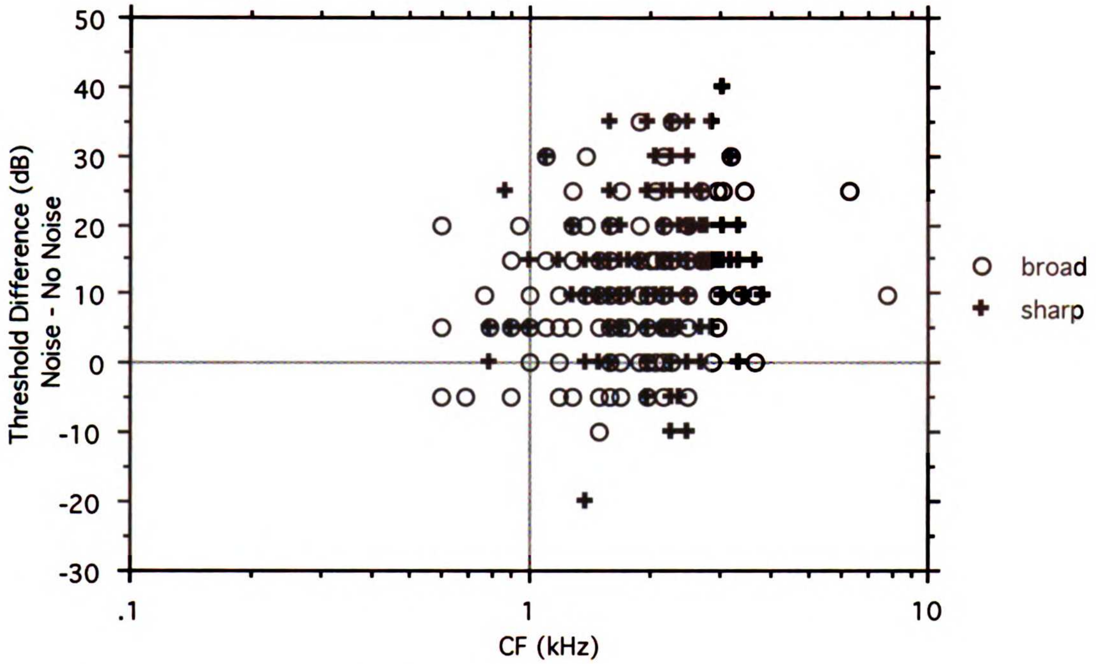
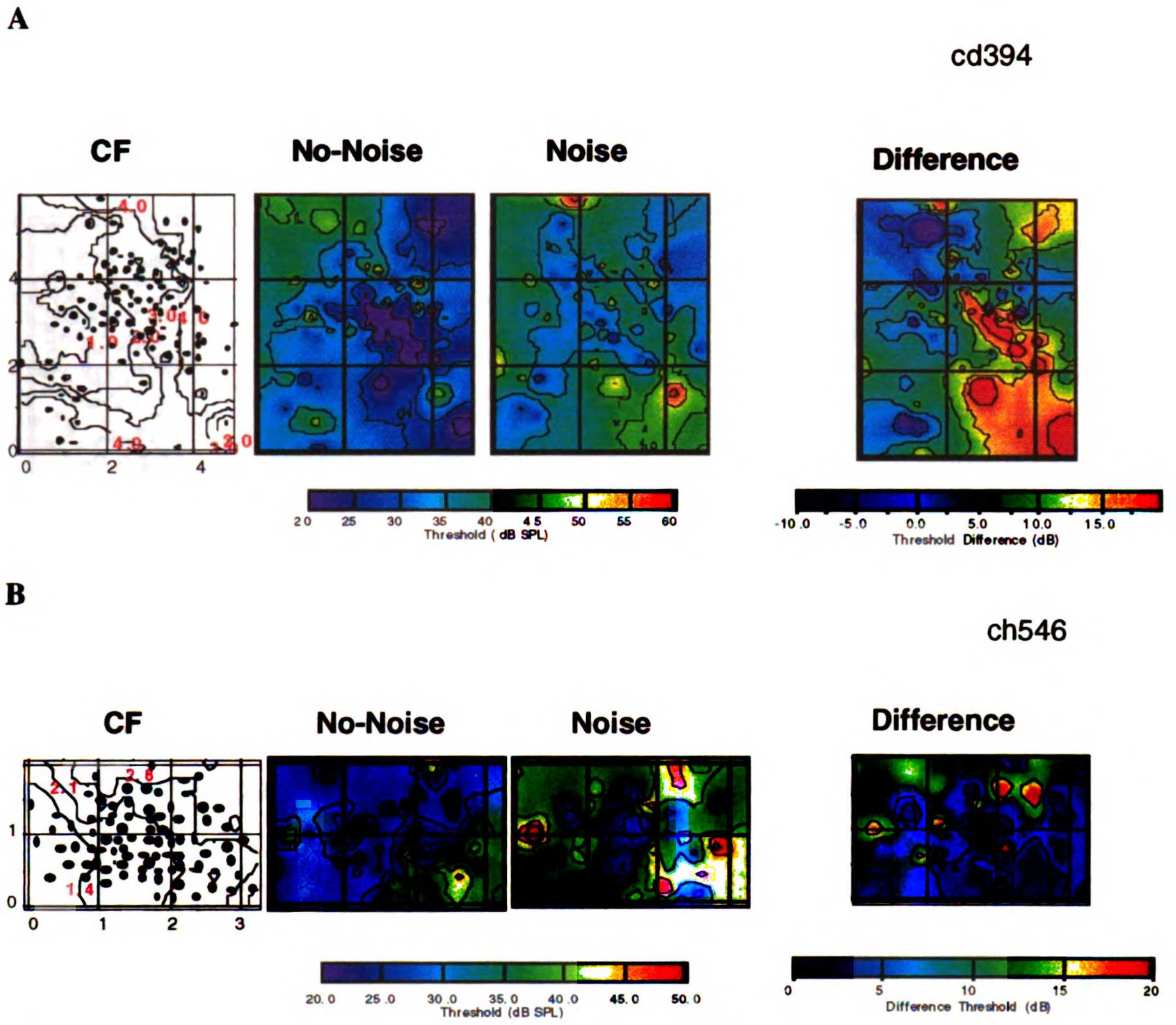


Figure 2.4 Distribution of threshold difference

The distribution of threshold difference, which measures the total threshold shift when background noise is present, is plotted as a function of CF. The threshold difference values were split into two subgroups on the basis of the unit's sharpness of tuning in the absence of noise. The distribution of threshold difference showed a positive shift in threshold that was similar for both broadly- and sharply-tuned neurons. There tended to be more broadly-tuned neurons sampled in CF regions below 1 kHz.



Figures 2.5**A-G** Effects of background noise on the spatial topography of FRA threshold

The effects of background noise on the spatial topography of FRA threshold were examined in seven cats. For each animal, the frequency representation and locations of the electrode penetrations were plotted in terms of cortical distance (in mm) as a reference for the threshold maps reconstructed from the two stimulus conditions (no-noise and noise). The difference map, which depicts the spatial distribution of the amount of threshold shift, provides a quantitative measure of the background noise effect. In **A**, a 6x5 mm region of auditory cortex whose frequency representation extended from 1 to 4 kHz is represented in the CF map. In the absence of background noise, the spatial distribution of threshold revealed a cortical regions of mid- to high-thresholds with a localized region of the lowest thresholds (in purple) in the central region crossing several CF contours. In the presence of background noise, the spatial distribution of threshold revealed a global elevation of thresholds across the sampled cortical regions (as evidenced by more regions of green-to-red). Although the spatial distribution of the difference in threshold revealed the greatest amount of threshold shift the central region, the topography of high sensitivity in the center of the isofrequency contours flanked by regions of lower sensitivy was not changed by the presence of background noise. In **B**, a smaller region of auditory cortex (2.1 x 3.3 mm) with 84 penetrations and a frequency representation extending from 1.4 to 2.8 kHz was

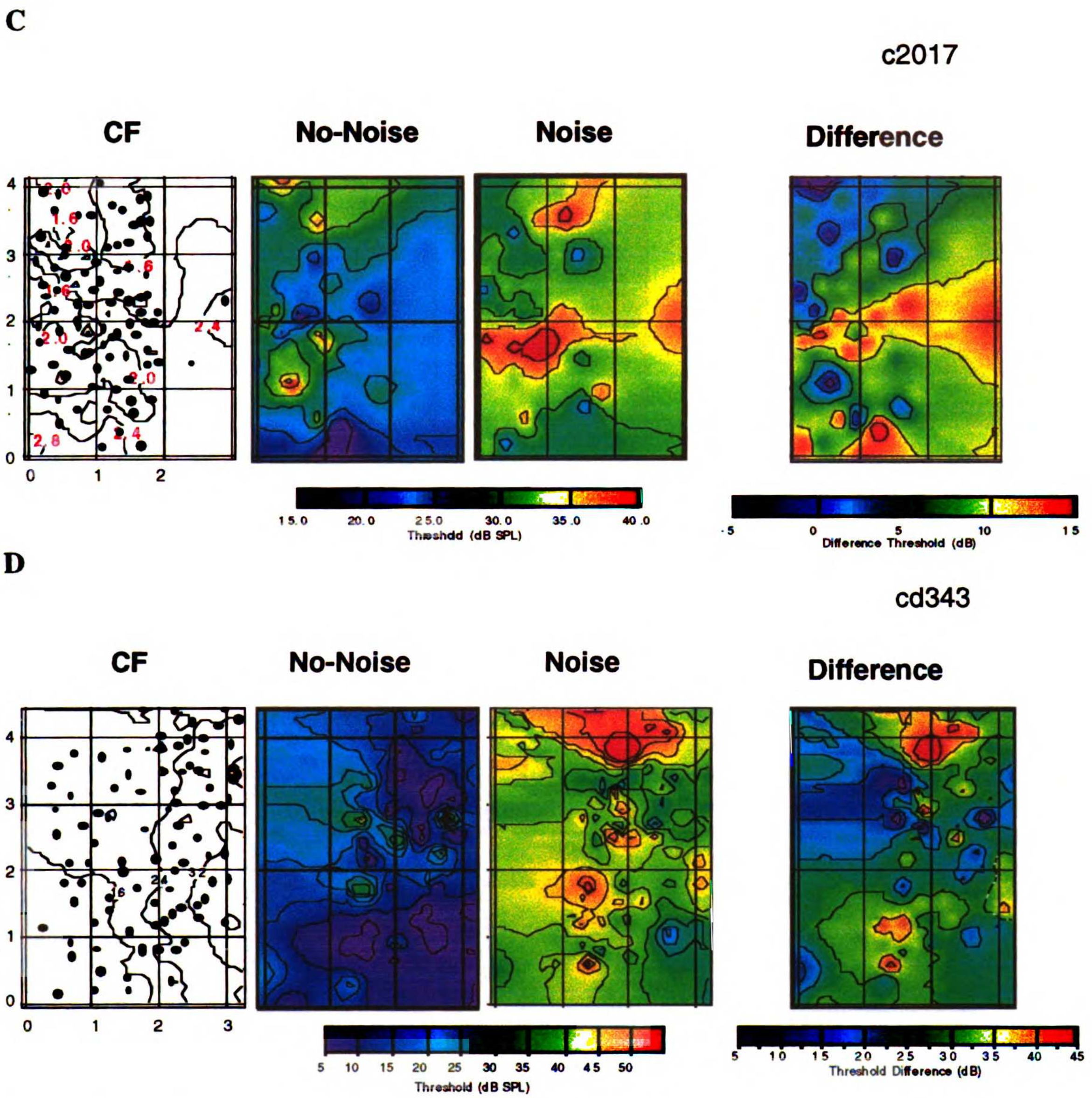


Figure 2.5 **A-G** (cont.)

characterized by mid- to high-threshold units in the absence of noise. In the presence of background noise, thresholds were globally elevated. The spatial distribution of the threshold difference revealed a localized region which exhibited minimal threshold shift flanked dorsally by regions of elevated threshold. Again though, the spatial topography of FRA threshold in the presence of background noise was not altered. In **C**, 85 penetrations within 4.2 x 3.1 mm region of cortex revealed a frequency gradient extending from 1.6 to 2.8 kHz. In the absence of background noise, the spatial distribution of threshold revealed localized areas of high-threshold units separated by a patch of lower threshold units. The threshold difference again revealed globally elevated thresholds, but essentially an unaltered spatial topography. In **D**, 96 penetrations within a 4.5 x 3.3 mm region of cortex revealed a frequency gradient extending from 1.6 to 3.2 kHz. In the absence of background noise, the spatial distribution of threshold revealed two localized regions of lower thresholds cutting

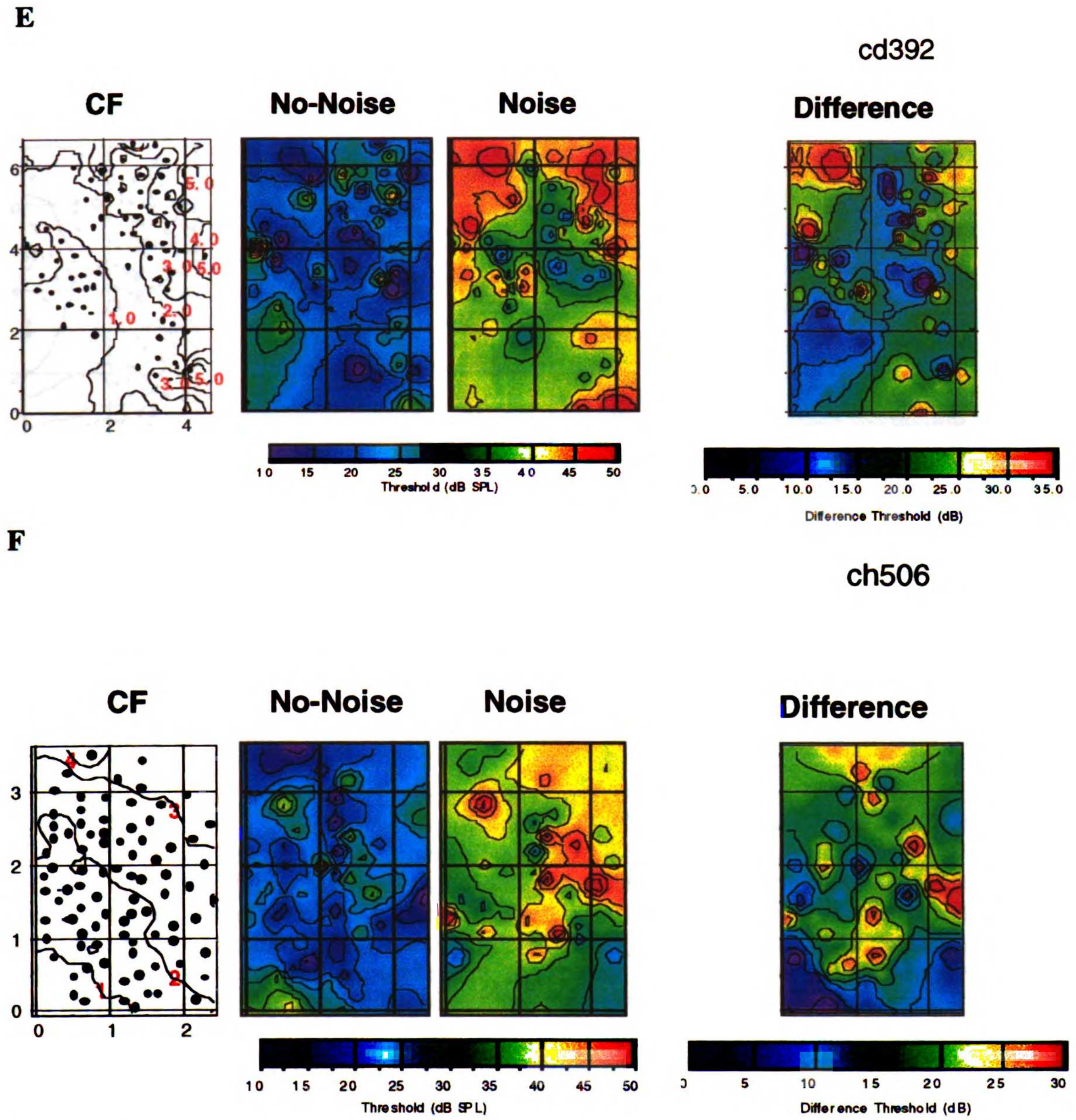


Figure 2.5 **A-G** (cont.)

across the isofrequency contours. The spatial distribution of the threshold differences again revealed globally elevated thresholds, but an unaltered spatial topography. In **E**, 76 penetrations within a 6.7 x 4.6 mm of auditory cortex revealed a frequency gradient extending from 1 to 4 kHz. In the absence of background noise, the spatial distribution revealed low- to high-threshold units. The spatial distribution of the threshold differences again revealed a global elevation of thresholds, but little change in the spatial topography of low- and high-threshold units. The next two examples (**F** and **G**), on the other hand, did show changes in the spatial topography of threshold in the presence of background noise. In **F**, 86 penetrations within a 3.7 x 2.5 mm region of auditory cortex revealed a frequency gradient extending from 1 to 3 kHz. In the absence of background noise, two thin strips of low-threshold units were found crossing several isofrequency contours. These two strips intersected near the 1.5 kHz

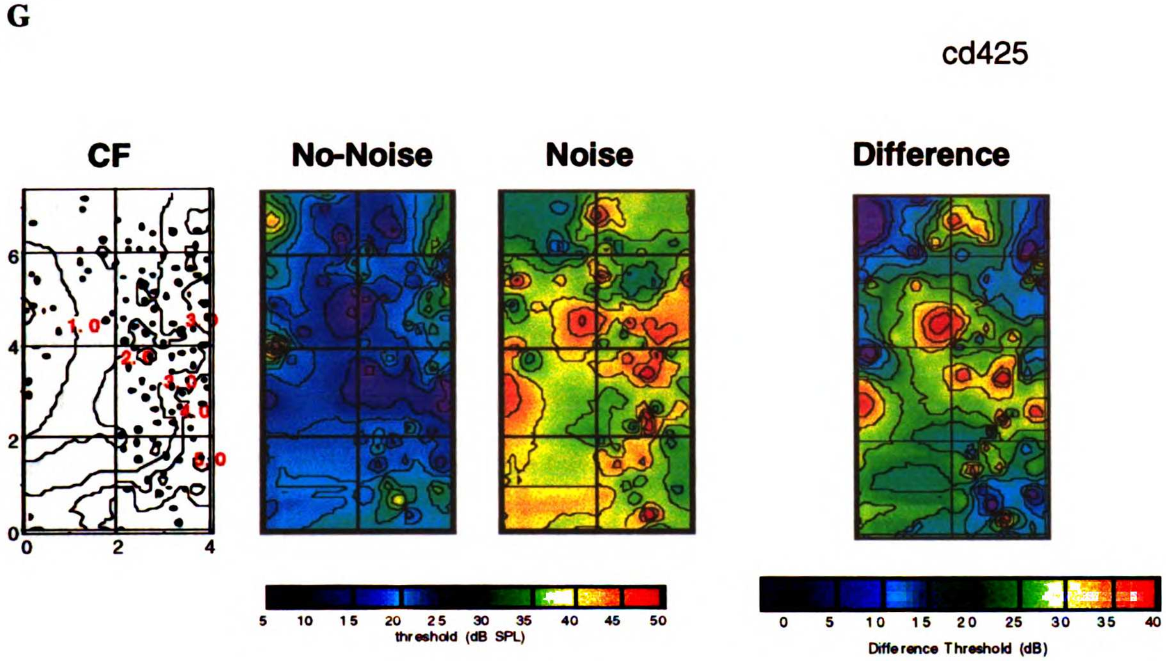


Figure 2.5 **A-G** (cont.)

isofrequency contour. In the presence of background noise, the spatial distribution of thresholds revealed a global elevation of thresholds; furthermore, the locations of the lowest thresholds relative to the other noise-elevated thresholds had changed. The difference threshold map revealed that the region with the greatest amount of threshold shift coincided with the initially low thresholds. In **G**, 159 penetrations within a 6.4 x 4.1 mm region of auditory cortex revealed a frequency gradient extending from 1 to 5 kHz. In the absence of background noise, the spatial distribution of threshold revealed two localized regions of low-threshold units, one located centrally across several isofrequency contours and another one running vertically down the center of the no-noise grid. Higher threshold units flanked these regions on either side. In the presence of background noise, the spatial distribution of threshold revealed a global elevation of threshold and a dramatic change in threshold topography. Regions which were initially lower threshold relative to other areas became higher threshold relative to neighboring locations with the addition of background noise. The difference threshold map illustrates the greatest amount of threshold shift occurring in the initially lower threshold regions.

majority of the penetration locations. In the difference map, a yellow-red region represents the greatest amount of threshold shift. The largest differences cut across CF contours and matched the very location of the lowest thresholds found in this example. Despite the great amount of threshold shift found in this central region, the topography of high sensitivity in the center of the iso-frequency contour flanked by regions of lower sensitivity was still evident in the noise condition.

The next four experimental cases, figures 2.5B-E, showed very similar results to the previous example in figure 2.5A. The CF grid in figure 2.5B (ch546) depicts 84 penetrations within a 2.1 x 3.3 mm region of primary auditory cortex with CFs roughly between 1.4 to 2.8 kHz. These units were characterized by mid- to high- thresholds (thresholds > 15 dB SPL) in the no-noise condition. The highest thresholds for the no-noise condition were concentrated in the rostral-ventral corner of the mapped area. Thresholds increased by about 10 dB with the addition of noise as was manifest in the threshold difference. The largest threshold shifts were again observed at the location with low thresholds in the no-noise condition. With noise, the lowest thresholds were seen in the center of the mapped area with a relatively large increase of threshold in the dorsal region.

Figure 2.5C (c2017) depicts 85 penetrations in a 4.2 x 3.1 mm region of cortical tissue with mid- to high-threshold units with CFs ranging between 1.6 to 2.8 kHz in the no-noise condition. The CF organization for this case did not follow the typical contour gradient of low-to-high frequencies which may be indicative of having sampled in the dorsal region AI where frequency organization is less strict. Despite the loose CF organization, the threshold map for the no-noise condition showed localized areas of high-threshold units separated by a patch of lower threshold units. The threshold difference map again manifested an increase in threshold of varying degrees, except for a few isolated regions (in purple) where there was no threshold difference or a slight decrease in threshold with the application of background noise.

Figure 2.5D (cd343) depicts a 4.5 x 3.3 mm cortical map with 96 electrode penetrations of low- to high-threshold units spanning a CF region of 1.6 to 3.2 kHz. Two regions of lower thresholds were seen along the iso-frequency axis in the no-noise panel. The threshold difference map showed an increase in threshold ranging from 5 to 45 dB with the smallest increments in high-threshold regions.

The CF grid in figure 2.5E (cd392) depicts 76 penetrations within a 6.7 x 4.6 mm region of auditory cortex with CF ranging from 1 to 4 kHz. The tone-evoked thresholds ranged from low to high values. The threshold difference map revealed mostly an increase in threshold in this cortical sample with a few isolated regions where no change in threshold was observed. As in the previous cases, the individual threshold values changed from very little to a large amount, however the overall spatial topography of threshold did not change substantially with the addition of background noise.

The next two examples, on the other hand, did reflect alterations in the spatial topography of threshold with addition of background noise. This suggests context dependency in frequency threshold maps. The CF grid in figure 2.5F (ch506) showed that there were 86 cortical penetrations made in a 3.7 x 2.5 mm region of auditory cortex. The topography in the no-noise condition of figure 2.5A showed low- to high-threshold units spread across iso-frequency strips ranging from approximately 1 to 3 kHz. A thin strip of low-threshold units (in purple) positioned orthogonal to the CF gradient was found running diagonally in the lower right section of the no-noise spatial topography of threshold. Another strip of low-threshold units cut across the center section of the no-noise spatial topography of threshold. In the background noise condition, the thresholds at the majority of the penetration sites shifted to higher values. Most significantly, the lowest threshold region in the noise map was no longer located at the same locations as without noise. The original low-threshold regions became areas of intermediate thresholds relative to the other neural thresholds in the cortical space.

The threshold difference in figure 2.5F depicted an increase in threshold of at least 15 dB particularly in the low-threshold units. Very little threshold shift was seen in the CF region below 1 kHz (bottom left corner) and in some mid-threshold units scattered across the 2 kHz iso-frequency strip. Therefore, this example illustrates that the pattern of threshold shift with background noise is not constant across CF or even within a CF contour and that regions of high sensitivity may shift relative to other groups of cells when tonal stimulation occurs in other listening contexts.

The next example also clearly showed the same effect of an altered spatial topography of threshold as tonal stimulation was embedded in background noise. In figure 2.5G (cd425), the CF grid shows 159 electrode penetrations in a 6.4 x 4.1 mm piece of auditory cortex whose units ranged in CF from 1 to 5 kHz. The no-noise condition showed a horizontal belt of low-threshold units (purple) positioned orthogonal to the CF gradient with higher threshold units flanking either side along each iso-frequency strip. With the addition of background noise, not only did the thresholds become elevated, but the relative position of the region of highest tonal sensitivity also changed as indicated by the regions of yellowish-orange in the central belt region. The threshold difference map showed patchy regions of threshold increases, ranging between 20 to 40 dB, with the greatest amount of threshold shift occurring in the regions that were originally highest in neural sensitivity. There were a few isolated regions where no increase in threshold was observed. Again, this example clearly shows the effect of background noise on the spatial topography of threshold as is revealed by the change in relative threshold among neighboring aggregates of cells.

2.2.3 Sharpness of Tuning (Q10)

The Q10 of each unit was reported to reflect the sharpness of tuning 10 dB above the minimum threshold. The quality factor, Q10, ranks the sharpness of tuning from being broadly-tuned with a value starting near zero to being sharply-tuned with a Q10 value greater than two. The distribution of Q10 values measured in both noise and no-noise conditions are displayed in figure 2.6. A comparison of the distribution of Q10 values showed no qualitative difference between the two stimulus conditions. Both distributions peaked at a Q10 of one which was really a range of values between one and two, but not inclusive of two. This reveals that the majority of units displayed broad tuning in both the no-noise and background noise condition. The scatterplot comparison in figure 2.7 showed no significant linear correlation in Q10 values measured in the no-noise and noise conditions.

The Q10 difference was calculated to determine whether or not the addition of background noise yielded a quantitative difference in Q10. The Q10 difference was calculated for both broadly- and sharply-tuned units (> 2) and plotted as a function of CF in figure 2.8A. A positive Q10 difference was defined as an increase in bandwidth in the presence of background noise while a negative Q10 difference was defined as a decrease. In this plot, the sharply-tuned units became 1) broadly-tuned (Q10 difference < 0), 2) more sharply-tuned (Q10 difference > 0), or 3) stayed roughly the same in bandwidth tuning (Q10 difference = 0). The broadly-tuned units, on the other hand, either became more sharply-tuned or maintained roughly the same bandwidth. As unit CF increased for sharply-tuned neurons, the distribution of Q10 not only broadened around zero, but also became more concentrated in difference values between 0 and -5. This concentration of negative values in the sharply-tuned units at higher CF contributed to the observed trend of bandwidth broadening (Q10 decreasing) with the addition of background noise. For broadly-tuned neurons, as CF increased, the distribution of Q10 also broadened, but with the concentration of Q10 difference values above zero contributing to the trend of bandwidth sharpening (Q10 increasing) in the presence of background noise.

A scatterplot comparison of the Q10 difference and Q10 measured in the no-noise condition, manifest in figure 2.8B, revealed a significant negative correlation (R= -0.7,

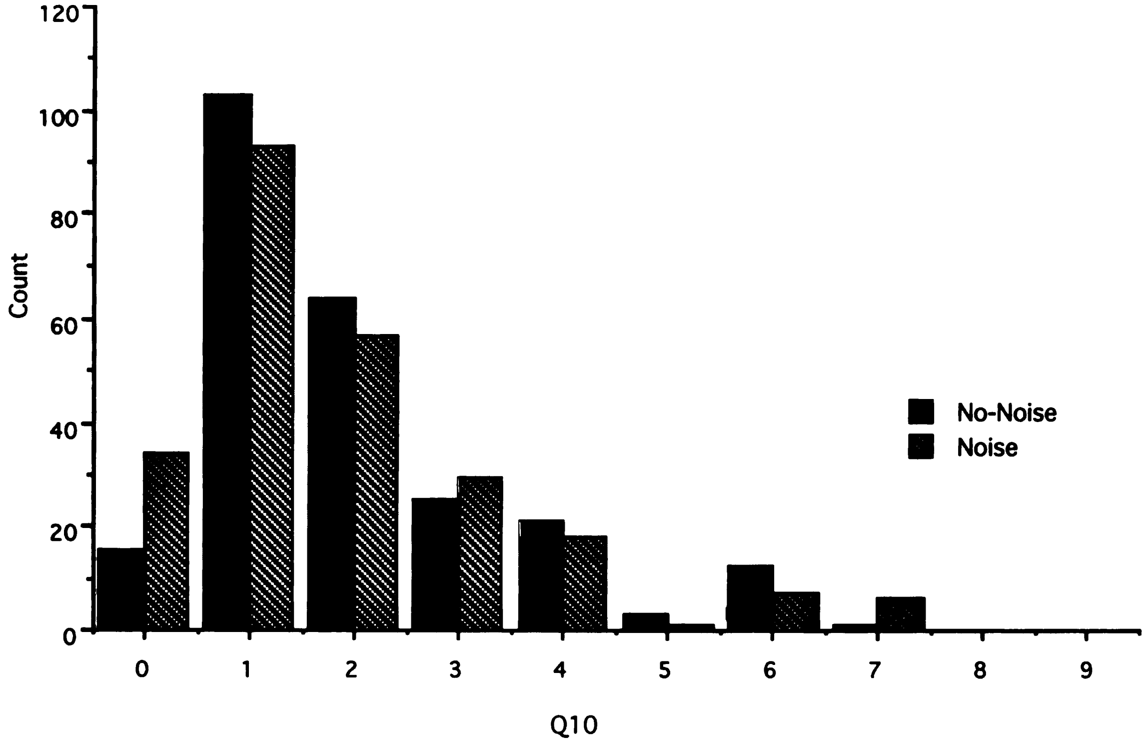


Figure 2.6 Frequency distribution of Q10

The frequency distribution of Q10 measured from FRAs recorded in quiet and noisy backgrounds revealed a peak around 1. The peak and width of the Q10 distributions for either stimulus condition were similar.

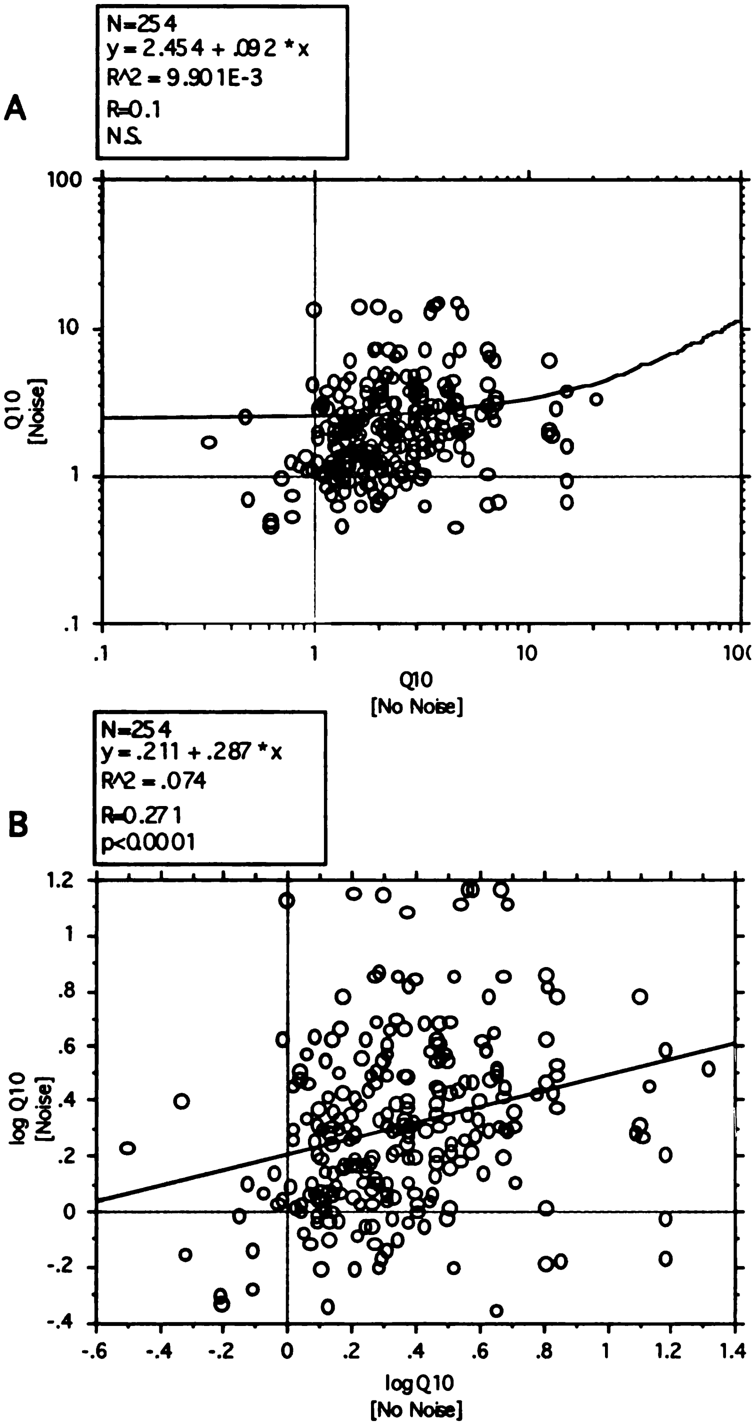


Figure 2.7 Scatterplot of Q10 measured in quiet and noisy backgrounds

In **A**, the distribution of Q10 measured in a noisy background is plotted as a function of Q10 from a quiet background. Regression analysis of Q10 revealed no significant correlation between the two conditions. When Q10 was converted to log units, as represented in **B**, regression analysis of Q10 revealed a weak positive correlation.

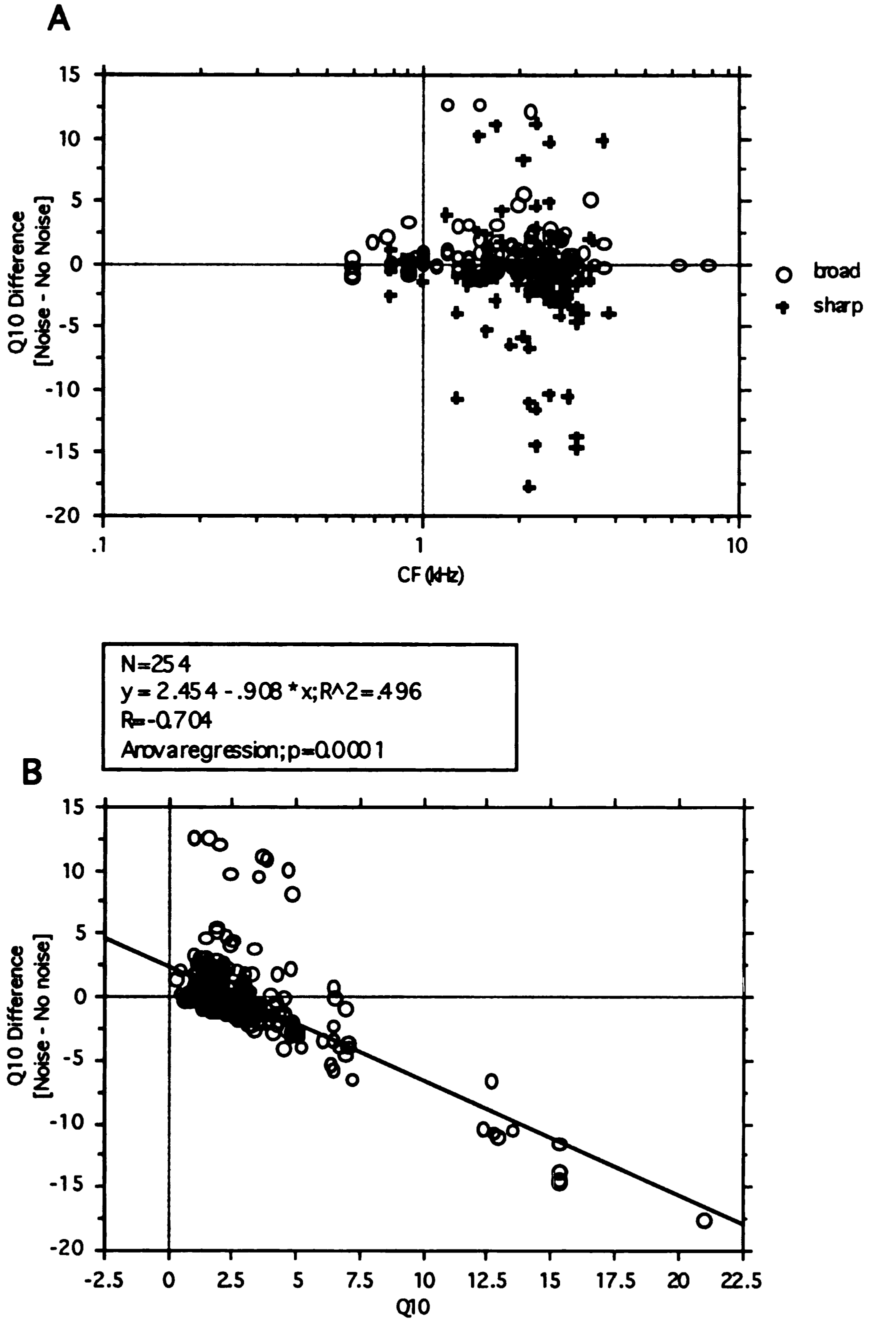


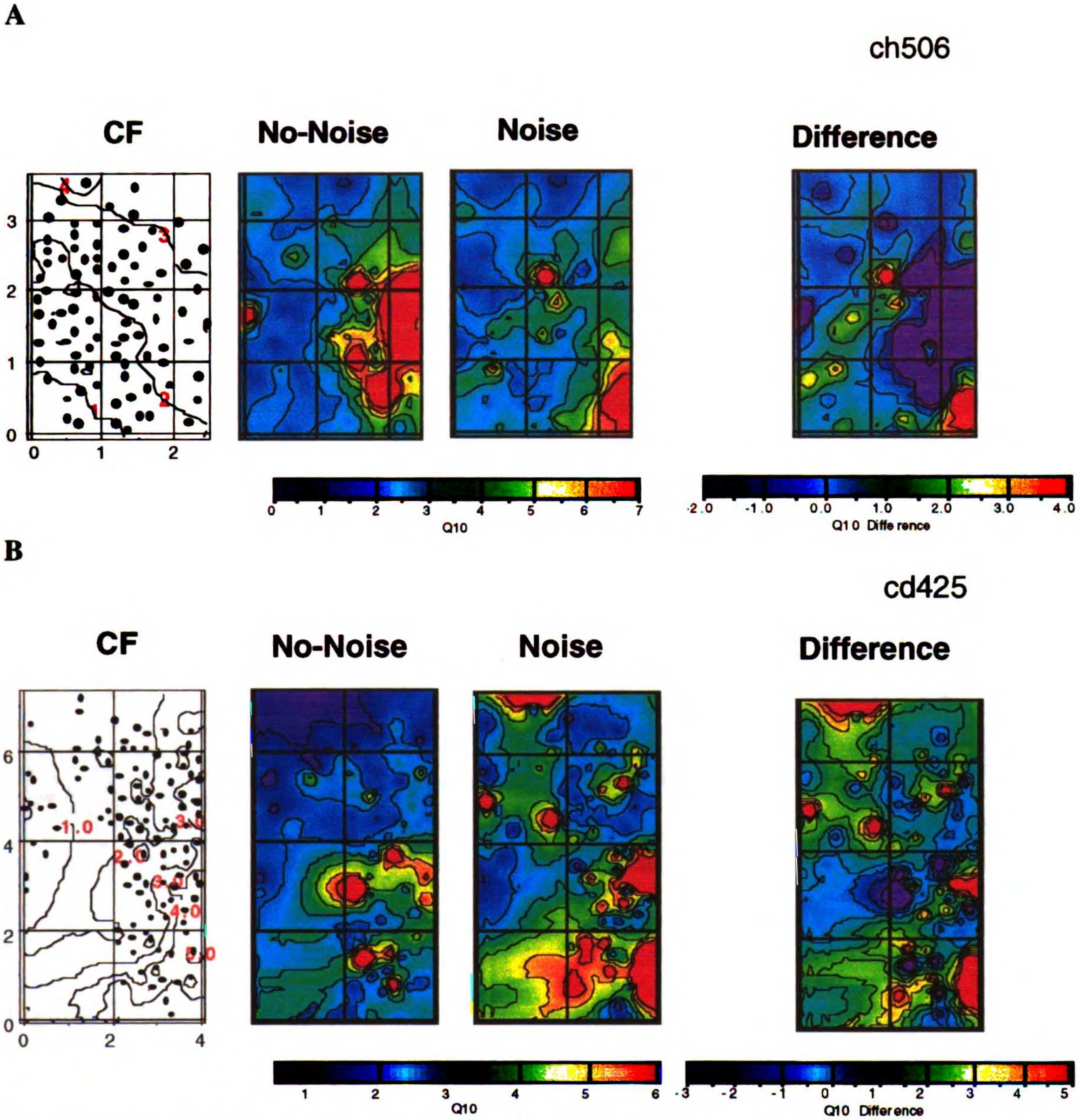
Figure 2.8 Distribution of Q10 difference and background noise effect on Q10

In **A**, the distribution of Q10 difference is plotted as a function of CF. Negative values indicate a lower Q10 or broader bandwidth in the presence of background noise, while positive values indicate a higher Q10 or sharper bandwidth. The distribution of Q10 difference for sharply-tuned units revealed changes in either direction or no change across the frequency axis. The greatest dispersion in the Q10 difference was found in units with CF greater than 1 kHz. The distribution of Q10 difference for broadly-tuned units revealed bandwidth change mostly in the positive direction or no change at all. In **B**, regression analysis of the Q10 difference, which is plotted as a function of initial Q10, revealed a significant negative correlation indicating an inverse change in bandwidth in units that experience change in Q10 in noisy backgrounds.

$p < 0.0001$). With a negative slope, the regression line depicted the average trend that as Q10 increased the Q10 difference decreased and became increasingly more negative. The high negative correlation further underscored the finding that sharply-tuned units became more broadly-tuned with the addition of background noise while broadly-tuned units became more sharply-tuned. There was, however, a small group of sharply-tuned units that became even more sharply-tuned; yet, this group also showed the general trend of a decreasing Q10 difference as the Q10 measured in the no-noise condition increased.

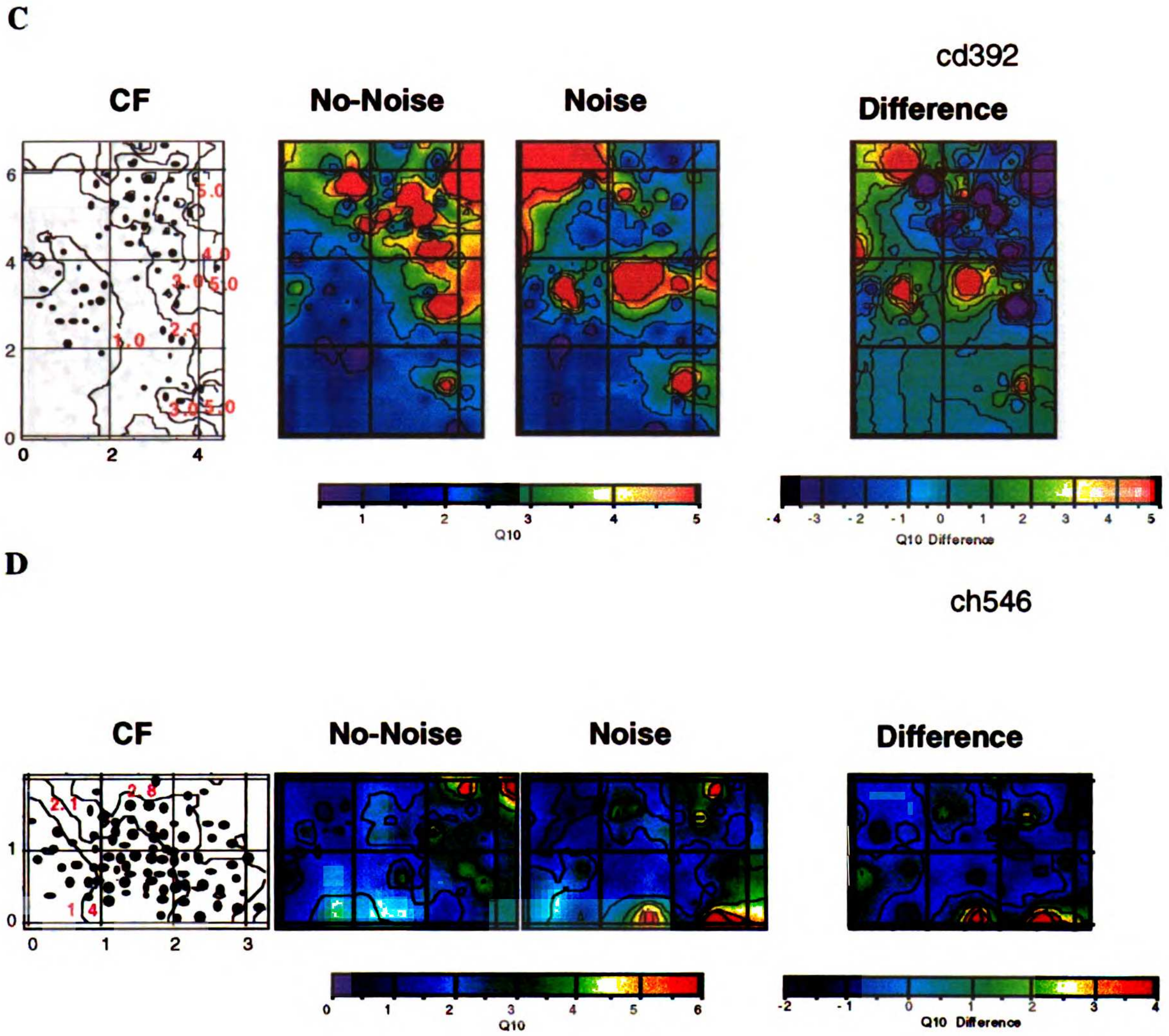
The spatial topography of Q10 from six out of the seven experimental cases further supported the observation of a differential effect of background noise in the listening environment. The bandwidth "inversion" occurred only in select subregions within an iso-frequency contour and spanned across the CF gradient of the cortical space. Figures 2.9A-G depict the CF map, Q10 maps for the no-noise and noise condition, and a map of the Q10 difference (noise - no-noise). The clearest example of a background noise-induced change in the spatial topography of Q10 was found in figure 2.9A (ch506). In this case, the spatial topography of Q10 in the no-noise condition showed a localized region of sharply-tuned neurons (yellow-red region) that coincided with the lower threshold units (see figure 2.5F). When continuous background noise was added to the tonal stimulus, the sharply-tuned region changed its Q10 characteristic and shifted to a broader bandwidth tuning (dark-blue/green region) while the broadly-tuned sides ventral to the sharply-tuned region became sharply tuned. The Q10 difference map underscored this change in bandwidth tuning while also portraying the lack of Q10 change in the dorso-caudal cortical area.

The next two examples also clearly demonstrated a bandwidth inversion with the addition of background noise. However, due to the doubling in size of the cortical space sampled, the affected subregions appeared smaller in these examples. The spatial topography of Q10 in the no-noise condition as seen in figure 2.9B (cd425) revealed a



Figures 2.9**A-G** Effects of background noise on the spatial topography of Q10

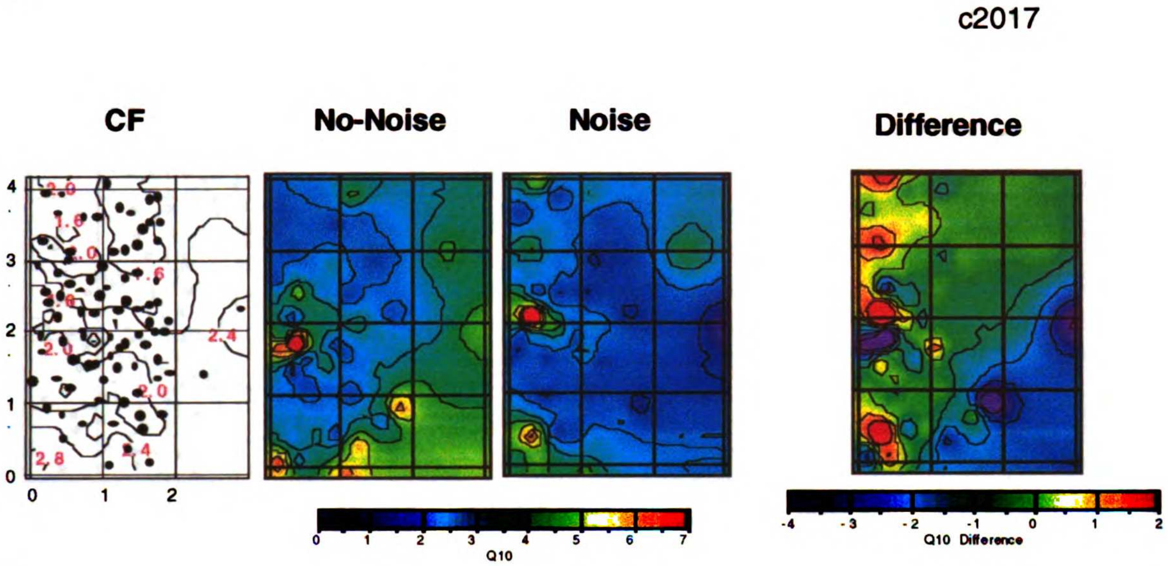
The effects of background noise on the spatial topography of Q10 were examined in seven cats. The general approach to studying the spatial topography of Q10 was the same as was described earlier in figure 2.5. Briefly, the frequency representation and electrode penetration locations were plotted in cortical distance as a reference for the spatial topographies of Q10 and the difference topography. The spatial topography of the Q10 difference, which measures the amount of change in Q10 with respect to addition of background noise, provides a quantitative measure of the background noise effect. In **A**, 86 cortical penetrations within a 3.7 x 2.5 mm region of auditory cortex revealed a frequency gradient extending from 1 to 3 kHz. In the absence of background noise, the spatial distribution of Q10 manifested a localized region of sharply-tuned neurons (indicated in yellowish-red) spanning across the isofrequency contours and flanked on either side by more broadly-tuned regions. In the presence of background noise, the spatial distribution of Q10 changed reflecting broad tuning in the region that was initially sharply- tuned. On either side of this new broadly-tuned area



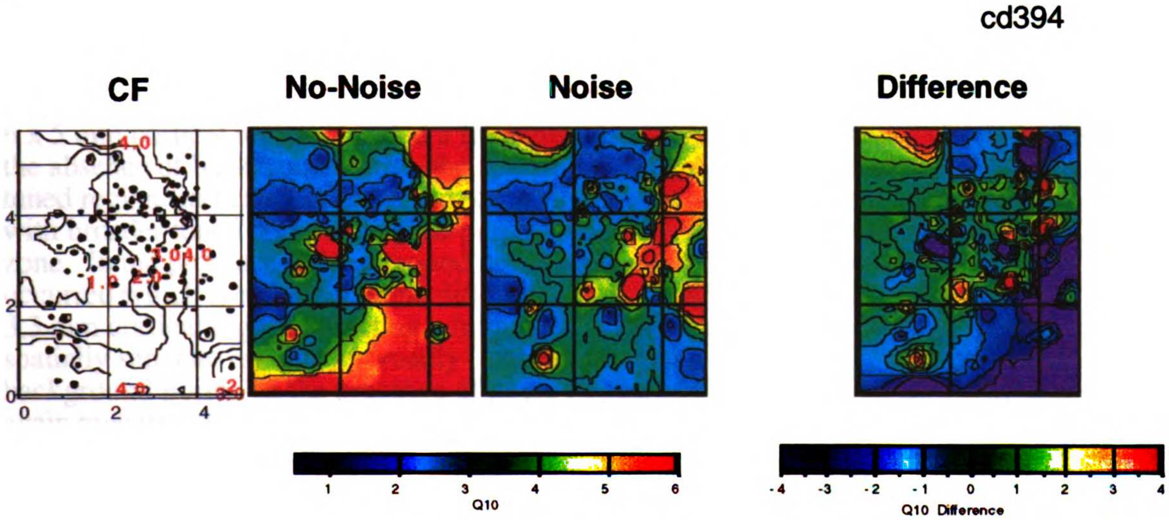
Figures 2.9**A-G** (cont.)

emerged two localized sharply-tuned regions. The spatial distribution of the Q10 difference manifested the inverse change in Q10 topography along the isofrequency dimension (purple indicates bandwidth broadening in the presence of background noise; green-red indicates bandwidth narrowing in the presence of background noise). In **B**, 159 penetrations within a 6.4 x 4.1 mm cortical region revealed a frequency gradient extending from 1 to 5 kHz. In quiet backgrounds, the spatial distribution of Q10 revealed a localized region of sharply-tuned neurons (indicated in green-red) across the center of the isofrequency contours flanked on either side by more broadly-tuned regions (indicated in blue-purple) and a second, smaller localized sharply-tuned area ventrally. In noisy backgrounds, the spatial distribution of Q10 again changed to reflect sharper tuning in the initially broadly-tuned areas which flanked the central region and broader tuning in the central and ventral regions which were initially more sharply-tuned. The spatial distribution of the Q10 difference illustrates this inverse change in Q10 topography along the isofrequency dimension. In **C**, 76 penetrations within a 6.7 x 4.6 mm cortical region revealed a frequency gradient spanning 1 to 4 kHz. In quiet backgrounds, the spatial distribution of Q10 revealed a large, but localized sharply-tuned region cutting across several isofrequency contours (indicated in green-red) with broader tuned areas on either side (indicated in blue-purple). Again

E



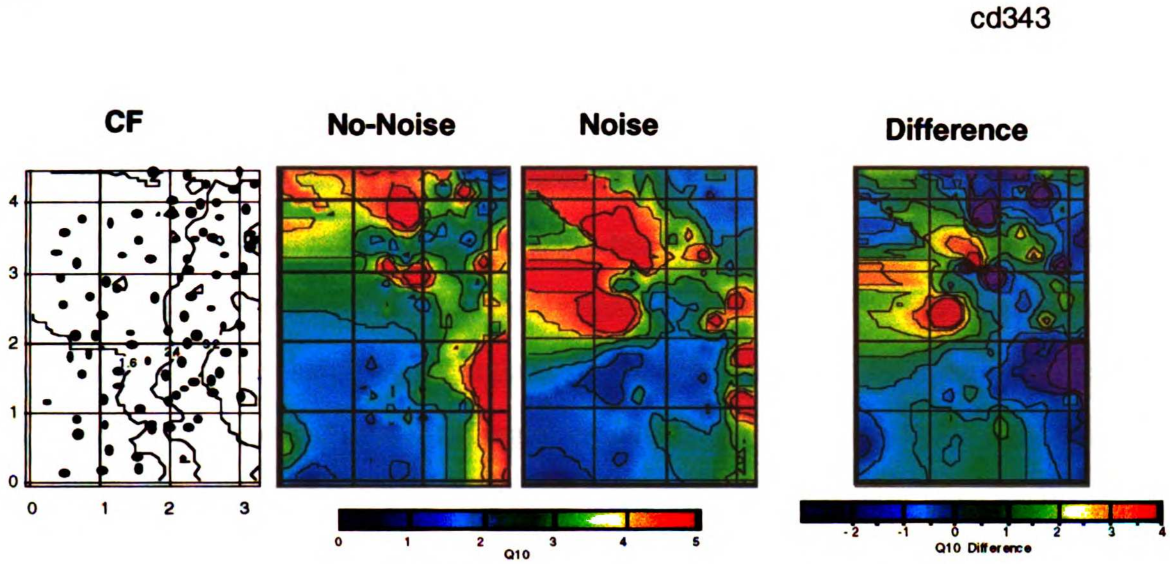
F



Figures 2.9**A-G** (cont.)

in noisy backgrounds, the spatial distribution of Q10 revealed an inverse change in Q10 in localized regions. That is, there were areas where no change in Q10 occurred as well. Again, the spatial difference of the Q10 difference manifested the inverse change in Q10 only in localized areas in the presence of background noise. In **D**, 84 penetrations within a 2.1 x 3.3 mm cortical region revealed a frequency gradient extending from 1.4 to 2.8 kHz. In quiet backgrounds, the spatial distribution revealed a sharply-tuned region (indicated in green-red) next to a broadly-tuned area (indicated in blue-purple) which spanned across the frequency gradient. The inverse change in the Q10 topography was reflected in the Q10 difference map. In **E**, 85 penetrations in a 4.2 x 3.1 cortical region revealed a frequency gradient extending from 1.6 to 2.8 kHz. In quiet backgrounds, the spatial distribution revealed a localized sharply-tuned area in the central portion of the lower frequency contours (indicated in green-red) with broader regions surrounding it on both sides along the isofrequency dimension. In noisy backgrounds, the spatial distribution of Q10 changed as a result of inverse changes in

G



Figures 2.9**A-G** (cont.)

bandwidth relative to spatial location only in localized areas. The difference map clearly manifested cortical areas where no change in Q10 occurred. In **F**, 138 penetrations in a 6 x 5 mm cortical region revealed a frequency gradient extending from 1 to 4 kHz. In the absence of background noise, the spatial distribution of of Q10 revealed a sharply-tuned region (indicated in green-red) centrally located across the isofrequency contours with broader tuned regions (in blue-purple) on either side of the central, sharply-tuned zone. In the presence of background noise, an localized region of inverse Q10 change occurred which is reflected in the difference map. In **G**, 96 penetrations within a 4.5 x 3.3 mm cortical region reflecting a frequency gradient spanning 1.6 to 3.2 kHz revealed spatially separate regions of broadly- and sharply-tuned neurons in the absence of background noise. In the presence of noisy backgrounds, inverse changes in Q10 were again revealed clearly in the difference map.

sharply-tuned strip in the central region of the iso-frequency line positioned orthogonal to the frequency gradient (yellow-red region) flanked on either side by a broader-tuned region (purple-blue region). A secondary sharply-tuned region restricted to the ventral section of the 3 kHz contour was also apparent in the no-noise condition. With the addition of background noise, the prior sharply-tuned central region dropped in Q10 to become broader in bandwidth tuning while neighboring regions which were relatively broader now became sharply-tuned regions. In the noise condition, three main patches of sharply-tuned neurons (yellow-red region) resulted from the change in the contextual environment for tonal stimulation. The banded or patch reversal pattern was clearly visible in the difference map for Q10. A large central patch in purple was surrounded on dorsally, ventrally, and rostrally by yellow-red patches indicating an increase in Q10 with the addition of background noise.

Another example illustrating an obvious change in spatial topography of Q10 with the addition of noise was found in figure 2.9C (cd392). In the no-noise panel, a large subregion of cortex in the upper right corner, which cuts across several CF lines, contained a difference pattern of sharply-tuned neurons. For the lower CF (< 1 kHz) region, a small patch of sharper-tuned units relative to its neighboring area was depicted in green-yellow. When noise was present in the background of tonal stimulation, two main sharply-tuned regions (in yellow-red) emerged sandwiching a more broadly-tuned region (in purple-blue) in the middle. As was seen in the noise condition panel, these regions of either sharply- or broadly-tuned neurons were approximately oriented orthogonally to the iso-frequency lines. The difference map showed more clearly the effect of noise on Q10 in this example at both high- and low-frequency cortical regions. Patches of increased sharpness of tuning with addition of noise (red) alternated with patches of decreased sharpness of tuning along the iso-frequency axis. It should be noted that not all regions showed an inversion of tuning sharpness.

The next three examples similarly depicted changes in spatial topography of Q10 when tonal stimulation was given in a noisy background. The no-noise panel in figure 2.9D (ch546), shows a sharply-tuned region (green-red region in the upper right corner) next to a broadly-tuned area (blue region below the green-red one). With the addition of background noise, the cortical region which was previously broadly-tuned became sharply-tuned while a small portion of the sharply-tuned region became broadly-tuned as illustrated in the Q10 map for the noise condition. However, overall the spatial pattern of sharpness of tuning in this case was not well expressed, perhaps because of the small size of the mapped region.

The complementary nature of the sharpness of tuning changes with noise was more clearly expressed in the next case. Figure 2.9E (c2017) shows a small sharply-tuned region in the left portion of the no-noise map under a broadly-tuned region whose Q10 characteristic inverted with the addition of background noise as can be seen explicitly in the difference map. The spatial topography of Q10 in figure 2.9F (cd394) again shows a large sharply-tuned region on the right portion of the no-noise condition map and in the left center block in the low-frequency region of cortex. As illustrated in the noise condition panel, only a portion (lower right section) of this sharply-tuned region in the no-noise map became broadly-tuned when background noise was introduced. In the noise condition panel, only a very restricted strip of cortex resulted as being sharply-tuned when background noise was present. The difference map clearly showed a decrease in Q10 at the lower section of the 4 kHz line and at a small section in the left central block in the low-frequency region of cortex. Figure 2.9G (cd343) reflected spatially segregated regions of broadly- and sharply-tuned neurons. In the presence of background noise, inverse changes in Q10 were again revealed in the difference map.

Therefore, in the six examples described above, the effect of background noise upon the spatial organization of Q10 has been seen to profoundly change the positions of local aggregates with similar bandwidth tuning. The overall differences in Q10 seen as a collective neural group may not have been tremendous. However, locally a clear anti-correlation of sharpness of tuning with and without noise was evident although this appears to be more strongly expressed in the regions near the spatial concentration of sharply-tuned neurons. The locations in which these changes in Q10 occurred may be reflective of some systematic organizational change essential to compensate for an altered contextual background.

2.2.4 Onset Latency

The minimum onset latency for tonal stimuli as measured in the primary auditory cortex occurs typically around 10 to 20 ms post-stimulus. Figure 2.10 shows two distributions of minimum onset latencies measured for both no-noise and noise stimulation conditions. The peak of the onset latency distribution for the no-noise condition was 11 ms while the peak of the distribution for the noise condition was 13 ms. Not only did the peak shift up by 2 ms, but the distribution of onset latencies became broader when background noise was present. A scatterplot comparison of minimum onset latency measured in both stimulus conditions seen in figure 2.11 showed a weak positive correlation ($R=0.3$, $p < 0.0001$). Although the majority of the units with onset latencies between 10 to 15 ms showed longer latencies in the presence of background noise, there were a few longer latency units which showed a shorter onset latency when background noise was present.

To see the effect of background noise on minimum onset latency in subpopulations of neurons, the minimum onset latency difference (noise - no-noise) was calculated for both broadly- and sharply-tuned units. The distributions of onset latency differences were fairly similar between the two groups across CF as seen in figure 2.12A. Although onset latency differences were widely distributed across CF and centered around zero, the scatterplot shows that the concentration of onset latency difference shifted above zero as unit CF increased; that is, the higher the CF, the larger was the

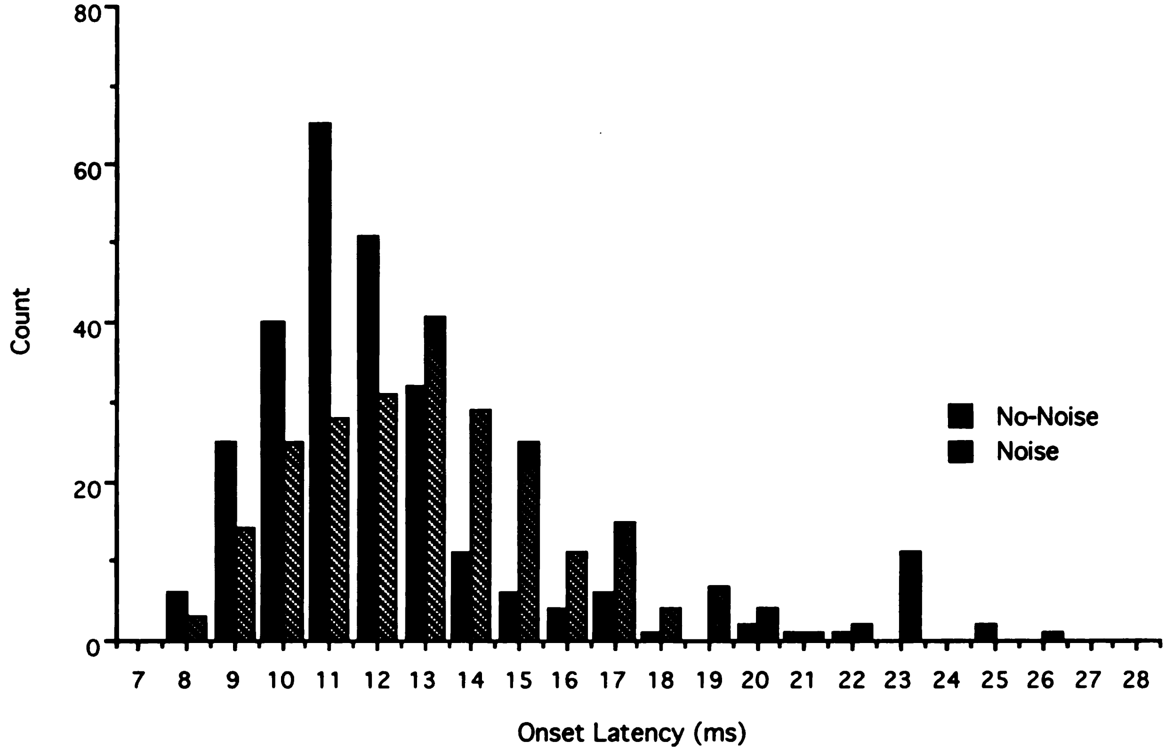


Figure 2.10 Frequency distribution of minimum onset latency

The distributions of minimum onset latency measured in quiet and noisy backgrounds both shared a similar range of latency values. The peak of the distribution shifted up by about 2 ms while the width of values immediately surrounding the peak broadened in the presence of background noise.

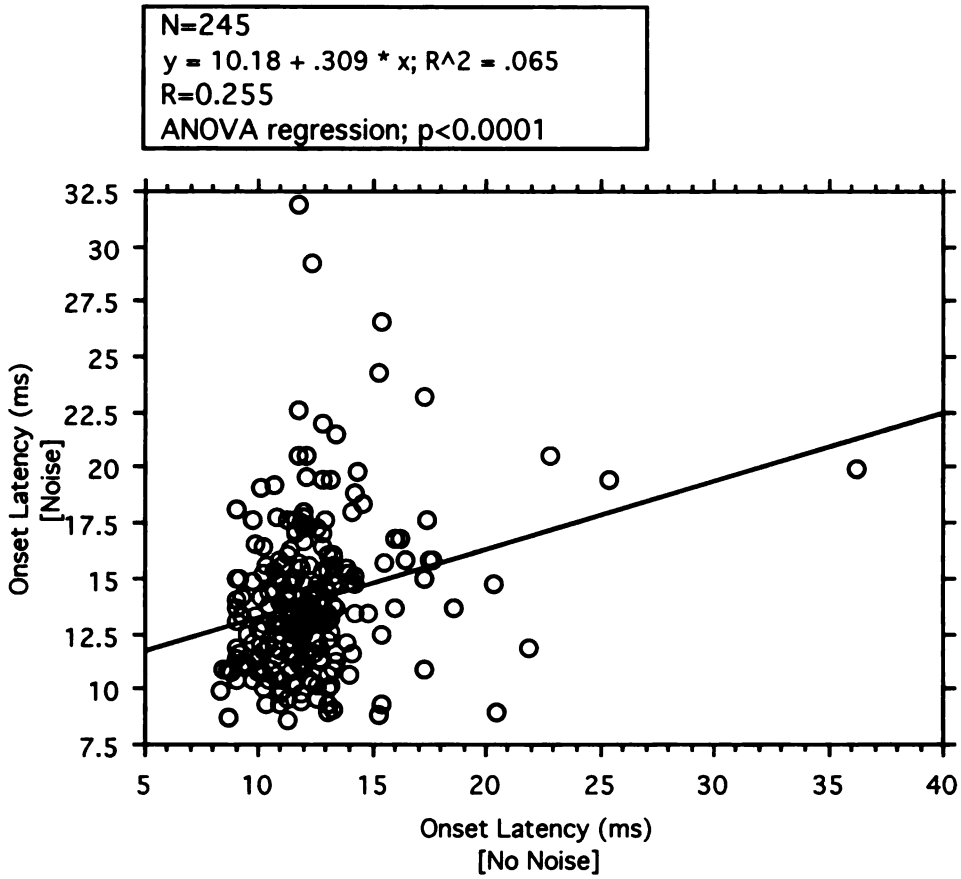
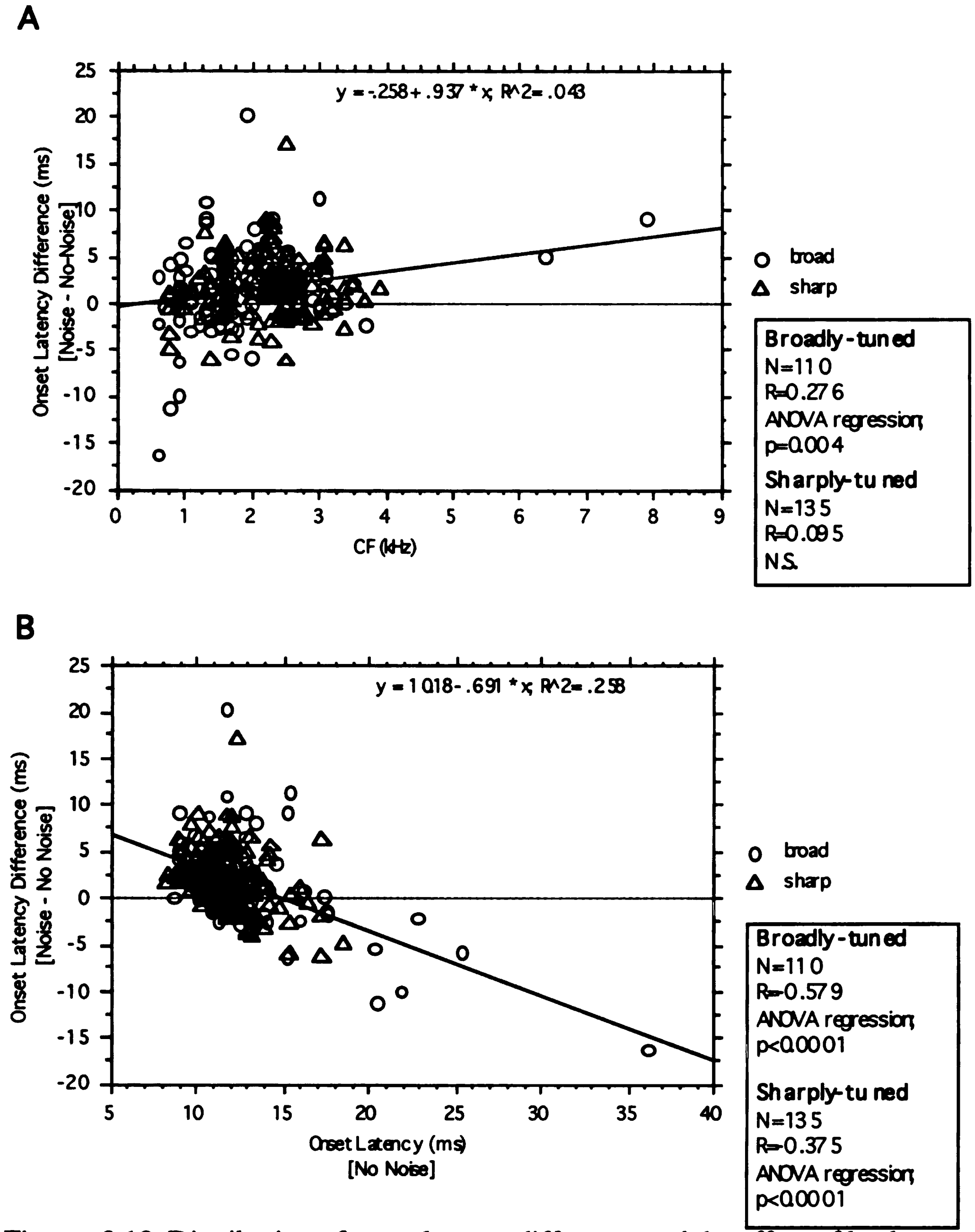


Figure 2.11 Scatterplot of onset latency measured in quiet and noisy backgrounds

Multiple-unit minimum onset latencies measured in noisy backgrounds are plotted as a function of onset latencies measured in quiet backgrounds. Regression analysis of the onset latencies revealed a weak positive correlation between the two stimulus conditions.



Figures 2.12 Distribution of onset latency difference and the effect of background noise on onset latency

In **A**, the distribution of onset latency difference is plotted as a function of CF. Negative difference values indicate lower onset latencies with background noise while positive differences indicate higher latencies. The distribution of onset latency difference from broadly-tuned units revealed a weak positive correlation with CF indicating a tendency towards longer onset latency delay with increasing CF. The distribution of onset latency difference from sharply-tuned units were evenly spread across CF. In **B**, regression analysis of the onset latency difference, which is plotted as a function of initial onset latency, revealed a significant negative correlation, which was stronger for broadly-tuned neurons, indicating an inverse change in onset latency in noisy backgrounds.

latency in the noise condition. A plot of the onset latency difference as a function of onset latency in the no-noise condition seen in figure 2.12B revealed a significant negative correlation (R=-0.6, p<0.0001). The negative correlation in onset latency suggests that background noise prolongs latencies of short latency units and shortens latencies of longer latency units. The distributions of broadly- and sharply-tuned units in this plot again did not appear different from each other.

The negative linear regression relationship found for onset latency difference and onset latency measured in no-noise was also evident in the spatial topography of onset latency. Figures 2.13A-G display a grid of cortical penetrations illustrating the CF gradient for each cat and the spatial topographies minimum onset latency in the absence and presence of background noise and the onset latency difference. In figure 2.13A (ch506), the cortical region below 2 kHz had longer onset latencies on average than the higher CF areas as seen in the no-noise panel. This latency gradient of shorter latencies at higher frequencies may partly be a reflection of travel time effects originating from the auditory periphery. Latency values across iso-frequency lines appear to have a nonhomogeneous distribution with adjacent aggregates of cells manifesting latency differences of at least 2 ms. When background noise was present, the global cortical activity appeared delayed. However, in a few areas the onset time either shortened or remained unchanged. This was most noticeable in the region around 1 kHz and below and the dorsal extent of the 3 kHz-4 kHz region. The difference map illustrates the effect of background noise upon the spatial distribution of latencies by depicting the reversal of global activation times in low-frequency auditory cortical neurons. In addition, background noise exerted a change in the organization of latency within an iso-frequency contour, not a maintenance of nonhomogeneous distribution at an elevated latency level.

In figure 2.13B (ch546), the cortical region below 2.1 kHz had longer onset latencies on average than the higher CF areas (except for in the lower left corner) as

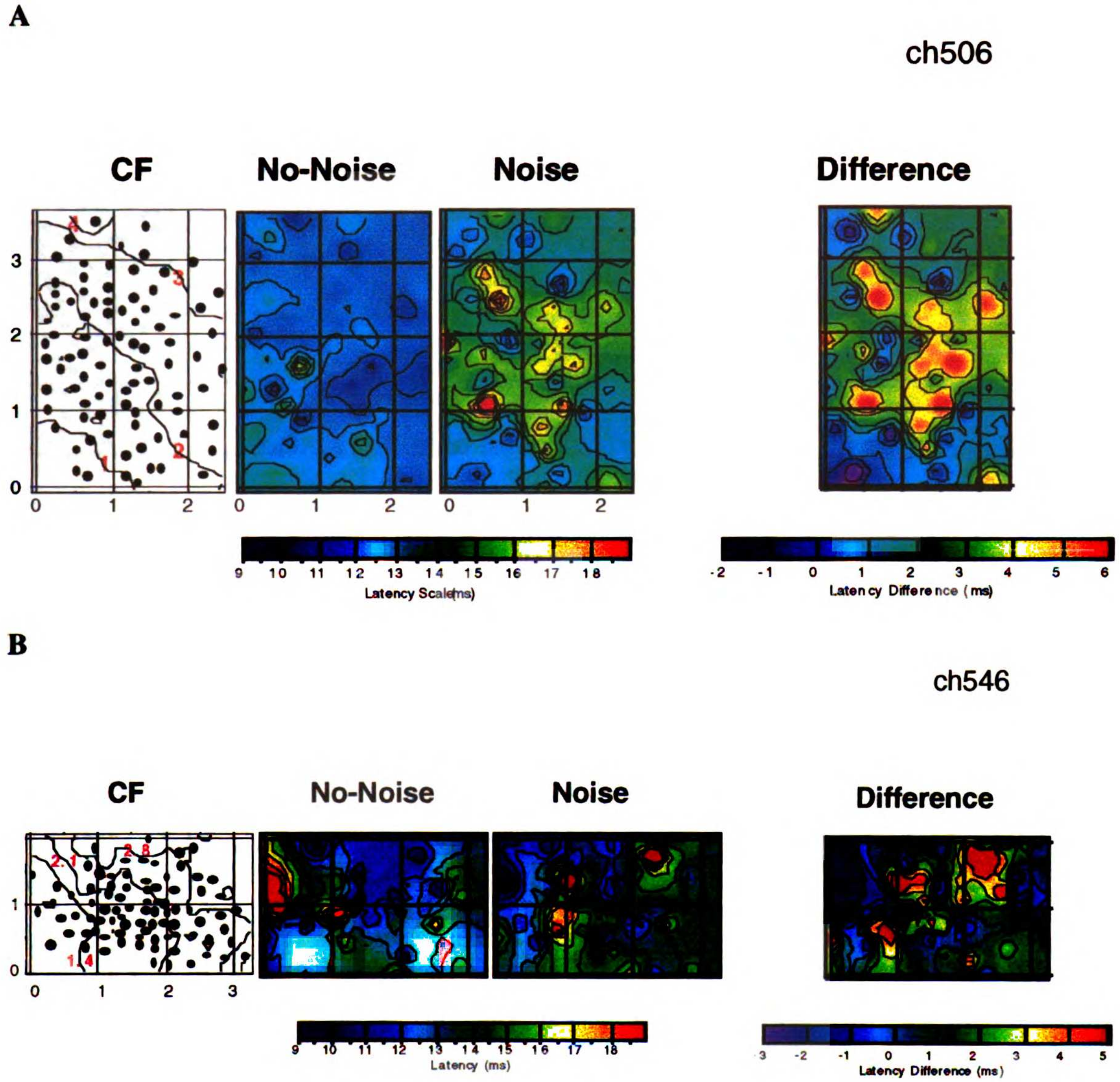


Figure 2.13**A-G** Effects of background noise on the spatial topography of onset latency

The effects of background noise on the spatial topography of onset latency were examined in seven cats. Again, the methods for studying changes in the topography have been described earlier in figure 2.5. Briefly, the spatial distribution of onset latencies were plotted as a function of cortical distance (in mm) and the amount of latency change resulting from the addition of background noise was depicted by the onset latency difference. In **A**, 86 penetrations within a 3.7 x 2.5 mm cortical region manifested a CF gradient from 1 to 3 kHz. In quiet backgrounds, a nonhomogeneous spatial distribution of onset latencies were found with longer latencies generally in the lower frequency regions. In noisy backgrounds, most regions showed an increase in minimum onset latency while localized regions, predominantly from initially longer latency areas, showed a decrease in minimum onset latency. The inverse change in onset latency was manifest in the difference map (green-red regions indicate increased onset latency with background noise while purple regions indicate decreased onset latency). In **B**, 84 penetrations within a 2.1 x 3.3 mm cortical region showed a CF

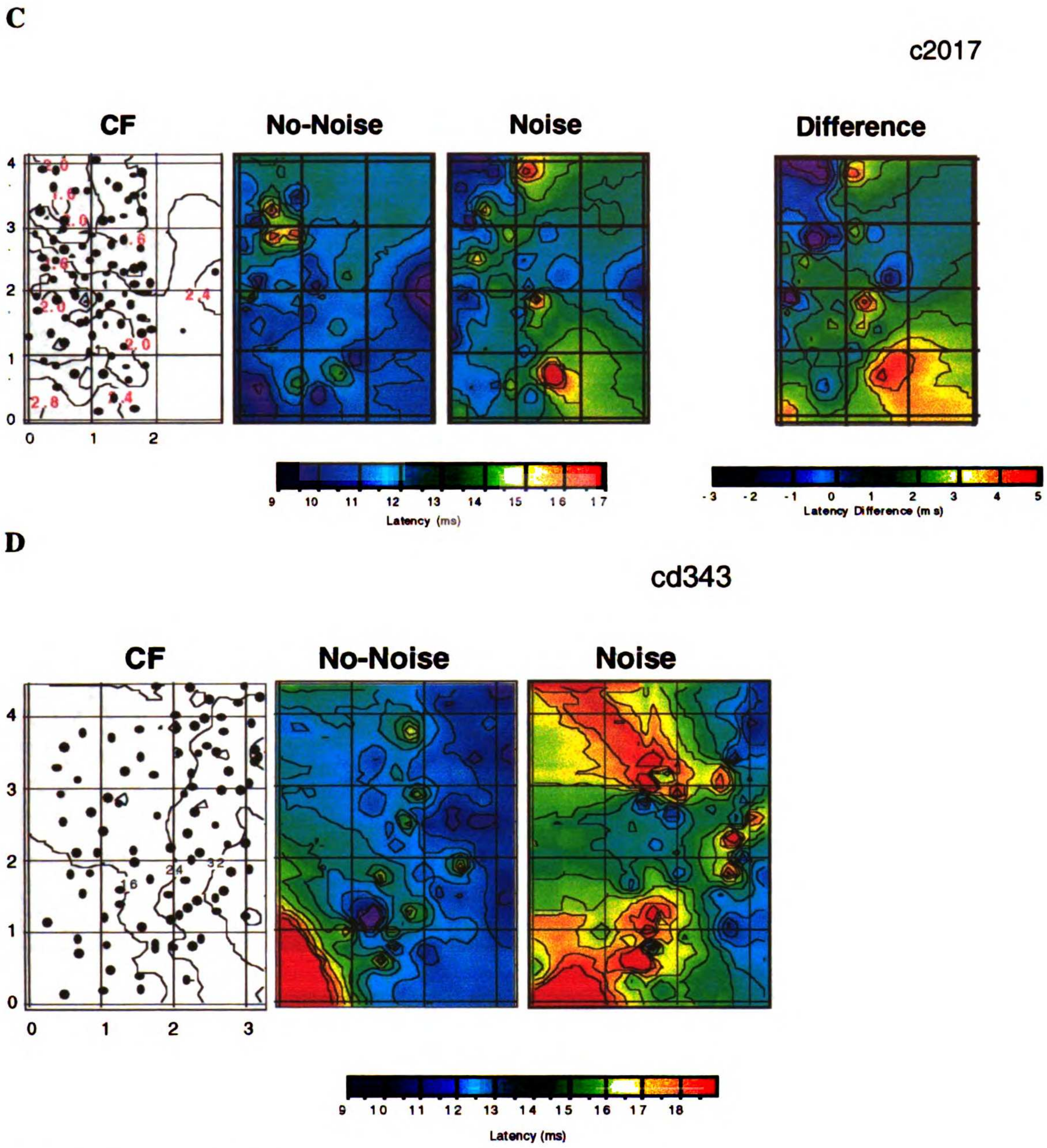


Figure 2.13**A-G** (cont.)

gradient from 1.4 to 2.8 kHz. In quiet backgrounds, a nonhomogeneous spatial distribution of onset latency was found across the cortical region. In noisy backgrounds, most of the areas increased in onset latency while some localized regions, dispersed unevenly across the isofrequency contours, decreased in onset latency. This inverse change in onset latency was depicted more clearly in the difference map. In **C**, 85 penetrations in a 4.2 x 3.1 mm cortical region revealed a CF gradient from 1.6 to 2.8 kHz. In quiet backgrounds, the latency gradient across the frequency gradient is similar to the previous examples. The spatial distribution along the isofrequency contours did not reveal any clear organization. In noisy backgrounds, most regions increased in latency, while some regions decreased in latency. In **D**, 96 penetrations in a 4.5 x 3.3 mm cortical region revealed a frequency gradient extending from 1.6 to 3.2

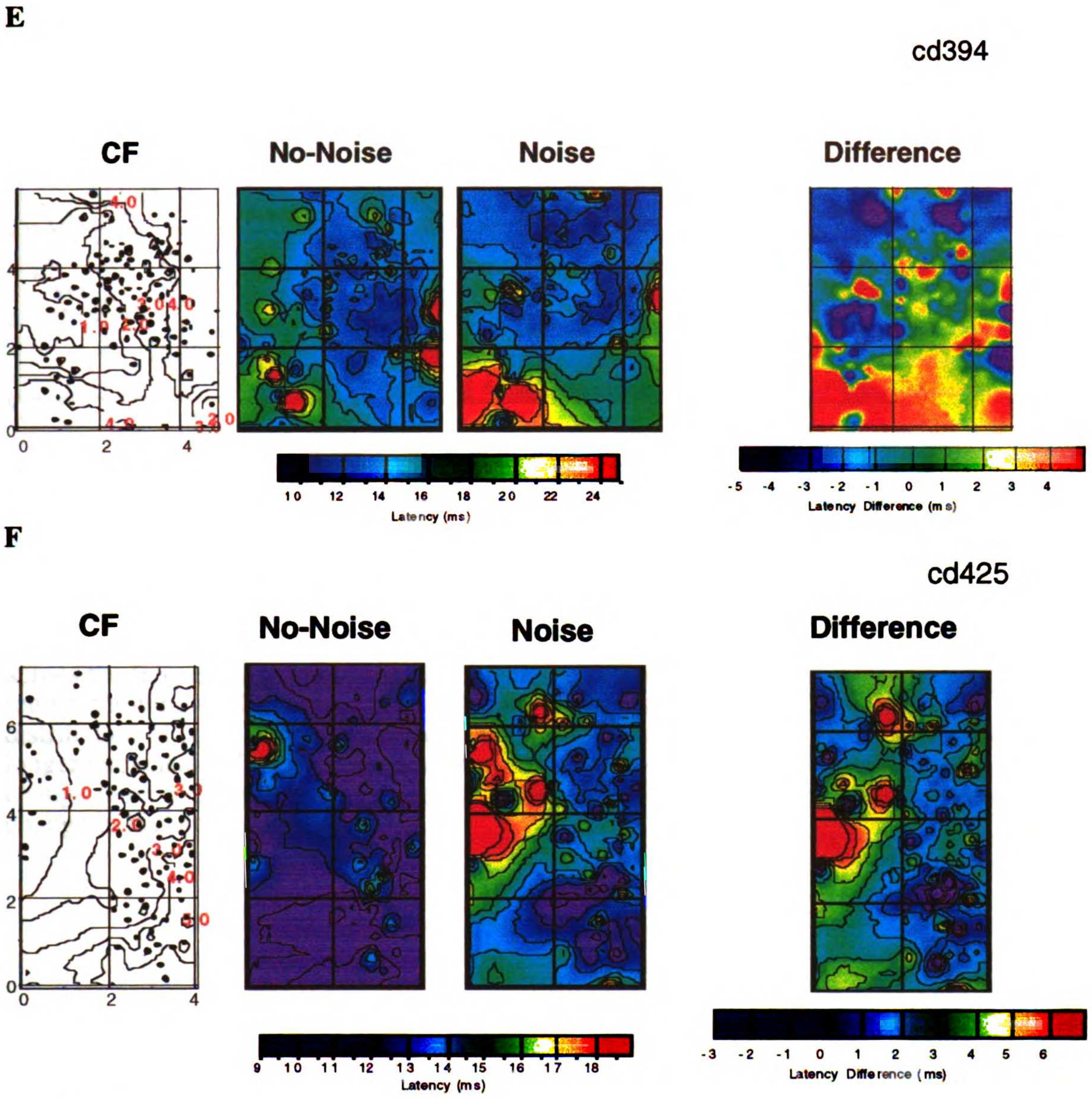


Figure 2.13**A-G** (cont.)

kHz. In quiet backgrounds, the spatial distribution of onset latency revealed a latency gradient (long to short) along the frequency axis with patchy, localized areas of longer latencies interspersed in shorter latency regions. In noisy backgrounds, the majority of the cells showed prolonged latencies, while others had shortened. In **E**, 138 penetrations in a 6 x 5 cortical region showed a CF gradient of 1 to 4 kHz. In quiet backgrounds, a localized region of shorter onset latency units were found to be centrally located across isofrequency contours between 1 kHz and 4 kHz. In noisy backgrounds, onset latencies were prolonged in the central shorter latency area while shortened onset latencies were found in neighboring areas. In **F**, 159 penetrations within a 6.4 x 4.1 mm cortical region revealed a CF gradient of 1 to 5 kHz. In quiet backgrounds, the majority of the units revealed onset latencies shorter than 12 ms, while localized patches of longer latency units were found cutting across isofrequency contours. In noisy backgrounds, spatially distinct onset latency regions emerged that showed either shortening or broadening of onset latencies which was dependent upon

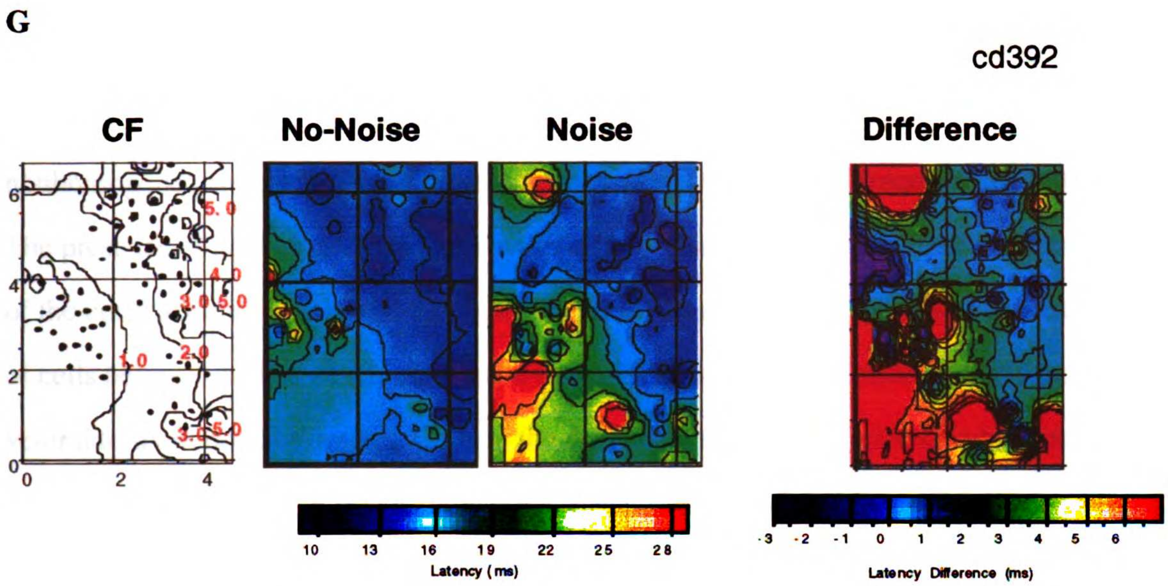


Figure 2.13**A-G** (cont.)

the initial onset latency. Although there was great variability in the response to background noise, patches of lengthened and shortened onset latency were found across the frequency gradient. In **G**, 76 penetrations within a 6.7 x 4.6 mm cortical region showing a CF gradient of 1 to 4 kHz. In quiet backgrounds, the spatial distribution of onset latency revealed a latency gradient across the frequency axis with longer latencies in the lower frequency region. In noisy backgrounds, a differential change in onset latency was observed in the low frequency region while there was less change in higher frequency regions.

seen in the no-noise panel. Again, adjacent aggregates of cells along the iso-frequency contours showed latency differences of up to 3 ms. There appeared to be a nonhomogeneous distribution of onset latency values along the iso-frequency extent. The presence of background noise again yielded differential changes in the onset times of the cortical neurons across the frequency and iso-frequency dimensions. Latencies of cells below 1.4 kHz, which were more dorsally located, became shorter while those ventral to them became longer. Cells in the frequency region above 1.4 kHz had longer latencies in general when background noise was present. The difference map for onset latency illustrates the differential change in latency across both the frequency and iso-frequency dimensions as was exemplified along the 1.4 kHz contour. Along the 2.1 kHz contour two groups of cells whose latency response in the presence of background noise differed were found to be spatially distinct and segregated. Again, this example demonstrates that background noise effectively alters the nonhomogeneous distribution of onset latency across frequency and within iso-frequency contours.

Although the spatial CF distribution was more noisy in the example shown in figure 2.13C (c2017), the latency gradient across frequency contours followed that of the previous examples. The iso-frequency contours were too disjointed in this example to reveal any clear organization of latency within a frequency line. The addition of background noise again revealed differential changes in onset latencies across AI. The regions of extended or shortened latencies were particularly clear in the difference map. Most of the lower half of the difference map showed a positive latency difference indicating prolonged latencies in the background noise condition, whereas the upper left region manifest shortened latencies with background noise.

In figure 2.13D (cd343), the latency gradient across frequency contours is consistent with the previous examples. The longest latency values were found in the frequency region below 1.6 kHz while the shortest values were above 2.4 kHz. The nonhomogeneous distribution of latency values across CF was seen along the isofrequency axis at 1.6 kHz and 2.4 kHz where there were alternating patches of short and long latency neurons. When noise was present, the majority of the cells showed prolonged onset latencies while only a minority of cells showed shortened onset latencies.

Response latencies to tones in a quiet background as seen in figure 2.13E (cd394) were nonhomogeneously organized across CF. The shortest latency values recorded were in the frequency regions above 2 kHz. Along the iso-frequency dimension, response latencies were also nonhomogeneously organized as seen along 2 kHz and 3 kHz where the more ventrally-located region had longer latencies than the rest of the cells along the contours. Along the 4 kHz iso-frequency strip, latency differences of up to 6-7 ms were seen between adjacent aggregates of cells. The spatial distribution of the difference in onset latencies again illustrated this differential change in values across CF and along the iso-frequency contours as a result of background noise.

The spatial representation of response latencies to tonal signals in figure 2.13F (cd425) again showed a nonhomogeneous distribution of latency values within the iso-frequency line. In the region below 1 kHz, there was a patch of cells with short latencies sandwiched between two patches of cells with longer latencies. In other frequency regions, there was a small patch of cells with longer latencies in between shorter latency cells. Again a differential change in onset latency was seen with the addition of background noise that altered the spatial representation of response latencies across frequency laminae and within an iso-frequency lamina. This differential change was better visualized in the spatial distribution of the onset latency difference.

Figure 2.13G (cd392) again showed the latency gradient across the frequency representation. The longer latencies were below 1 kHz while the shortest latencies resided in the higher CF region. The spatial distribution of onset latency within a CF contour was also nonhomogeneously organized. Within the iso-frequency contour at 1 kHz, the most dorsal region had the longest latencies. Along the 2 kHz contour there were alternating patches of cells with onset latency differences of 2 ms. The addition of noise brought about a differential change in onset latencies which was, again, most prominently seen within the iso-frequency contour. Along the 1 kHz contour, the most dorsal extent showed a reduction in onset latency while the adjoining region of cells showed a lengthening of greater than 7 ms in the onset latency. Along the other frequency contours, there were regions of no change in onset latency and regions of increased onset latency.

In summary, background noise acts to change the nonhomogeneous distribution of onset latency in a differential manner both across frequency laminae and within iso-frequency representations.

2.3.0 Discussion

The goal in this series of experiments was two-fold: 1) To study the effects of a fixed background noise level on receptive field properties of AI neurons and more importantly, 2) to compare the effect that a fixed continuous noise level has on the spatial distribution of functional properties in AI neurons with known functional topographies for quiet backgrounds. The methodology in this set of experiments differed from that in the previous chapter in which the continuous noise masker was set at a relative level for each single unit. By setting the continuous noise masker at a fixed stimulus level, which was done in these experiments, it is possible to compare functional receptive field properties relative to each cell sample in a collective, topographic manner.

The results are summarized as follows: 1) Neural sensitivity decreased in the presence of continuous noise. The relative spatial distribution of tone thresholds was preserved in most cases with the presence of continuous noise. 2) The change in sharpness of tuning with continuous noise was dependent on the tone-evoked excitatory bandwidth in quiet settings. As seen for single units (chapter 1), broadly-tuned multiple units tended to become more sharply-tuned when tested in background noise, and narrowly-tuned multiple units became more broadly-tuned. Since narrowly- and broadly-tuned neurons show spatial segregation, the differential change in bandwidth was reflected by locally restrictive, inverse changes in the spatial distribution of Q10. 3) The differential change in minimum onset latency with continuous noise was dependent upon the minimum onset latency in quiet settings. The spatial distribution of minimum onset latency in the presence of background noise was altered locally by the differential change in onset latency as a function of onset latency in a quiet setting.

The following can be concluded. 1) Continuous background noise affects the spectral sensitivity of AI cells not only through response threshold shifts, but also through differential bandwidth and onset latency changes. 2) The effect of continuous noise in AI neural activity is prevalent in the distinct topographic reorganization of specific spectral receptive field features such as bandwidth and onset latency. In other words, basic aspects of functional organization of AI, as reflected in its distribution of receptive field parameters, is not fixed but depends on contextual aspects such as the presence of background noise.

2.3.1 Technical Considerations

LIMITATIONS OF MULTI-UNIT RECORDING. The present study was conducted by recording multi- instead of single-unit activity since the primary goal was to look at relative local effects on the global spatial organization of receptive fields in AI when continuous noise is present. Although the data collected in these experiments were visually monitored and response filtered to ensure that spike waveforms were of neural origin, the absolute number of individual neurons sampled was unknown. Since the data collected reflects multi-unit activity, action potentials from cell types other than pyramidal cells (i.e., stellate cells) may have been incorporated in the data. Although comparison of the results from the previous chapter with those in this chapter indicate comparability of the multi-unit group data with those from the single-unit data, the goal was to study spatially-distributed, locally-averaged effects of tonal responses in continuous background noise. Thus, single-unit recording was not considered. In fact, in order to gain finer resolution of spectral parameter maps, single-unit recording would have been ineffective due to the small number of single units that can be isolated within a given animal. In addition, since the multi-unit activity recorded at each electrode penetration represents the sum of activity integrated at each site, the response recorded from a multi-unit cluster would be a more accurate representation of the local relative effects of background noise than would any single unit.

ANESTHETIC STATE. Unlike previous mapping studies which used sodium pentobarbital for anesthesia (Mendelson, Schreiner, and Sutter, 1997; Schreiner, Mendelson, and Sutter, 1992; Schreiner and Mendelson, 1990), all of these experiments were conducted with the use of ketamine in conjunction with diazepam. The effective difference between the use of ketamine and sodium pentobarbital on the mean spectro-temporal characteristics of AI neurons, as reported in other studies which have compared the use of both types of anesthesia (Brosch and Schreiner, 1997; Calford and Semple, 1995), is minimal if any. Therefore, the interpretation of the effect of a constant continuous noise source upon the functional topography in AI may be made without considering the effect of using a different anesthetic. However, whether the effect of a constant level, continuous noise on AI topography would be to the same degree had the animal been awake has yet to be determined by further studies.

CONSTRUCTION OF FUNCTIONAL MAPS. The representation of spectral parameters in a spatially-distributed manner is an attempt to display a topography of neural filter properties that would allow a parametric study of how perturbations to tonal stimulation backgrounds may affect not only feature parameters of individual neural filters, but also the systematic organization of those features along defined-CF boundaries. The spatial distribution of response parameters across the cortical surface was generated by using a gridding software (Surfer, Golden Software) that takes into account the actual spatial locations of the recording sites and makes a two-dimensional grid of the data. The third dimension represents the magnitude of the response feature measured. Since the gridding program uses a weighted, inverse-distance, squared algorithm that interpolates among the four closest neighboring points within a quadrant, extreme data values are reduced by the weighted, local average which results in a slight compression in the functional range. One of the disadvantages in using this technique is that local distortions in the map may arise with non-uniform sampling of recording sites within the cortical space. This problem is typically more pronounced at the corners and edges of the smoothed spatial pattern. However, since the objective in this study was to make intra-subject comparisons of functional maps derived from different background stimulation conditions, the effect of spatial distortions represented in the map is minimized in the interpretation of the data since the degree of compression in the functional range would be the same within each animal.

CONSTANT LEVEL OF CONTINUOUS NOISE. Whereas the results obtained in the previous chapter were of continuous noise effects where the noise level was fixed relative to the response threshold, the present results which in some instances reflect similar findings are in response to a constant level of continuous noise across all units of varying sensitivity. By using a constant level of continuous noise, one can compare relative frequency sensitivity across neurons in a topographic manner which was the primary goal of this study. However, by using this method, the level of effective masking will understandably be different for neurons of varying sensitivity. Therefore, the interpretation of the data will require looking at the combined effects of continuous noise on several receptive field parameters and not merely on any individual spectral feature.

2.3.2 Relation to previous studies

As was discussed in the first chapter, previous studies have shown that neurons respond to tone burst stimulation in the presence of continuous background noise through the displacement of their tone-intensity profile toward higher tone levels and through prolonged latent periods for tones (Phillips and Cynader, 1985; Phillips, 1985), and with sensitivity loss of greater than 15 dB, rate-intensity slope steepens thereby compressing the dynamic range (Phillips and Hall, 1986). This effect has also been reported to be comparable in either monotonic or non-monotonic cortical cells (Phillips, 1985). In the present study, continuous background noise was also found to increase response threshold, although multiple repetitions of CF-tones presented at several levels were not conducted in these experiments to see the shift in the rate-level function or compression of the dynamic range as previously reported. As a group, minimum onset latency was also found to be prolonged in the presence of continuous background noise. In addition, unit CF was found to influence the degree of prolonged latent period. High-CF neurons (> 2 kHz) are known to exhibit shorter onset latency than low-CF neurons to tone burst stimulation in quiet backgrounds due to a CF-dependent trend originating from the auditory periphery. Therefore, they show a greater shift in latency than low-CF neurons. Furthermore, the CF-dependent prolongation of onset latency with continuous background noise was found in broadly-tuned (Q10 < 2) neurons and not so for sharply-tuned neurons. Although categorization of neurons based upon their sharpness of tuning characteristic is not equivalent to a monotonicity categorization of non-monotonic and monotonic neurons whose response differences has been discussed in regards to intensity discrimination (Heil, Rajan, and Irvine, 1994; Phillips, 1990; Sutter and Schreiner, 1995), monotonicity topography is related to sharpness of frequency tuning topography as reported in previous studies (Schreiner and Mendelson, 1990; Schreiner and Sutter, 1992) and may be a subset of neurons classed by their sharpness of tuning. Based upon the superposition of sharpness of tuning and monotonicity maps, the outer regions of the dorso-ventral isofrequency contour is represented by broadly-tuned, monotonic neurons in the dorsal region and by broadly-tuned, non-monotonic neurons in the ventral region. Therefore, it is conceivable that both monotonic and non-monotonic neurons show the same trend in CF-dependent latency prolongation.

Previous continuous noise studies merely discuss the similarity of response between monotonic and non-monotonic units which was also the case in the present study; however, the present study suggests that a subclassification based upon sharpness of tuning of near-threshold responses may be a more informational marker in studying continuous background noise effects on receptive field parameters. The negative correlation of onset latency shift and tone-evoked onset latency in a quiet background is stronger in broadly-tuned neurons, whose regression line is weighted by the longer latencies of the low-frequency neurons collected in the sample, than in sharply-tuned neurons, whose latency values are more tightly clustered within a 10 ms range. Regardless of the difference in correlation strength, both classes of neurons display this relation between noise-induced latency shift and initial onset latency which suggests the global impact of a constant level of continuous background noise in this aspect. Although the latency data presented in this section was pooled across several animals with varying topographies sampled, it is worth noting that this correlation was found to be significant in the multi-unit data, but not in the single unit data presented in the previous chapter. A strong possibility exists that this effect could have been found in the single unit had their been a larger sampling of units, especially of low-frequency, longer latency neurons.

Although no other study has actually looked at the effects of continuous noise on the excitation bandwidth of the frequency response area of cortical cells, Phillips and Hall (1986) did present rate-level data that were constructed to form partial threshold tuning curves that suggests that continuous noise acts to broaden frequency selectivity near threshold. Partial threshold frequency tuning curves, composed of responses to 5-7 tone frequencies relatively close to each cell's CF and for which tone thresholds were within 10-20 dB of CF threshold, were constructed and the Q10 of these tuning curves were measured and compared in the absence and the presence of continuous background noise. Observations made from the current study extend those from the preliminary data collected by Phillips and Hall. The current findings reveal that broadened frequency selectivity, as measured by the Q10 differences between the absence and presence of background noise, occurs in only one class of cells. There are another subclass of cells which demonstrate a sharpening of frequency selectivity with the presence of background noise. Furthermore, these lines of cell classification need not be divided primarily by monotonicity, but by sharpness of tuning as was suggested in the prior chapter.

The idea that the mechanism controlling for frequency selectivity in the presence of continuous background noise and revealed through changes in spectral bandwidth is not dependent on the excitatory or inhibitory nature of the inputs activated by the mask is not novel (Phillips, 1985). Phillips and colleagues have shown that monotonic and non-monotonic cells, whose hypothesized differences lie in the different interplay of excitatory and inhibitory processes affecting spectral bandwidth at high- and low-SPL (Phillips, 1988), have displayed opposite responses to broadband noise (Phillips, et al., 1985). However, when continuous broadband noise is coupled with the presentation of a tone burst, both classes of cells show suppression of tone-evoked discharges (Phillips and Cynader, 1985). If both classes of neurons show similar effects to noise, then it would seem necessary to find a different neural subclassification that would manifest categorical responses to tones in the presence of continuous background noise. It has been suggested in the previous section that such a subdivision may exist in the classification based upon sharpness of tuning as represented by Q10.

As the results of this chapter revealed, the distributions of Q10 in either quiet or noisy background conditions are relatively the same, with respect to width and peak, when analyzing the neurons as a whole. However, when the distributions of Q10 difference, which measures the differential response between the two stimulation conditions, are distinguished by their tone-evoked spectral bandwidth categorization, the distributions become identifiably distinct from each other. That is, the broadly-tuned neurons experience an increase of Q10, or narrowing of bandwidth, when continuous background noise is present; whereas, the narrowly-tuned neurons show a bi-directional change in Q10, although the majority decrease in Q10, or broaden in bandwidth. Although bandwidth appears to be weakly dependent on CF, relative noise effects on Q10 should still be reflected in the topographic organization of bandwidth. The expansion and contraction of bandwidth for near-threshold activity in response to background noise is a reflection of the balance between excitatory and inhibitory processes that are hypothesized to be modulated by intrinsic connections of neighboring broadly-and sharply-tuned cells along the dorso-ventral extent of a frequency lamina (Read, per communication) which may be coupled by another parameter known to be organized in the primary auditory cortex such as binaural interaction bands (Imig and Adrian, 1977; Middlebrooks, Dykes, and Merzenich, 1980).

2.3.3 Spatial Distributions: Effect of continuous background noise

The systematic organization of receptive field properties in AI of cats have been reported for frequency (Merzenich, Knight, and Roth, 1975; Reale and Imig, 1980), excitatory bandwidth (Schreiner and Mendelson, 1990), and onset latency (Mendelson, Schreiner, and Sutter, 1997). In addition, the functional topographies of stimulus features such as tone intensity (Heil, Rajan, and Irvine, 1994; Phillips, et al., 1994; Schreiner, Mendelson, and Sutter, 1992) have been found along the isofrequency domain of the auditory cortex. The existence of multiple topographic representations of sensory input has been considered one of the foremost fundamental organizational principles in the mammalian forebrain (Schreiner, 1991) which may provide a framework for understanding cortical representations of complex auditory information such as speech. Since the eventual and more exciting goal is to understand how complex sounds are represented in the auditory cortex, the robustness of functional topographies need to be tested via the use of stimulation conditions that more closely resemble realistic listening environments such as continuous background noise.

In the present study, the receptive field parameter of response threshold was mapped in both quiet and noisy backgrounds. Previous studies have found minimum threshold, dynamic range, best SPL, and monotonicity to be correlated and spatially distributed along the isofrequency domain (Heil, Rajan, and Irvine, 1994). Heil and colleagues found that for low to moderate stimulus intensities, the spatial pattern of activity remains constant along the isofrequency strip; but, at higher intensities, which may activate SPL-dependent inhibition, the spatial frequency pattern changes markedly. In the present study, the effect of noise on the response threshold is dependent upon the changing signal-to-noise ratios within the response area of the cortical neuron. Although response thresholds shift in differing increments to the presence of continuous background noise, the resulting spatial distribution of response threshold along the dorso-ventral isofrequency contour remains basically the same in five of the seven animals. For the other two animals, the spatial topography for threshold not only changed with the addition of the continuous background noise but the topography appeared to invert in restrictive, local areas while remaining the same in other areas. Implications for this detected change in only two of the seven animals can only be speculative at this time. The conclusion gathered from data reported in this study is that the spatial pattern governing response threshold to tone bursts does not change with low levels of continuous background noise.

Previous studies analyzing the organization of excitatory bandwidth (Schreiner and Mendelson, 1990) found a segregation of broadly- and sharply-tuned neurons along the dorso-ventral isofrequency domain. The ventral and dorsal regions of the isofrequency axis were defined by broadly-tuned cells while the dorso-ventral center was defined by sharply-tuned cells. The data as reported by the results found similar spatially-segregated areas of sharply- and broadly-tuned neurons consistent with previous reports. When continuous background noise was applied along with the tone burst, the spatial segregation of sharply- and broadly-tuned neurons changed remarkably. The inversive bandwidth relationship that was reported first in the single unit statistics from the previous chapter, and next in the multi-unit statistics from this chapter is systematically grounded in the topographic inversion of bandwidth as reflected by the Q10 values. This pattern change was unmistakably identifiable in six of the seven cases. By using a continuous broadband stimulus in conjunction with a tone burst, inhibitory and excitatory processes are activated simultaneously. The finding that excitatory bandwidth organization may change with a variant background listening environment offers the interesting possibility that context-dependent spectral analysis is governed by complex interactions between a robust global frequency gradient and locally phase-shifted spatial frequencies of receptive field properties such as excitatory bandwidth.

The last receptive field property for which its spatial topography was challenged by continuous background noise is onset latency. Mendelson and colleagues (1997) reported a statistically significant correlation between minimum onset latency and Q10 and spatial locations of shorter onset latency coexistent with the sharply-tuned central region of the dorso-ventral isofrequency axis. They reported a high degree of regional variability in onset latency along the dorso-ventral extent and across animals. Some animals that were studied did not even show a clear relationship between minimum onset latency and other spectral parameters analyzed. The present study finds patchy nonhomogeneous distributions of minimum onset latency consistent with the earlier report. The inversive relation between neurons with short and long minimum latent periods is reflected in the spatial distributions of onset latency with and without continuous background noise. In both background listening conditions nonhomogeneous patches of neurons with similar latent periods are manifest; however, the spatial organization of these patches are reversed. Despite these changes, there are distinct regions where no change in onset latency occurred with the addition of continuous background noise. The finding that inversive effects in both Q10 and onset latency occur with this alteration in background listening environment is consistent with the idea that sharpness of tuning and onset latency are integrally related for spectral analysis processing. The current data suggests that by changing the latency of neurons through the presence of continuous background noise, the members forming the cell assemblies that show synchronized/correlated responses regroup in response to the change in acoustic background. That is, cell assemblies may recruit different neurons into short or long-latency assemblies based upon changes in the signal-to-noise ratio of the acoustic context.

2.3.4 The effect of binaural interactions on functional topography in AI

Previous studies examining the organization of receptive field properties to tone burst stimulation have carried out their research using contralateral stimulation as opposed to binaural stimulation as was the case in the current study. The higher spatial frequency observed in the Q10 distribution along the isofrequency contours manifest across several animals in this data may be a reflection of a modulation of frequency selectivity by the binaural interaction properties of these low-frequency cells. Roughly 50-65% of the cells within AI are EE in binaural behavior while EI cells are represented by 33-34% of the cells in AI (Imig and Adrian, 1977; Middlebrooks, Dykes, and Merzenich, 1980; Phillips and Irvine, 1983). While EE cells are sensitive to interaural disparities in the phase of low-frequency tonal stimuli (Reale and Brugge, 1990), EI cells are sensitive to interaural disparities in the intensity of high-frequency tonal stimuli (Brugge, et al., 1969; Phillips and Irvine, 1981). The topographies reported in this study are representative of neurons with primarily low CF whose responses may not only be shaped by the summation behavior of the cells, but also by a function of neural sensitivity governed by interaural phase disparity. It has been proposed by others that the unaligned, but superimposed functional maps created by convergent inputs may be essential for the cortical representation of multiple characteristics of incoming complex sounds (Mendelson, Schreiner, Sutter, 1997; Schreiner, 1991). Under conditions such as binaural stimulation, the neural responses of cells when spatially distributed across cortex may reflect modulatory influences from different stimulus and receptive field parameters.

2.3.5 Evidence of convergence in primary auditory cortex

Convergent input onto cortical cells from subcortical and cortical contributors have previously been demonstrated physiologically (Anderson, Knight, and Merzenich, 1980; Phillips and Cynader, 1985; Phillips and Hall, 1986), anatomically (Code and Winer, 1985; Imig and Reale, 1981; Imig and Brugge, 1978), and neuropharmacologically (McKenna et al., 1988; Metherate and Weinberger, 1989). Evidence from these studies has shown that multiple convergent inputs to a cortical cell shape the excitatory and inhibitory areas of its receptive field space (Phillips and Hall, 1986). They also contribute to different temporal components of the cortical neuron's response (McKenna et al., 1988) and responses to different tone frequencies within the response area (Metherate and Weinberger, 1989).

The present findings of systematic shifts in Q10 and minimum onset latency in the presence of continuous background noise, particularly in the spatial distributions of these receptive field parameters within each isofrequency contour, are more evidence indicative of convergence onto the auditory cortical cell. Not only are the findings consistent with previous reports, but they show how functional organizations within primary auditory cortex are systematically modulated by convergent inputs.

2.3.6 Implications on representational invariance of complex sound processing

In this chapter, continuous background noise has been found to change the spatial distribution of receptive field parameters expressed by Q10 and minimum onset latency. It is possible that these changes found in the functional topographies of Q10 and latency due to a change in the acoustic environment reflect a mode by which the cortical system is capable of making generalizations of complex sounds. The mechanism by which complex sounds are represented in the auditory cortex must include a process that eventually renders them invariant in different acoustic environments. A normalization process such as a shift in the spatial distribution of spectral bandwidth under different background listening contexts may be an example of how invariance of sensory stimuli are created by cortical cells. This sort of normalization process necessitates a neural assembly approach of complex sound processing in the primary auditory cortex.

2.3.7 Closing remarks

In summary, the effects of continuous background noise on functional topographies in AI can be considered a result of binaural interactions of and from converging inputs onto cortical cells. The topographic changes in spectral parameter organization due to a change in the acoustic context may reflect a normalization process which allows the cortical system to generalize auditory information in various background environments. Further study of the effects of continuous background noise on the binaural processing of complex sounds will elucidate whether or not generalization of complex sounds across various signal-to-noise contexts are evident in the primary auditory cortex.

Chapter 3: The cortical representation of speech syllables and the effects of continuous background noise on its representation

3.0.0 Introduction

The search for a neural representation of communication sounds has inspired auditory research to study neurophysiological aspects of complex sound processing at various levels of the central nervous system in the mammalian model. In the analysis of neural representations of speechlike sounds and animal vocalizations, the spectral and temporal domains of the acoustic signal have emerged as being fundamentally encoded stimulus features represented at various stages of the auditory nervous system. At the level of the auditory nerve in cats, it has been shown that spectral and temporal features of speech or speechlike sounds are represented by rate-place or temporal-place coding of neural discharge patterns (Delgutte and Kiang, 1984; Sachs and Young, 1979; Sinex and Geisler, 1983; Young and Sachs, 1979). Spectral and temporal characteristics of vowel sounds have also been manifest in firing rate and temporal response pattern of neuronal subpopulations in the anteroventral cochlear nucleus of cats (Blackburn and Sachs, 1990). At the level of the primary auditory cortex in marmoset monkeys, spectro-temporal characteristics of the species-specific vocalization were found to be temporally integrated and spectrally represented by the synchronization of neural activity from spatially-dispersed cortical cell assemblies (Wang, et al., 1995).

In the study of communication calls, the issue of behavioral relevancy of the vocalization has been argued to be a factor in shaping neural representations of stimulus features at the level of the cortex due to significant influences from auditory environment and experience (Merzenich, 1988, 1990). Exposure to specific extractable acoustic-phonetic features of communication calls such as speech has been found in human infants to be critical for developing perceptual boundaries specific to their native language; while general, more basic auditory processing, which is capable of discerning differences between all phonetic units based upon acoustic differences, has been found to exist for infants and mammals with no prior speech experience to shape these perceptual boundaries (see review by Kuhl, 1994). It has been shown that the perceptual feature, voice onset time (VOT), of speech syllables is represented by time-locked neural activity in the primary auditory cortex of monkeys (Steinschneider et al., 1982, 1990) and cats (Eggermont, 1995). In these experimental studies, however, the animals had not been trained or exposed to the speech stimuli in order to give these physiological responses any behavioral salience. Thus, the findings from these previous studies merely reflect the ability of the primary auditory cortex to respond to subtle changes in the acoustics of the speech stimuli. Addition of the behavioral component to physiological experiments examining the neural representations of speech stimuli would not only provide behavioral salience to neural responses but would also provide insight to specific cortical reorganizations to trained stimuli.

The question of how varying signal-to-noise levels affect the neural representation of speech is essential in discovering and understanding neuronal mechanisms that allow for speech intelligibility in various contextual backgrounds. Studies of background noise effects on vowel-like stimulus representation in the auditory nerve have shown that neural adaptation effects, resulting from discharge to the continuous background noise, causes a reduction in response amplitudes (Delgutte and Kiang, 1984; Sachs, Voigt, and Young, 1983) resembling background noise effects on simple stimuli (Geisler and Sinex, 1980; Rhode et al., 1978; Smith, 1979). The effect of background noise was found to have differential impact on the rate-place encoded, spectral profile which was dependent upon the spontaneous rate fiber type (Sachs, Voigt, and Young, 1983). The temporal encoding of the vowel-like stimuli (Sachs, Voigt, and Young, 1983) and consonant-vowel syllables (Geisler and Gamble, 1989; Silkes and Geisler, 1991), in contrast to rate encoding, was more robust to changes in signal-to-noise levels. At the level of the auditory cortex, the effect of continuous background noise on tonal stimulation has been found to be fundamentally different than in the periphery (Phillips and Cynader, 1985). That is, cortical cells do not fire tonically in the presence of continuous background noise. From experiments described in the first chapter of this dissertation, one of the basic experimental findings of background noise effects on response properties was the inversion of bandwidth near threshold in single units. In the second chapter, it was found that the topography of sharpness of tuning near threshold based upon multi-unit responses also manifest inversions in bandwidth tuning at spatially-localized areas in the presence of background noise. Based on these findings, it was hypothesized that subpopulations of neurons based upon their sharpness of tuning characteristic play a role in creating generalization mechanisms in the presence of varying signal-to-noise ratios in the environment. Little is known, however, about how neural representations of speechlike sounds are affected by changes in signal-to-noise levels.

Therefore, there were three specific aims in the current series of experiments. The first objective was to study how complex sounds such as speech syllables are represented in the primary auditory cortex; second, how does continuous background noise affect the cortical representation of speech; and third, what is the effect of behavioral training on the cortical representation of speech.

3.1.0 Methods

3.1.1 Stimuli

Four consonant vowel (CV) stimuli were synthesized using a Klatt-model speech synthesizer (SenSyn) which is based on the source-filter theory. The four CV stimuli were /be/, /pe/, /ke/, and /ko/. Each CV stimulus was 250 ms in duration. The fundamental frequency declined linearly during the stimulus duration from 120 Hz to 100 Hz. The beginning and endpoints of the three formant frequencies for each CV stimulus were defined by:

/be/: F1(350-550 Hz), F2(1400-1700 Hz), F3(2100-2500 Hz)
/pe/: F1(NV-550 Hz), F2(1400-1700 Hz), F3(2100-2500 Hz)
/ke/: F1(NV-550 Hz), F2(2150-1700 Hz), F3(2150-2500 Hz)
/ko/: F1(NV-550 Hz), F2(2150-1700 Hz), F3(2150-2500 Hz)

where NV is the absence of voicing and thus, no F1 at the onset of the consonant burst for the three unvoiced CV stimuli. With the exception of /be/ where the onsets of the consonant burst and voicing are simultaneous, the onset of voicing in the unvoiced CV stimuli occurred 60 ms after the consonant burst. The other acoustic parameters used to construct the CV stimuli were defined by:

/be/: AF(85, 0), (90, 60), (95, 0);
AB (175,0), (180, 60), (220,60), (225, 0)
/pe/: AH (95, 0), (100, 60), (140,48), (160, 0);
AF (85, 0), (90, 60), (95, 0);
AB (175, 0), (180, 60), (220,60), (225, 0)
/ke/: AH (100, 20), (140, 50), (160, 50), (180, 0);
AF (95, 0), (100, 60), (160, 54), (165, 0);
A2F (95, 0), (100, 60), (160, 54), (165, 0);

A3F (95, 0), (100, 40), (160, 60), (165, 0)

/ko/: acoustic parameters were the same as for /ke/

where A(2 or 3)F controls the amplitude of frication and AH controls the amplitude of aspiration.

3.1.2 Behavioral Training

Four cats were trained to detect the presence of a CV syllable and to discriminate between two CV syllables. The cats were trained in an acoustically transparent test cage housed within a single walled acoustic chamber lined with echo-attenuating foam. Auditory CV stimuli were delivered in the free-field from a single speaker located directly over and in front of the animal's head. Randomly generated continuous background noise was delivered in the free-field from another speaker located directly in front of the animal's head. All aspects of the stimulus generation and data collection were automated and computer controlled.

The cats were trained to initiate a trial by licking food from a metal spoon. The duration of each stimulus trial was 1 s. The CV stimuli was presented every 2 s. In the detection task, the cat was required to maintain contact with the spoon through continuous licking, and to break contact upon the detection of the CV stimulus within the last 200 ms of the stimulus trial window. Failure to release before this time window was scored as a "miss" and was punished in a non-noxious electric shock to the cat's tongue. A release in the absence of a CV stimulus was scored as a "false-positive" with no behavioral consequence. A release prior to this time window in the presence of a CV stimulus was scored as a "hit", and no shock was delivered. CV stimuli were presented in 20% of the trials during a given session. The frequency of stimulus delivery was chosen to minimize the amount of time the cat spent away from the food.

In the case of the discrimination task, the cat was required to lick a food-dispensing spoon continuously during the presentation of a "safe" CV stimulus, and to release contact from the spoon during the presentation of a "warn" CV stimulus before the last 200 ms of the stimulus trial session. Behavioral responses to the "warn" CV stimulus were scored the same as in the detection task. The durations of the stimulus trial interval and interstimulus interval were the same as in the detection task.

Two of the four animals were trained on two discrimination tasks: 1) to discriminate between /pe/ vs. /ko/ and then 2) /pe/ vs. /be/. The other two animals were trained on only one discrimination task which was to discriminate between /be/ vs. /ko/. All four animals were required to undergo the detection task for all four of the CV stimuli. Three of the animals were also required to perform the speech detection and discrimination tasks in the presence of continuous background noise set at one to three signal-to-noise ratio levels.

Detection and discrimination thresholds were obtained using the method of constant stimuli. In the detection task, the level of CV stimulus was stepped randomly in increments of 2 dB over a 20 dB range. For the discrimination task, in order to prevent intensity cues resulting from spectral differences in the CV stimuli, the stimulus presentation level of both the "safe" and "warn" CV stimuli were stepped randomly in increments of 3 dB over a 21 dB range which was the same for both stimuli. Performance (P) across test sessions was defined by:

$$P=(H-FP)/(1-FP)$$

where H is the "hit" rate for that CV (H=# correct responses/# "warn" stimulus trials) and FP is the "false positive" rate (FP=# incorrect spoon releases/#total stimulus trials). The psychometric function, which was determined across testing sessions due to the few number of stimulus trials at each stimulus level per session, was plotted as a function of "warn" stimulus level and fit to a nonlinear sigmoid function (S) defined by:

$$S=1/(1+e^{\wedge}(-4*T*(P-R)))$$

where T is the slope of the sigmoidal function and R is the psychometric threshold at 50% of maximum performance. The statistical nonlinear fit also calculated the X^2 statistic to test goodness of fit (Mathematica 2.2.2).

3.1.3 Electrophysiology

The primary auditory cortex on the right hemisphere was defined electrophysiologically in the four cats trained in the psychophysical task and three naive cats that had not been trained in any listening task. The animal surgery performed on these animals have been described in detail in chapter 2. Briefly, anesthesia was induced by an intramuscular injection of ketamine HCL (10 mg/kg) in combination with the tranquilizer acepromazine maleate (0.28 mg/kg). The femoral vein was cannulated and sodium pentobarbital was given i.v. to maintain an areflexic state of anesthesia (25-29 mg/kg induction: titrated to effect). Heart and respiration rate were monitored. The animal's core body temperature was maintained at 37.5 °C by wrapping a temperature-controlled heating pad around the animal's lower body. A craniotomy exposed the primary auditory cortex. Following the removal of the dura and the application of a film of silicon oil to protect the cortical tissue, a photographic image of the cortical vasculature and sulci was recorded as a reference for the topographical location of the electrode penetration sites. After surgery, subsequent delivery of the anesthetizing agents during the recording study were infused intravenously with ketamine (1-2 mg/kg) with diazepam (1-2 mg/kg). All electrode penetrations were approximately perpendicular to the cortical surface. The electrode insertion points were marked onto the image of the cortical surface with reference to the vasculature. Recording electrodes were lowered to depths of 700 to 1100 microns corresponding to cortical layers 3 and 4.

The stimulus generation and data acquisition methods for recording frequency response areas (FRA) have been described in detail in chapter 2. Briefly, single tone

bursts of 50 ms duration (3 ms linear rise/fall time) were generated by a microprocessor (TMS32010) and presented binaurally via calibrated earphones (STAX 54). The intensity of the tones varied in 5 dB steps over a range of 70 dB, while the 45 different frequency values covered a range of 4 to 5 octaves centered around the estimated CF of the unit. The pure tone series were summed with or without continuous background noise depending upon the stimulus condition. Following the recording of the FRAs at each penetration location, post-stimulus time histograms (PSTHs) were recorded using all four of the CV stimuli and two vowels, which were sectioned from the CV stimuli, delivered at three to five different stimulus levels and at five different signal-to-noise ratio levels. Each stimulus was presented once every second across twenty repetitions.

For analysis, the response properties that were measured from the FRAs included CF, threshold, Q10, onset latency, and firing rate at the transition point. Spectral parameter maps of CF, threshold, Q10, and onset latency in conditions with and without continuous background noise were constructed as a reference for the cortical maps of speech-activated neural activity. Summed spike activity from the PSTHs of each multi-unit were binned in 1 ms windows and normalized by the firing rate at the transition point of the unit's FRA. Normalizing the speech-driven neural activity to the best firing rate in the FRA reduces the size-bias of the multi-unit response and, therefore, allows the comparison of response magnitude across units. The averaged PSTH, spectral profile, population histogram, and cortical speech maps were derived for each CV at each signal level and signal-to-noise ratio level for each cat. The peak firing rate, latency of the peak response, and peak width were recorded for each response peak to each CV stimulus at each signal level and signal-to-noise ratio level for each cat. Analysis of variance measures were used to find significant parametric factors contributing to the variance found in the response peaks of the speech-activated neural activity.

3.2.0 Results

3.2.1 The Cortical Representation of Speech and the Effects of Background Noise on the Representation

Previous studies of sound processing in the primary auditory cortex of cats have discovered a complex, yet systematic organization of neural response properties that is believed to provide a basic framework for understanding how complex signals such as animal vocalizations and speech sounds may be represented at the cortical level. In ordinary listening situations, speech is commonly embedded in some level of background noise which may vary across time. At low levels of background noise, these variances in the speech-in-noise signal have minimal impact on the speech perception in the normal hearing individual. Neural processes that shape the normalization of speech sounds in various background listening contexts may be found in the modulation of general auditory processing that occurs in the primary auditory cortex. However, the role of the systematic distribution of receptive field properties in the coding of complex sounds and the generation of representational invariances have not been explored.

In chapters one and two, the effects of continuous background noise on tonal response properties in A1 neurons was most obviously seen in an increase of response thresholds and in the differential change in sharpness of tuning. It was hypothesized that the subdivision of A1 neurons based upon their sharpness of tuning characteristic may reveal a normalization process that allows for wide range of categorical invariances found in speech discrimination in the presence of continuous background noise. The following section explores the consequences of noise on the cortical representation of complex sounds by examining the effects of continuous background noise on A1 neural responses to synthesized speech syllables.

Data from seven cats were collected and analyzed in this series of experiments. Four of the cats had been trained to make discriminations of speech syllable sounds in order to assure that these complex, yet biologically-relevant sounds are acoustic stimuli that can be distinguished by non-human subjects. The other three cats were naive listeners of speech sounds. Comparison of data collected from naive and speech-trained cats were made to examine whether the biological relevance of select speech sounds served to emphasize particularly relevant parts of their cortical representation and to enhance the preservation of neural activity to the stimuli in the presence of background noise. Post-stimulus time histograms of spike responses to speech-in-noise stimuli were collected for at least four continuous background noise levels and then compared to those collected for speech stimuli without background noise. Cortical response maps to speech with and without continuous background noise were also compared for each speech syllable. Although the main emphasis of the current study was to study the effects of continuous background noise on speech sound representation in A1 neurons, comparisons of cortical responses to two different speech stimuli were also made. In addition, comparisons were made between signal level versus signal-to-noise ratio effects upon cortical responses to the speech stimuli.

3.2.2 Summary of the behavioral thresholds

Psychophysical thresholds for both a detection and discrimination task were obtained with a conditioned avoidance task at various background noise levels and are summarized for each of the four trained cats in figure 3.1. Thresholds were defined as the stimulus level at which the animal was able to respond correctly 50% of the time based upon the psychometric fit of the data (see Methods). In all four cases, the mean detection thresholds (averaged across all CV stimuli) were lower than the discrimination thresholds by 5 dB to 35 dB. Despite interanimal variability of the measured thresholds, detection and discrimination thresholds did significantly increase with

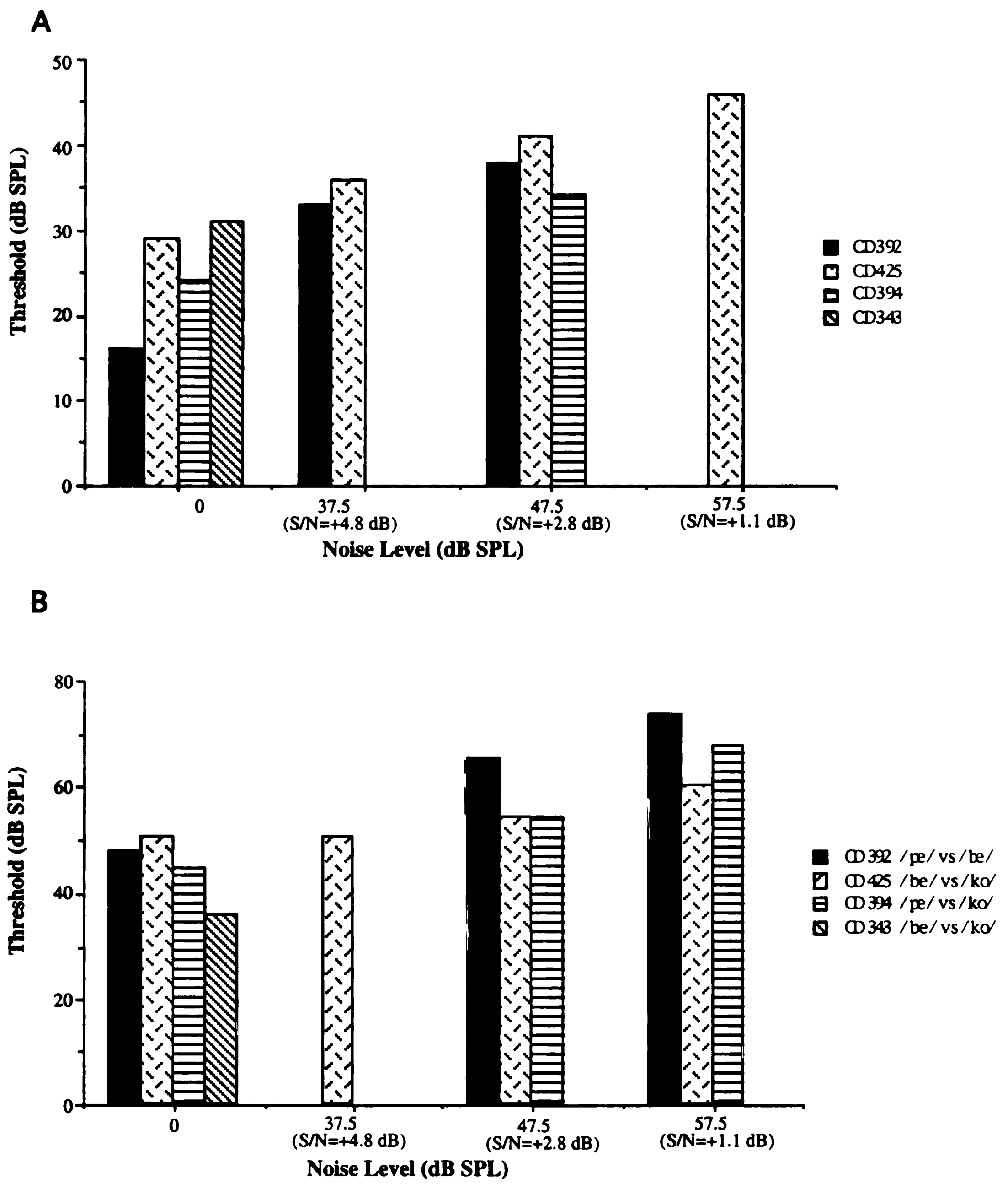


Figure 3.1 Behavioral thresholds from a speech detection and discrimination task

Psychophysical thresholds at four signal-to-noise levels for four animals obtained through behavioral training for (A) a detection and (B) a discrimination task. Both detection and discrimination thresholds increased ~3 dB per 10 dB increase in background noise. The psychophysical thresholds were defined as 50% of maximum performance for the summed across several sessions (see Methods for a detailed description of the threshold calculation). The discrimination thresholds from each animal were obtained from different speech discrimination tasks: two were trained on /be/ vs. /ko/ while the other two were trained on /pe/ vs. /ko/ and then /pe/ vs. /be/. One of the cats from the latter group was not able to yield behavioral thresholds for the more difficult discrimination task, and thus, thresholds for the easier task (/pe/ vs. /ko/) are depicted instead.

increasing levels of background noise. The average detection and discrimination thresholds increased by ~3.6 dB per 10 dB/Hz increase of the noise level.

3.2.3 Spectrograms and the mean cortical response histograms

The two speech stimuli selected for current analysis were the Klatt-synthesized syllables /be/ and /pe/ which share the same formant structure and differ only in the timing of their voicing onset. The spectro-temporal difference between the two speech syllables are observed by contrasting the spectrograms for the two stimuli in figure 3.2. The onset of voicing for the unvoiced consonant-vowel (CV) syllable /pe/ occurs 60ms after the onset of the consonant component of the stimulus. The formant frequencies were F1=550 Hz, F2=1700 Hz, F3=2500 Hz with the transitions starting at 350 Hz, 1400 Hz, and 2100 Hz respectively (see Methods for a detailed description of the acoustic parameters used in the synthesis of the CV stimuli).

The spectro-temporal differences between the two stimuli appear to be maintained in the accumulated cortical responses to /be/ and /pe/ which is manifest in the mean post-stimulus time histograms displayed below the spectrograms for each CV. The mean cortical response revealed one peak to the onset of the stimulus /be/ and two peaks for /pe/ which corresponds in time to stimulus onset and to voicing onset. The magnitude of the peak response to the CV stimuli were level dependent as is depicted in figures 3.3A-D. The mean magnitude of the peak to the CV stimulus /be/, is seen in figures 3.3A for a naive animal and 3.3C for a trained animal, decreased with decreasing stimulus level. The latency of the peak response changed very little, on the average 3 ms over a 30 dB reduction of level. The mean magnitude of the first peak to the CV stimulus /pe/, as seen in figures 3.3B for a naive animal and 3.3D for a trained animal, decreased with decreasing stimulus level while the mean magnitude of the second peak remains relatively unchanged until the first peak response disappears at which point the second peak decreased as well. Again, it was noted that although response magnitude

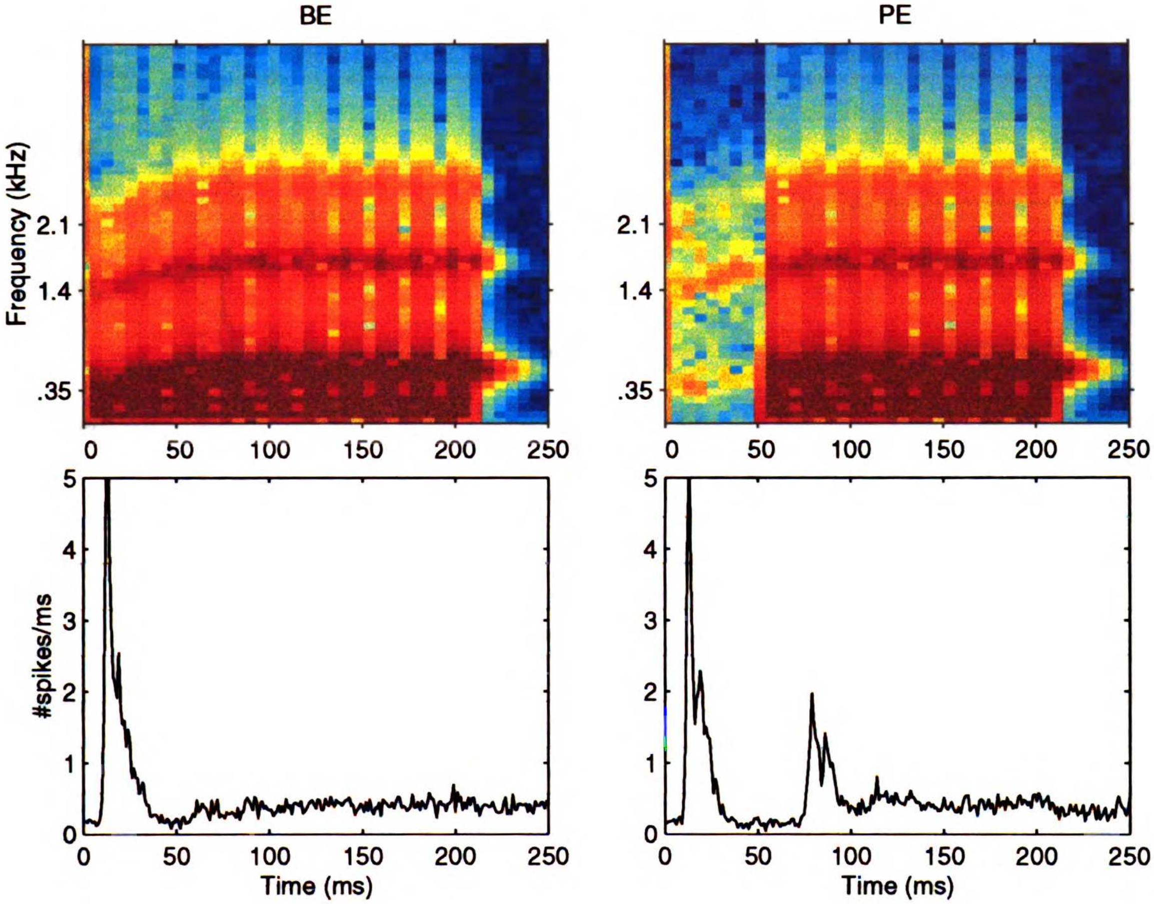
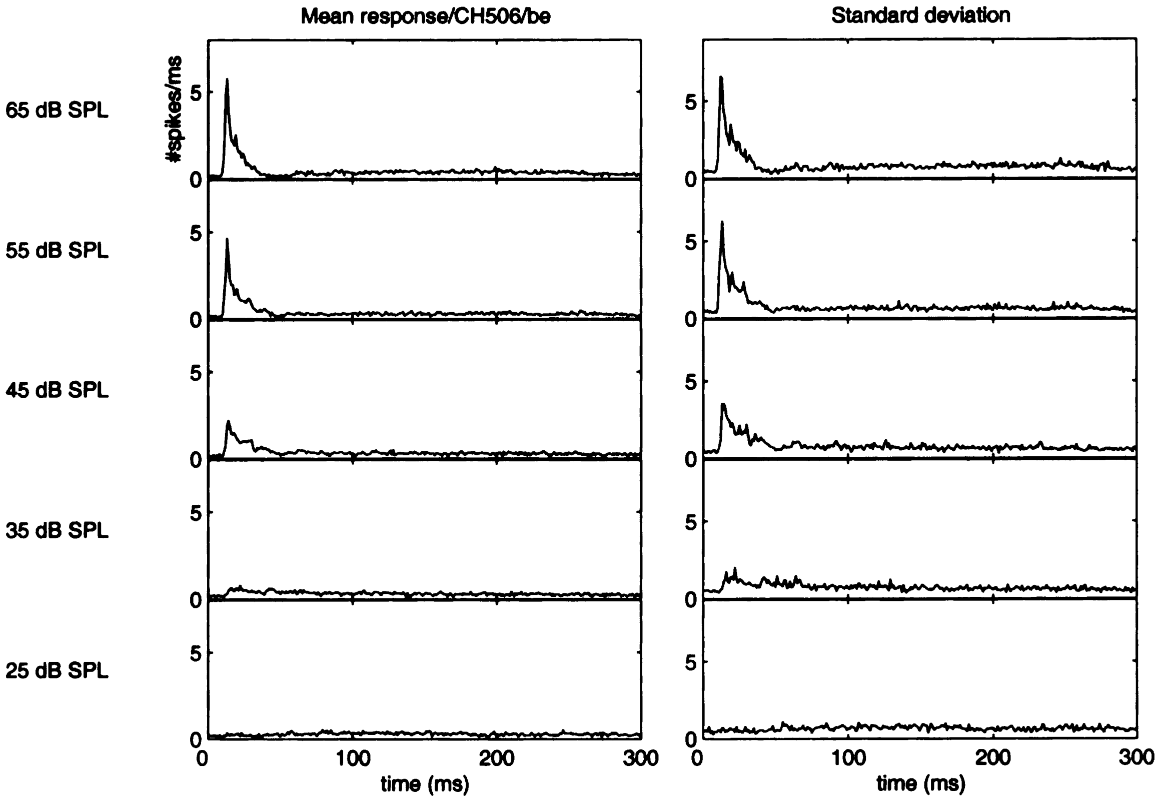


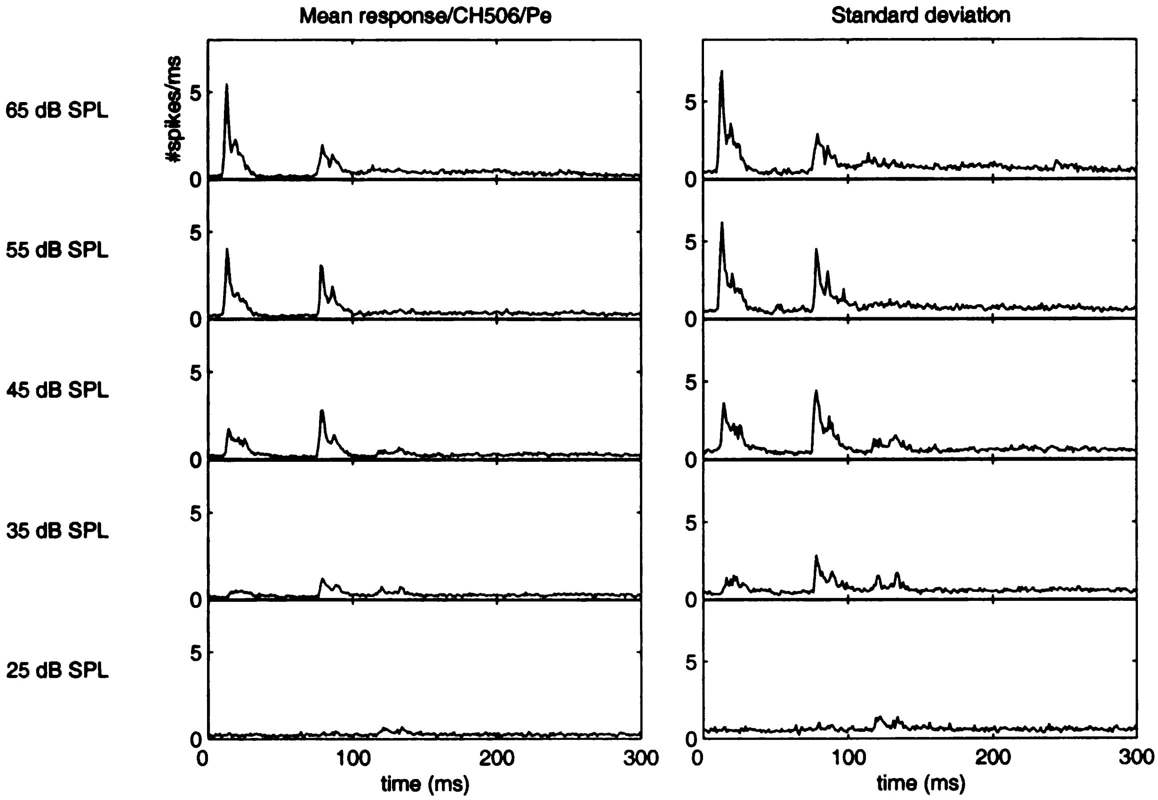
Figure 3.2 Spectrograms and the Mean Cortical Response

The spectrogram representation of the CV stimuli /be/ and /pe/ contrast in the onset of voicing that is delayed by 60 ms for /pe/. Underneath the spectrograms, the mean response post-stimulus time histograms (PSTH) for the CV stimuli contrast in the number of phasic responses evoked by the speech syllables. The morphology of the peak response to the stimulus onset was similar for the CV stimuli. An interval of suppressed neural activity immediately follows the first peak response for either stimulus. The second response peak to /pe/ coincides with the delayed onset of voicing.

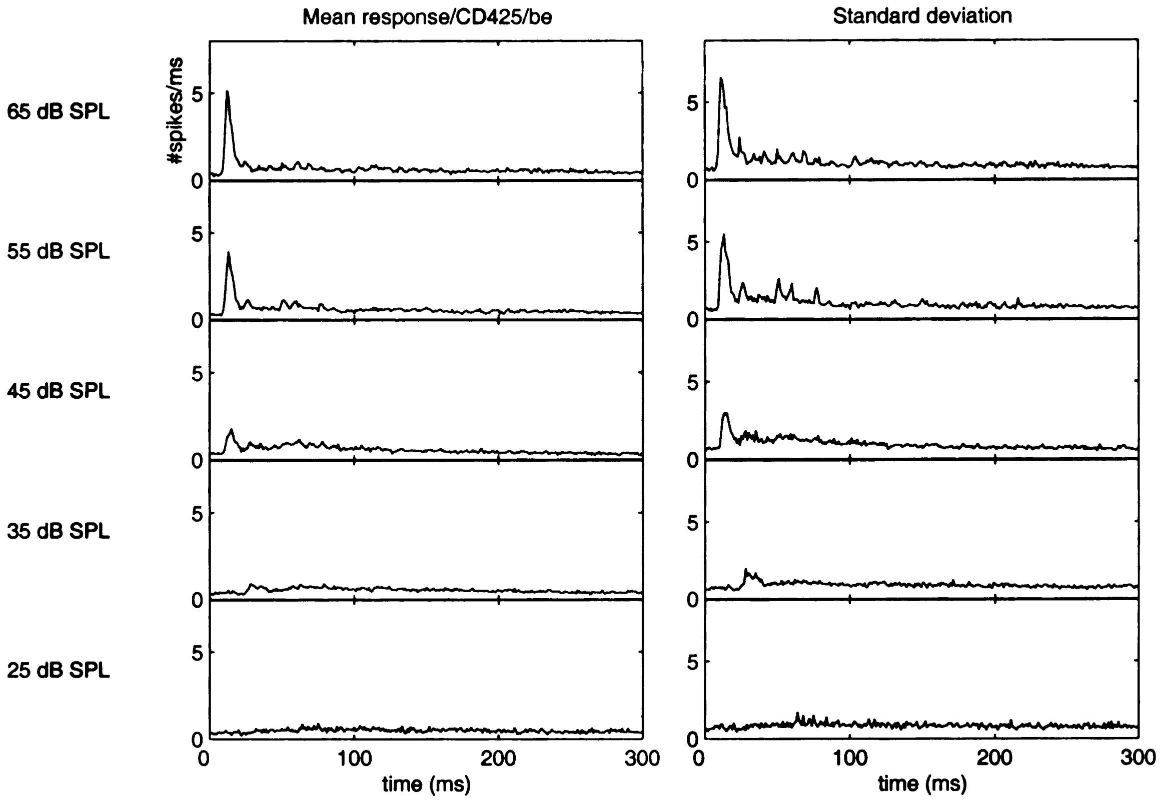
A


Mean response/CH506/be
Standard deviation
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
#spikes/ms
0
5
0
100
200
300
time (ms)

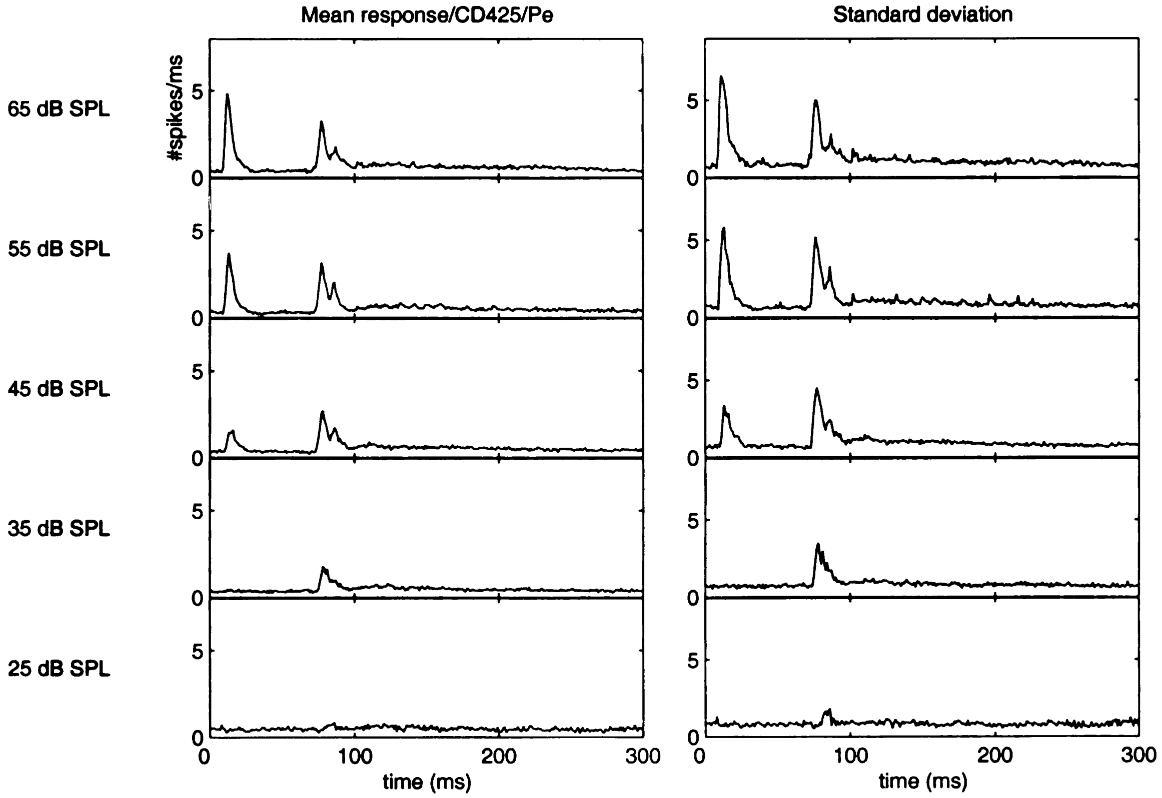
B


Mean response/CH506/Pe
Standard deviation
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
#spikes/ms
0
5
0
100
200
300
time (ms)

C


Mean response/CD425/be
Standard deviation
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
#spikes/ms
0
5
0
100
200
300
time (ms)

D


Mean response/CD425/Pe
Standard deviation
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
#spikes/ms
0
5
0
100
200
300
time (ms)

Figures 3.3A-D Stimulus Level Profile of Mean PSTH Response

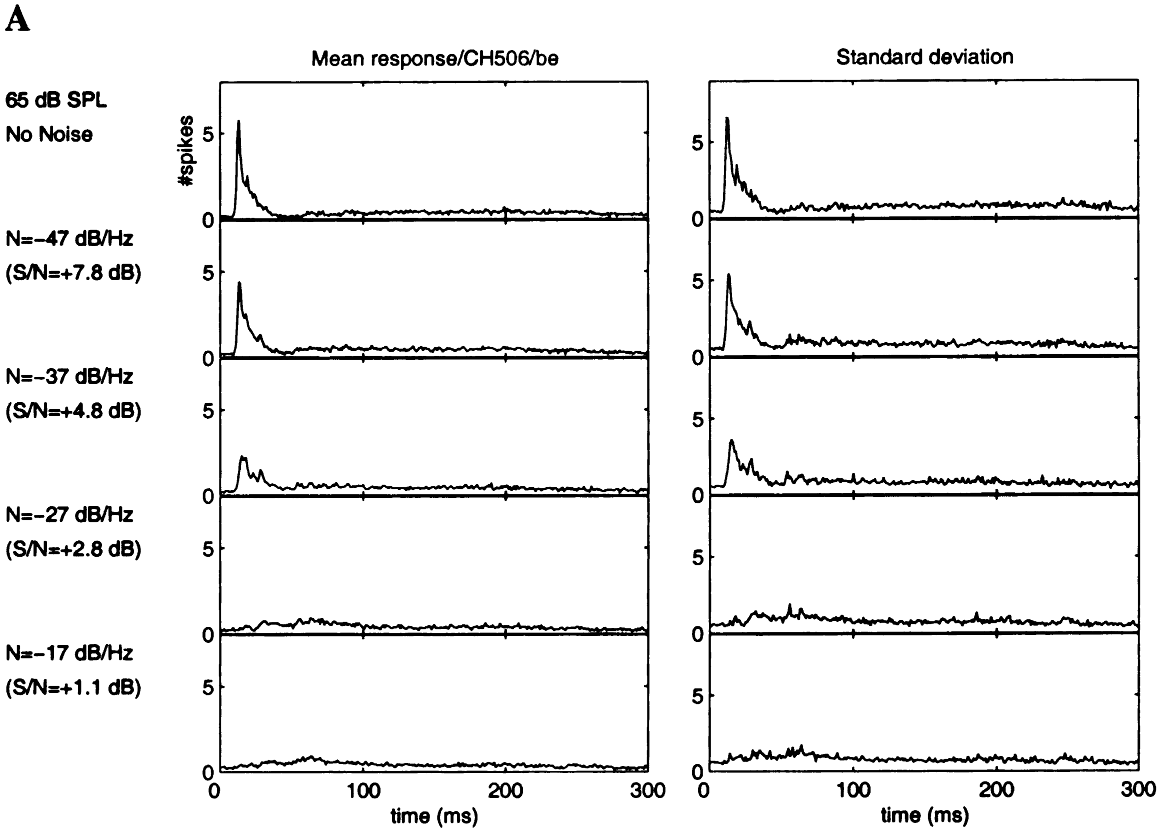
Stimulus profiles of the mean response and standard deviation to the presentation of /be/ and /pe/ for a naive animal are presented in **A** and **B** respectively and for a trained animal in **C** and **D** respectively. The mean peak response and standard deviation of the mean response in **A** and **C** decreased with decreasing stimulus level while the peak response latency changed minimally (~3 ms across the entire level range) at each stimulus presentation level. For the stimulus presentation of /pe/, the magnitude of the first peak in the mean biphasic response and standard deviation in **B** and **D** decreased with decreasing stimulus level while the second peak increased slightly with a slight decrease in stimulus level, then also decreased with subsequent lowering of the stimulus level. There was no difference in mean response between naive or trained animal.

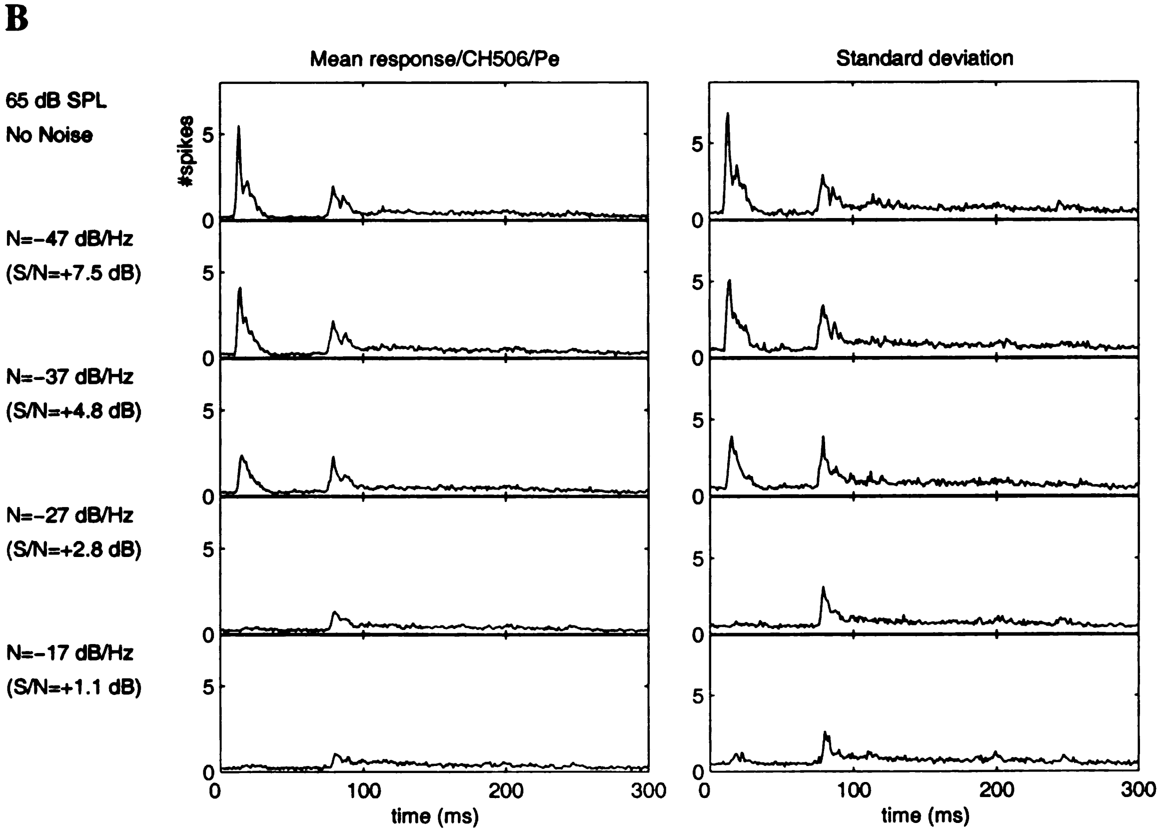
appeared to be level dependent, latency of either first or second peak response appeared to be less affected by level.

The changes in the mean PSTH, under stimulus presentation conditions with continuous background noise were incredibly similar to those by stimulus level changes. Again, the magnitude of the peak response to the CV stimuli were signal-to-noise level dependent as is depicted in figures 3.4A-D. The mean magnitude of the peak to the CV stimulus /be/, as seen in figures 3.4A (naive animal) and 3.4C (trained animal), decreased with the addition of continuous background noise and eventually disappeared with increasing levels of background noise. Although the magnitude of the mean response decreases with increasing background noise levels, the mean onset latency remains relatively steady at each level with only up to a 3 ms prolongation on average at low signal-to-noise levels.

The standard deviation of the mean response to /be/ was similar in magnitude, latency, and width to the actual mean response at each signal level and signal-to-noise level indicating a wide dispersion of peak response values to the stimulus. The dispersion found is attributable to the differences in neural response properties of the neurons sampled.

In the comparison of the mean PSTH response with regards to behavioral training, there appeared to be little difference in the temporal response of the mean PSTH to the CV stimulus /be/ between either the trained or naive animal. Similarly, the mean magnitude of the first peak response to the CV stimulus /pe/, as seen in figures 3.4B (naive animal and 3.4C (trained animal), decreased with the addition of continuous background noise and eventually disappeared with increasing levels of background noise. The second peak response occurring at roughly 80 to 85 ms post-stimulus onset, on the other hand, appeared to be more robust under the influence of background noise. The amplitude of the secondary peak was still visible at noise levels where the primary peak had already disappeared. Again, the response latencies of both the


A
Mean response/CH506/be
Standard deviation
65 dB SPL
No Noise
#spikes
N=-47 dB/Hz
(S/N=+7.8 dB)
N=-37 dB/Hz
(S/N=+4.8 dB)
N=-27 dB/Hz
(S/N=+2.8 dB)
N=-17 dB/Hz
(S/N=+1.1 dB)
0
5
0
100
200
300
time (ms)


B
Mean response/CH506/Pe
Standard deviation
65 dB SPL
No Noise
#spikes
N=-47 dB/Hz
(S/N=+7.5 dB)
N=-37 dB/Hz
(S/N=+4.8 dB)
N=-27 dB/Hz
(S/N=+2.8 dB)
N=-17 dB/Hz
(S/N=+1.1 dB)
0
5
0
100
200
300
time (ms)

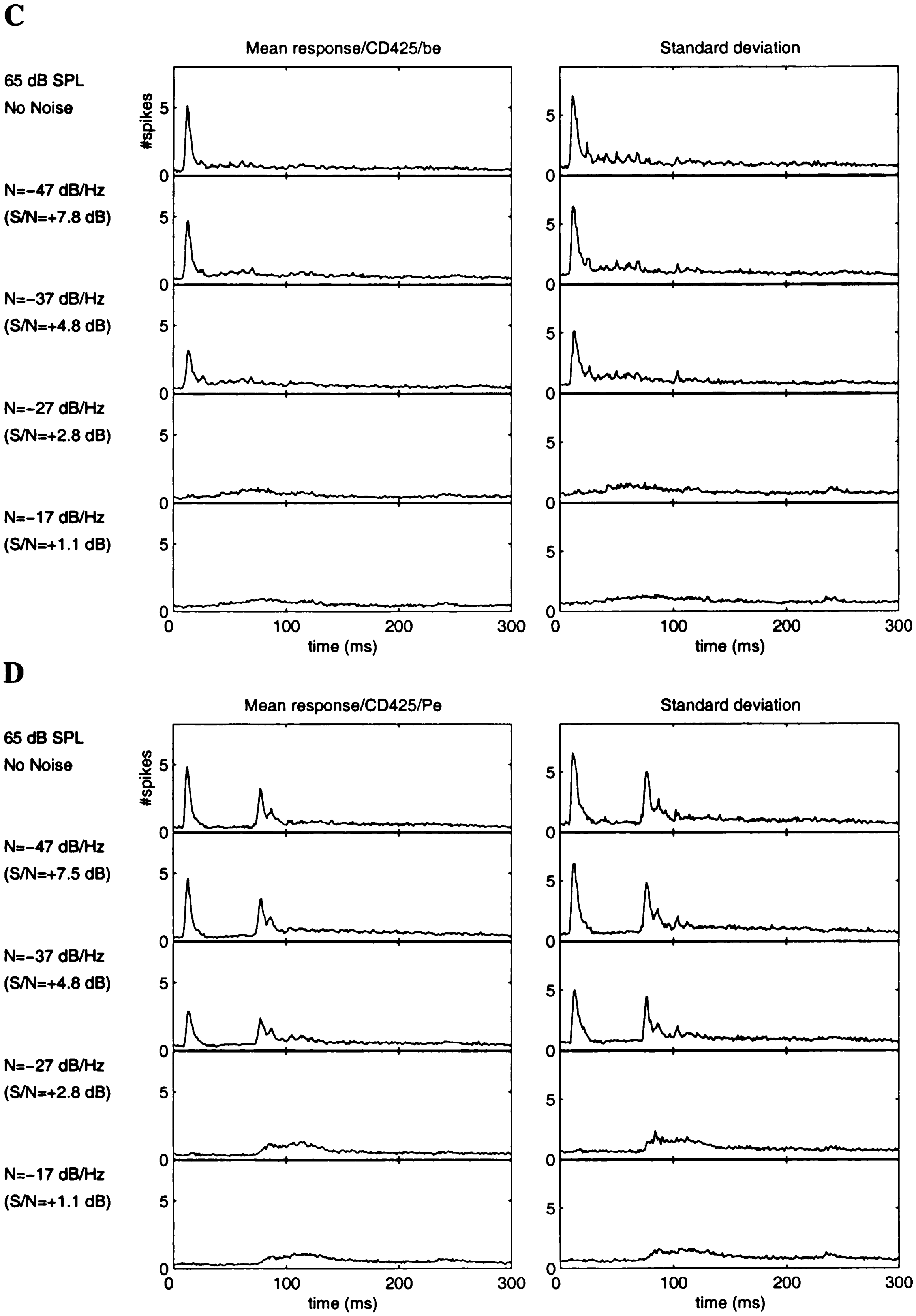

C
Mean response/CD425/be
Standard deviation
65 dB SPL
No Noise
N=-47 dB/Hz
(S/N=+7.8 dB)
N=-37 dB/Hz
(S/N=+4.8 dB)
N=-27 dB/Hz
(S/N=+2.8 dB)
N=-17 dB/Hz
(S/N=+1.1 dB)
#spikes
0
5
100
200
300
time (ms)
D
Mean response/CD425/Pe
Standard deviation
65 dB SPL
No Noise
N=-47 dB/Hz
(S/N=+7.5 dB)
N=-37 dB/Hz
(S/N=+4.8 dB)
N=-27 dB/Hz
(S/N=+2.8 dB)
N=-17 dB/Hz
(S/N=+1.1 dB)
#spikes
time (ms)

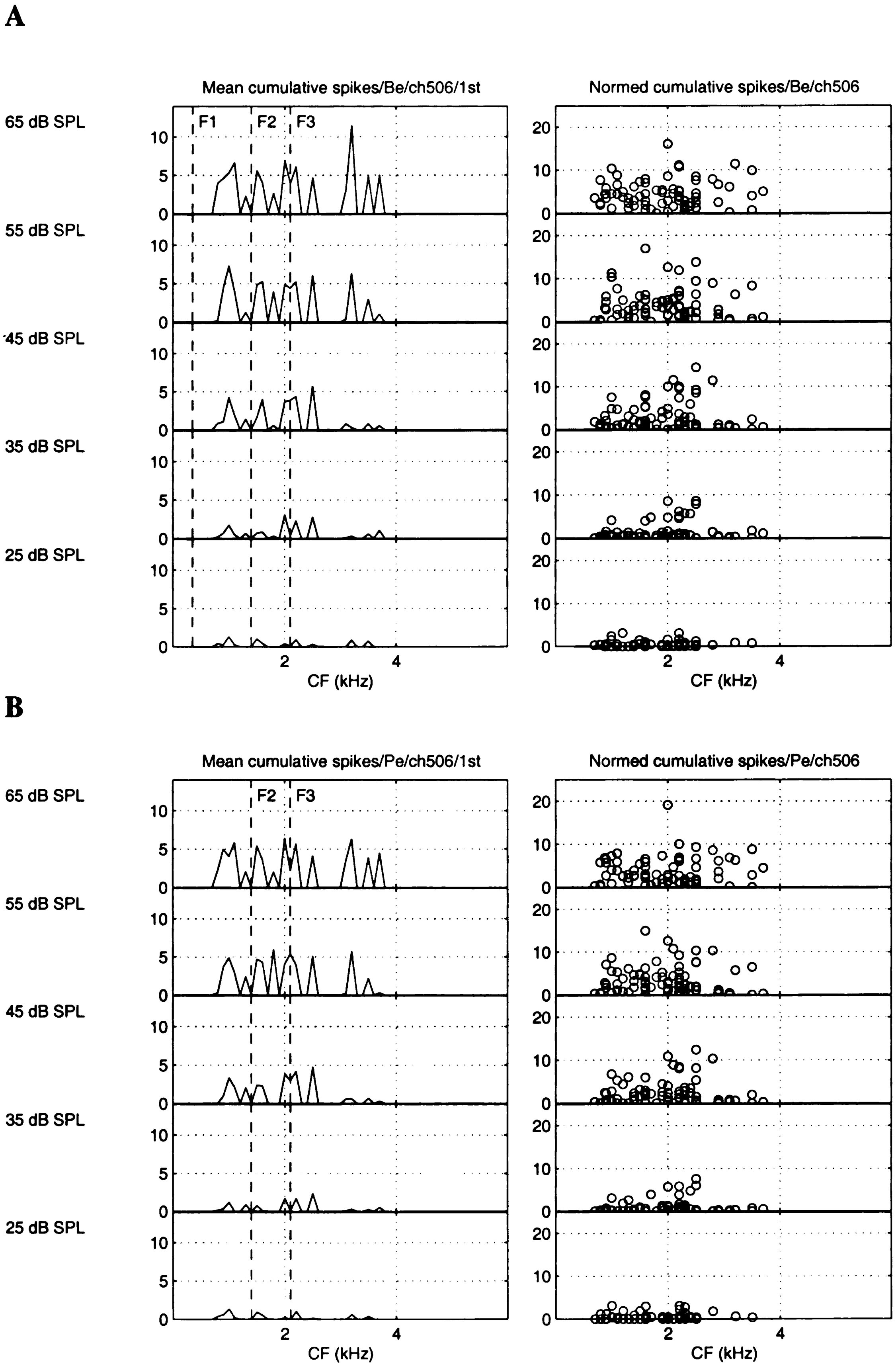
Figure 3.4A-D Signal-to-noise Level Profile of Mean PSTH Response

Signal-to-noise level profiles of the mean response and standard deviation to the presentation of /be/ and /pe/ for a naive animal are presented in **A** and **B** respectively and for a trained animal in **C** and **D** respectively. The mean peak response and standard deviation decreased with decreasing signal-to-noise level while the peak response latency changed minimally (~ 3 ms across the entire level range) with the 10 dB decrease each stimulus-to-noise level. For the stimulus presentation of /pe/, the magnitude of the first peak response decreased faster with 10 dB decreases in signal-to-noise level than for the second peak response. The response latency changed very little for either the first or second peak response. There was no observable difference in signal-to-noise level-dependent behavior in the mean PSTH response between naive and trained animals.

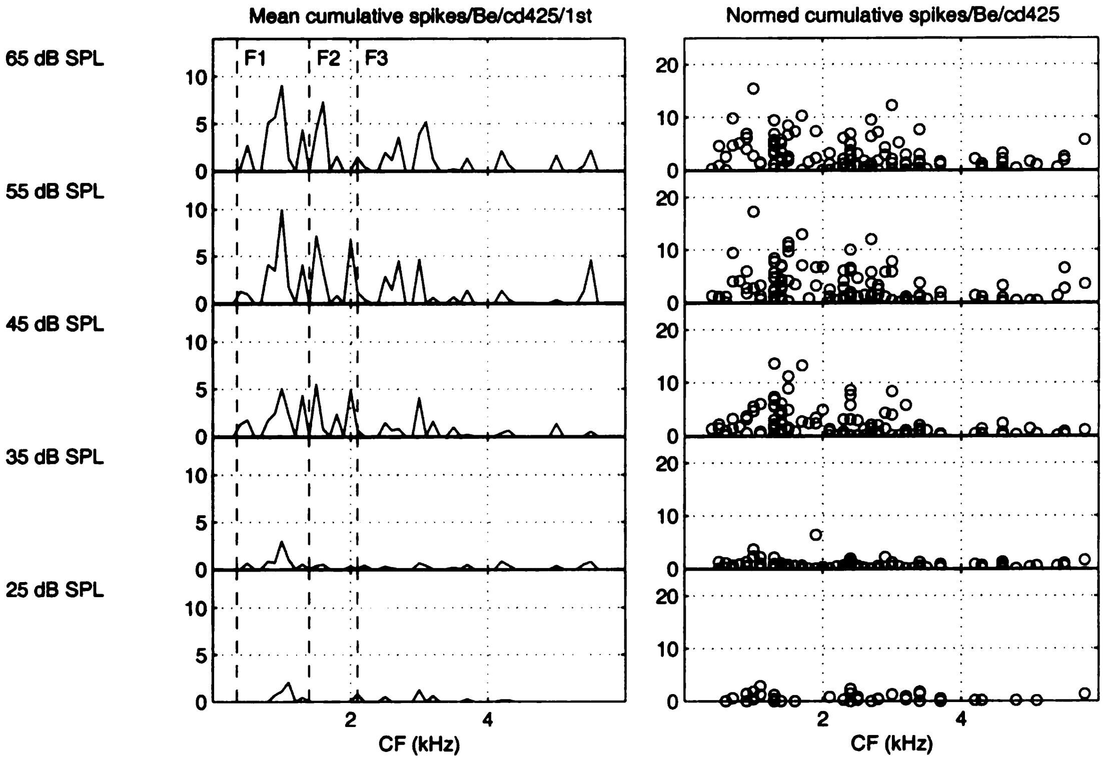
primary and secondary mean response peaks to the CV stimulus /pe/ were not affected by the increase in background noise levels. The standard deviation of both primary and secondary mean response peaks were similar in magnitude, latency, and width to the actual mean response peaks at each noise level where their corresponding peaks were visible. Again, there appeared to be no obvious difference in the temporal response of the mean PSTH to the CV stimulus /pe/ between either the trained or naive animal.

3.2.4 Spectral profile analysis

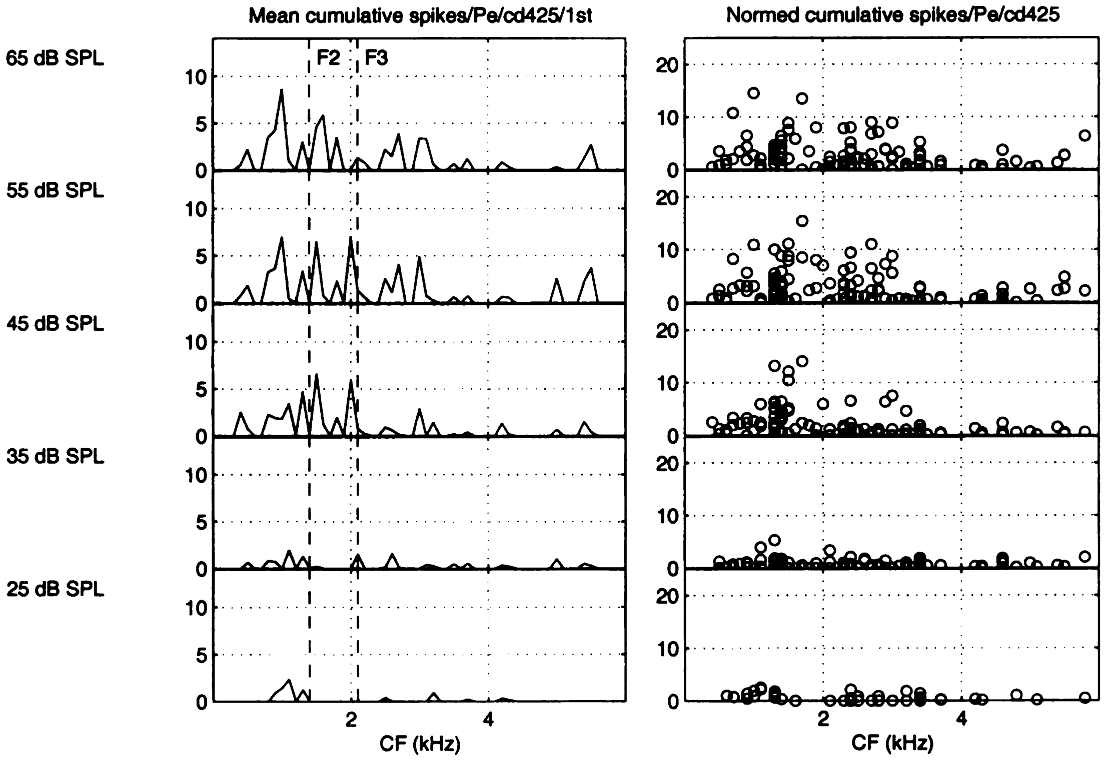
The basic tonotopic organization of AI suggests that spectral properties such as location and spacing of formants may be reflected in the firing rate distributable across AI. Thus, in an effort to examine whether the spectral content of the speech stimuli was captured by the neural activity of the cells classified by their CF, a spectral profile was constructed from the PSTH spikes windowed (5 ms window) around the mean peak responses to each CV stimulus at each stimulus and background noise level. These spectral profiles of the mean normalized cumulative spike activity and normalized cumulative spike activity for each stimulus level are displayed in figures 3.5A-H. Figure 3.5A shows the spectral profiles for the first peak response to /be/ from a naive animal. The three dotted vertical lines in the mean normalized cumulative spike plots mark the locations of the formant frequencies at the onset of the CV stimulus (F1: 0.35 kHz; F2: 1.4 kHz; F3: 2.1 kHz). The curve represents the mean cumulative activity at each CF. Areas where spike activity is zero imply that there were no cells with that CF sampled. An example of this occurrence is at the F2 location. The actual normalized cumulative spike data are displayed to the right side. The spike activity to the speech stimuli were normalized to the transition point of the cell's pure-tone response at CF. In the highest stimulus level condition (S=65 dB SPL), the peaks in the spectral profile of the neural response to speech, when plotted as a function of CF, appeared only to align with the formant frequencies of the speech stimulus at F3. However, due to the wide


A
Mean cumulative spikes/Be/ch506/1st
Normed cumulative spikes/Be/ch506
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F1
F2
F3
0
5
10
20
2
4
CF (kHz)
B
Mean cumulative spikes/Pe/ch506/1st
Normed cumulative spikes/Pe/ch506
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F2
F3
0
5
10
20
2
4
CF (kHz)

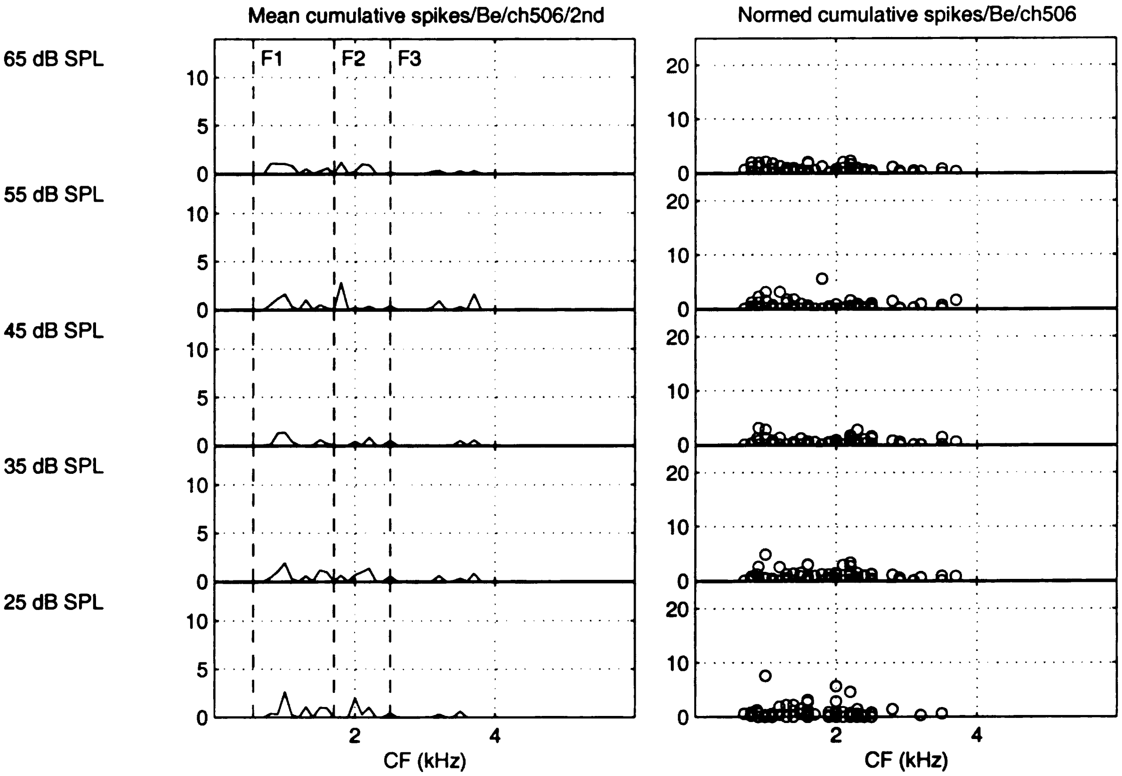
C


Mean cumulative spikes/Be/cd425/1st
Normed cumulative spikes/Be/cd425
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F1
F2
F3
0
5
10
20
2
4
CF (kHz)
CF (kHz)

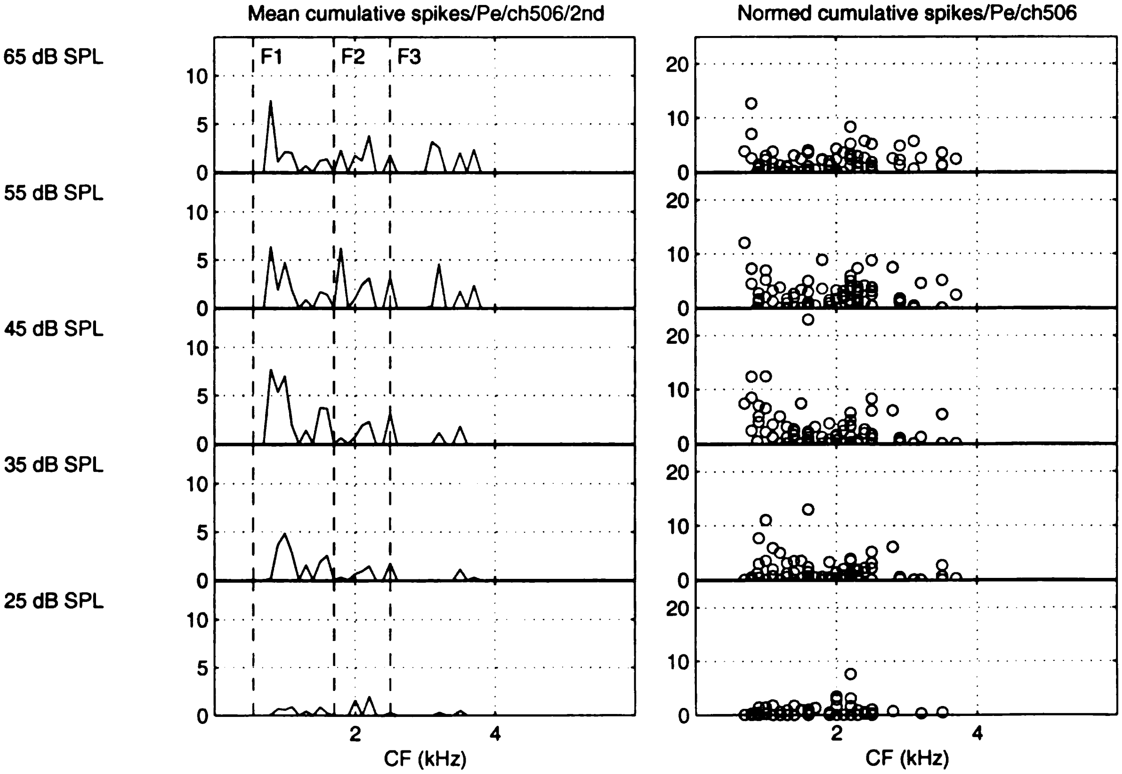
D


Mean cumulative spikes/Pe/cd425/1st
Normed cumulative spikes/Pe/cd425
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F2
F3
0
5
10
20
2
4
CF (kHz)
CF (kHz)

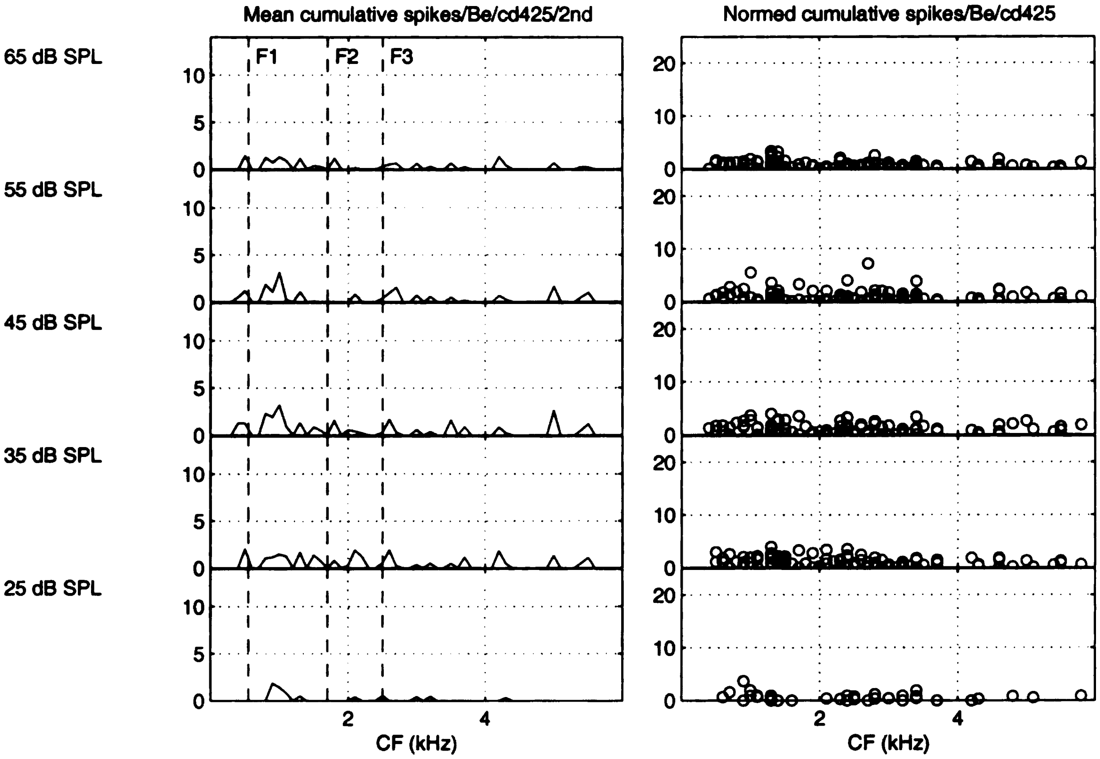
E


Mean cumulative spikes/Be/ch506/2nd
Normed cumulative spikes/Be/ch506
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F1
F2
F3
0
5
10
20
2
4
CF (kHz)
CF (kHz)

F


Mean cumulative spikes/Pe/ch506/2nd
Normed cumulative spikes/Pe/ch506
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F1
F2
F3
0
5
10
20
2
4
CF (kHz)
CF (kHz)

G


Mean cumulative spikes/Be/cd425/2nd
Normed cumulative spikes/Be/cd425
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F1
F2
F3
0
5
10
20
2
4
CF (kHz)

H

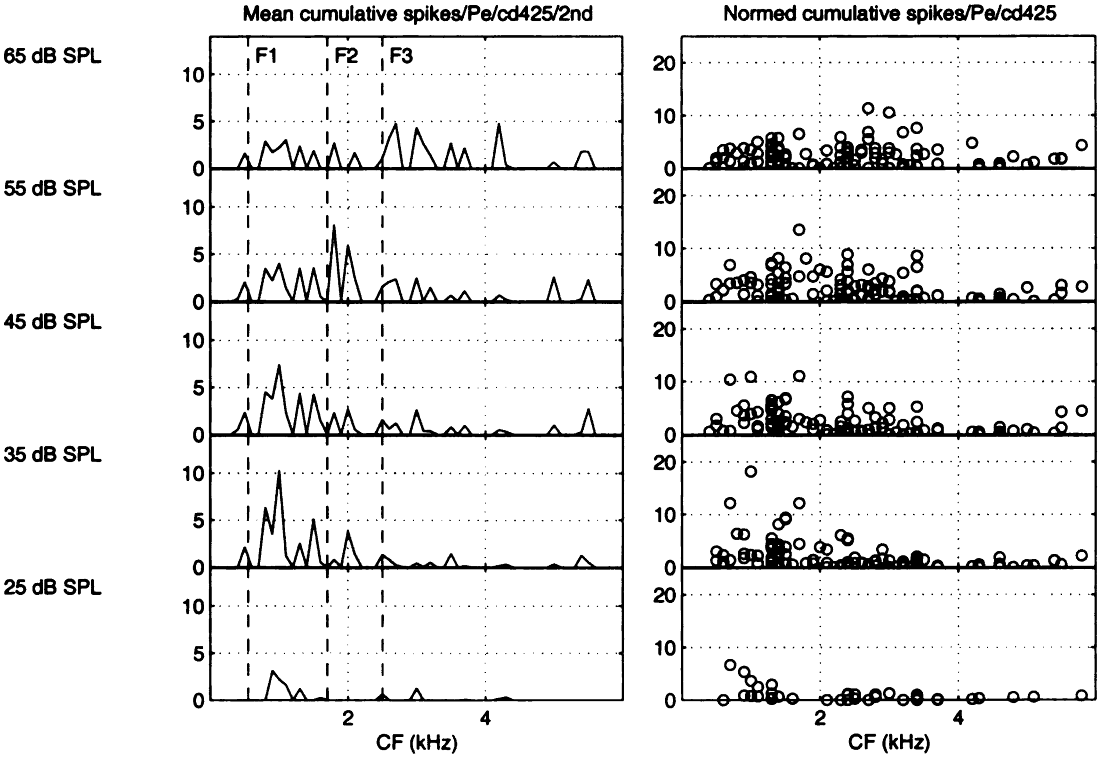

Mean cumulative spikes/Pe/cd425/2nd
Normed cumulative spikes/Pe/cd425
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
F1
F2
F3
0
5
10
20
2
4
CF (kHz)

Figure 3.5 Spectral Profiles of Mean Cumulative Spikes

The first and second response spectral profiles of mean cumulative spikes and normed cumulative spikes to the presentation of /be/ and /pe/ for a naive animal are presented in **A** and **B** (for the first response) and in **E** and **F** (for the second response) and for a trained animal in **C** and **D** (for the first response) and in **G** and **H** (for the second response). Spectral profiles, which plots average spike activity as a function of CF, are presented at five different stimulus levels for the peak response to each CV. In **A** and **B**, the spectral profiles for /be/ and /pe/ at 65 dB SPL were similar in appearance. Both showed a choppy spectral spread of neural excitation across the CF representation of the cortical area sampled. As stimulus decreased in 10 dB steps, not only did average spike activity decrease, but spectral spread of excitation was reduced also. In **C** and **D**, again the spectral profile for /be/ and /pe/ at 65 dB SPL were similar in appearance. The level-dependent behavior of the spectral profile for the first response in the trained animal was similar to the naive animal. In **E** and **F**, the spectral profile of the second response in the naive animal showed very little spectral energy for /be/ which is consistent with the absence of an identifiable second response in the mean PSTH response while for /pe/, the rate-encoded, spectral representation of the second response was different than for the first peak response. Although the spectral profile in the second peak response was less affected by decreases in stimulus level than in the first peak response, differential changes in average spike activity occurred in localized CF regions with decreases in stimulus level. In **G** and **H**, the spectral profile of the second response in the trained animal showed similar level-dependent behavior as in the naive animal.

extent and resulting overlap of frequency response areas between neurons with neighboring CFs, it is likely that peaks in the spectral profile such as between F1 and F2 and between F2 and F3 may be contributed to formant response activity within the frequency response area of the neighboring CF neurons. As stimulus level decreased, response magnitude across all CF locations decreased, with the exception of a slight increase at a restricted frequency location between F1 and F2 and at a frequency location slightly above F3 which then decreased with further decreases in stimulus level.

The spectral response profile of the neural response to /pe/ in the naive animal is displayed in figure 3.5B. The response profile was similar to that found in figure 3.5A for the CV stimulus /be/. The magnitude of the spectral response to /pe/ also decreased with decreasing stimulus level.

Figures 3.5C and 3.5D show the spectral response profiles to the first peak response for /be/ and /pe/, respectively, from a trained animal. In the highest stimulus level condition for either figures, two prominent spectral peaks were found. One between F1 and F2 and the other between F2 and F3. As in the case of the naive animal, neurons with CF exactly at the formant frequencies F1 and F2 were not sampled. The response to the F3 formant frequency was considerably less than in the naive animal. As signal level decreased by 10 dB, the spectral profile changed to reflect three prominent peaks: 1) the prominent peak below F2 sharpened and increased slightly in magnitude, 2) the peak slightly above F2 shifted its peak response to a lower frequency, and 3) a peak response near F3 appeared. The magnitudes of these spectral peaks decreased with subsequent decreases in signal level.

Figures 3.5E and 3.5F display the spectral response profiles of the second peak response for /be/ and /pe/, respectively, from a naive animal. As was seen in the mean PSTH response for the CV /be/, there was no secondary peak at any stimulus level; therefore, the spectral profile, as depicted in figure 3.5E, was correspondingly flat.

The mean PSTH response for the CV /pe/ did reveal a secondary peak around 80 to 85 ms. Therefore, the spectral profile for /pe/ at this later time frame differed from that for /be/. The formant frequency locations were shifted to reflect the endpoints of the formant transitions which would be represented in the spectrograms of the stimuli corresponding to this time frame (F1: 0.55 kHz; F2: 1.7 kHz; F3: 2.5 kHz). In figure 3.5F, there was only one prominent spectral peak located near F1 seen at the highest stimulus level. As the stimulus level decreased by 10 dB to 55 dB SPL, three changes in spectral profile occurred: 1) that prominent peak broadened in frequency composition, 2) a secondary peak slightly above F2 appeared, and 3) another peak at F3 increased in magnitude. As the stimulus level decreased another 10 dB to 45 dB SPL, the magnitude of the broadened spectral peak near F1 increased while the peak near F2 shifted to a frequency level slightly below F2. Further decreases in stimulus level resulted in decreases in the response magnitude.

Figures 3.5G and 3.5H display the spectral response profiles of the secondary peaks for /be/ and /pe/ from a trained animal. Again, the spectral profile differed between the two CV stimuli due to the absence of a prominent secondary peak for /be/. It is interesting to note, though, that with decreasing stimulus levels, a small spectral peak near F1 did appear for presentations of /be/ as manifest in figure 3.5G. In figure 3.5H, the spectral profile to /pe/ was best seen at lower stimulus levels. The spectral peaks appeared to shift from high CF to low CF with progressive decreases in stimulus level (e.g., peak near F3 at 65 dB SPL, peak near F2 at 55 dB SPL, peak between F1 and F2 at 45 dB SPL).

The spectral profiles of the mean normalized cumulative spike activity and normalized cumulative spike activity for each background noise level are displayed in figures 3.6A-H. Figure 3.6A shows the spectral profiles for the first peak response to /be/ from a naive animal. The spectral profile for the no noise condition was described earlier in figure 3.5A for stimulus level at 65 dB SPL. Compared to speech

stimulation in the absence of noise, at low continuous background noise levels (i.e., N=-47 dB/Hz), the mean activity increased at most CF sampled locations, excluding frequencies above 3 kHz where activity decreased in the presence of continuous background noise. Increasing the background noise level by 10 dB (i.e., N=-37 dB/Hz) reduced the mean activity back to levels similar to when no noise was present during stimulus presentation, excluding frequencies above 3 kHz. Further increases of the noise level virtually obliterated any neural response to the speech stimulus as depicted in the normed cumulative spike plots at the two higher noise levels (i.e., N=-27 dB/Hz and N=-17 dB/Hz).

The spectral response profile of the neural response to /pe/ in the naive animal is displayed in figure 3.6B. Not only were the spectral profiles for /pe/ and /be/ similar, which was mentioned earlier, but the effect of the continuous background noise on the mean neural response was also similar. There was a slight increase in mean activity with the addition of continuous noise at the lowest level, which decreases with subsequent increases in noise level.

Figures 3.6C and 3.6D show the spectral response profiles to the first peak response for /be/ and /pe/, respectively, from a trained animal. In the presence of low continuous background noise (i.e., N=-47 dB/Hz), the effect of noise was similar to that found in the naive animal. The magnitude of the response below 2.5 kHz increased, while that above 2.5 kHz decreased. Subsequent increases of background noise decreased the mean response for either CV stimuli, with the exception of an increase in mean activity near F3 at the intermediate noise level (i.e., N=-37 dB/Hz).

Figures 3.6E and 3.6F display the spectral response profiles of the second peak response for /be/ and /pe/ respectively from a naive animal. As was mentioned earlier, the spectral profiles of the secondary response differ between /be/ and /pe/ due to the absence of a secondary response for /be/ as was seen in the mean PSTH response. In figure 3.6F, the stimulus presentation in the absence of noise revealed only one

prominent spectral peak located near F1. The addition of continuous noise (N=-47 dB/Hz and N=-37 dB/Hz) again increased the magnitude of that peak and boosted the activity of the frequency area between F2 and F3. Further increases of noise continued to enhance neural activity near F1, while activity at the second peak declined.

Figures 3.6G and 3.6H display the spectral profiles of the secondary peaks for /be/ and /pe/ from a trained animal. Again, for similar reasons as discussed earlier, the spectral profile for /be/ was flat for the stimulus presentation in the absence of noise. However, in the presence of background noise and with progressive increases of the noise level, the activity level of neurons with the lowest CF appeared to increase resulting in the appearance of a prominent peak approximately at 1 kHz. For /pe/, the spectral profile of the response to the stimulus in the absence of background noise also resulted in a seemingly flat profile with a slight peak near F3. With the addition of background noise, neural activity increased in cells with the lowest CF, as was the case for /be/, and resulted in the formation of a spectral peak approximately at 1 kHz. The peak gradually decreased with further increases in the noise level.

In summary, spectral profile analysis of the cortical responses to speech sounds of moderate sound levels (~65 dB SPL) at periods of peak neural activity did not reflect any spectral semblance to the frequency content of the speech stimuli with or without the presence of background noise. At lower sound levels (~ 45 to 55 dB SPL), formant structure of the stimuli became apparent in the cortical responses which could be attributed to a reduction in peripheral adaptation and neural recruitment. Nonetheless, the limitations of using cumulative types of analysis for cortical responses is understood when considering the well-documented, systematic variation of neural response properties along the isofrequency dimension. Therefore, the spectral content of the speech stimuli may necessarily be reflected along receptive field dimensions other than frequency.

3.2.5 Neurogram: The population histogram

As an attempt to bypass the limitations that either the mean PSTH or spectral profile have, a population histogram of the data was constructed in order to visualize the temporal response to the speech stimuli across all CF units sampled. Figures 3.7A (naive animal) and 3.7B (trained animal) display a CF-dependent population histogram, or neurogram, in response to either /be/ or /pe/ underneath each mean PSTH response. The firing rate of each recording location is plotted as a function of time and CF. The degree of shading within the neurogram represents the magnitude of the neural response normalized to the pure-tone response at the CF-transition point. In response to /be/, there was one dark vertical band in the neurogram corresponding to the peak in the mean PSTH as opposed to the two dark bands corresponding to the two peaks in the mean PSTH for the unvoiced CV /pe/. In figure 3.7A, the latencies of the first peak for either CV appeared to lengthen with lower CF, whereas, the latencies of the second peak for /pe/ appeared relatively similar across frequencies. In figure 3.7B, the latencies of the first and second peaks (in the case for /pe/) were similar across frequencies. The resulting high synchronous activity across the frequency axis to speech stimulation is not a property of stimulus-relevant training since it was observed in trained as well as in untrained animals.

The response as depicted by the neurogram clearly showed level-dependent neural firing, as seen in figures 3.8A (naive animal) and 3.8C (trained animal). In these figures, the neurograms for /be/ and /pe/ were plotted for each 10-dB decrement in stimulus presentation level. Not only did the magnitude of the first peak response decrease with each 10-dB drop in stimulus level, but the frequency representation of the response decreased as well. The magnitude of the second response for /pe/ increased in certain frequency locations and decreased in others with decreasing stimulus levels.

Figures 3.8B (naive animal) and 3.8D (trained animal) display one neurogram per CV in the absence of noise and four additional neurograms per CV for the different

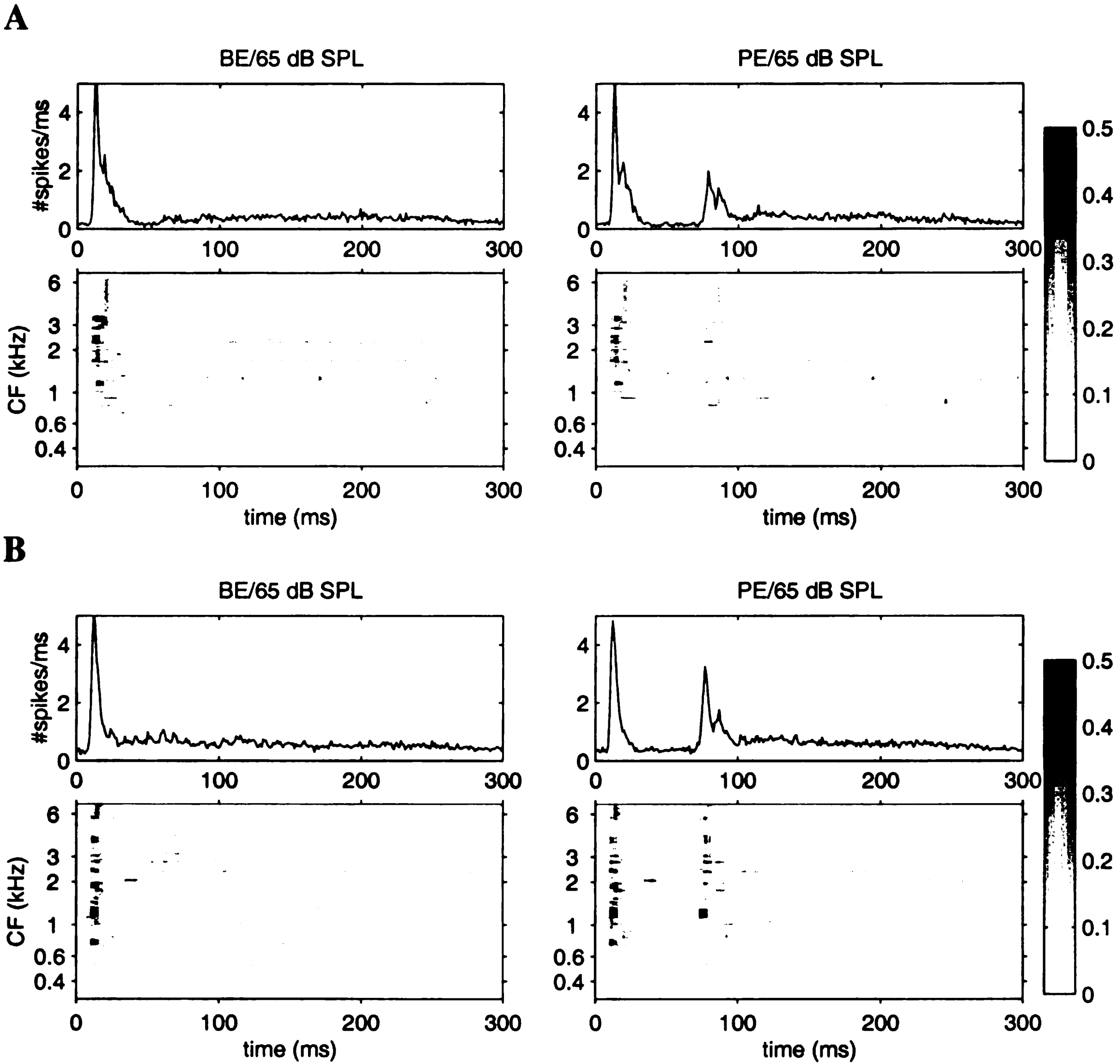
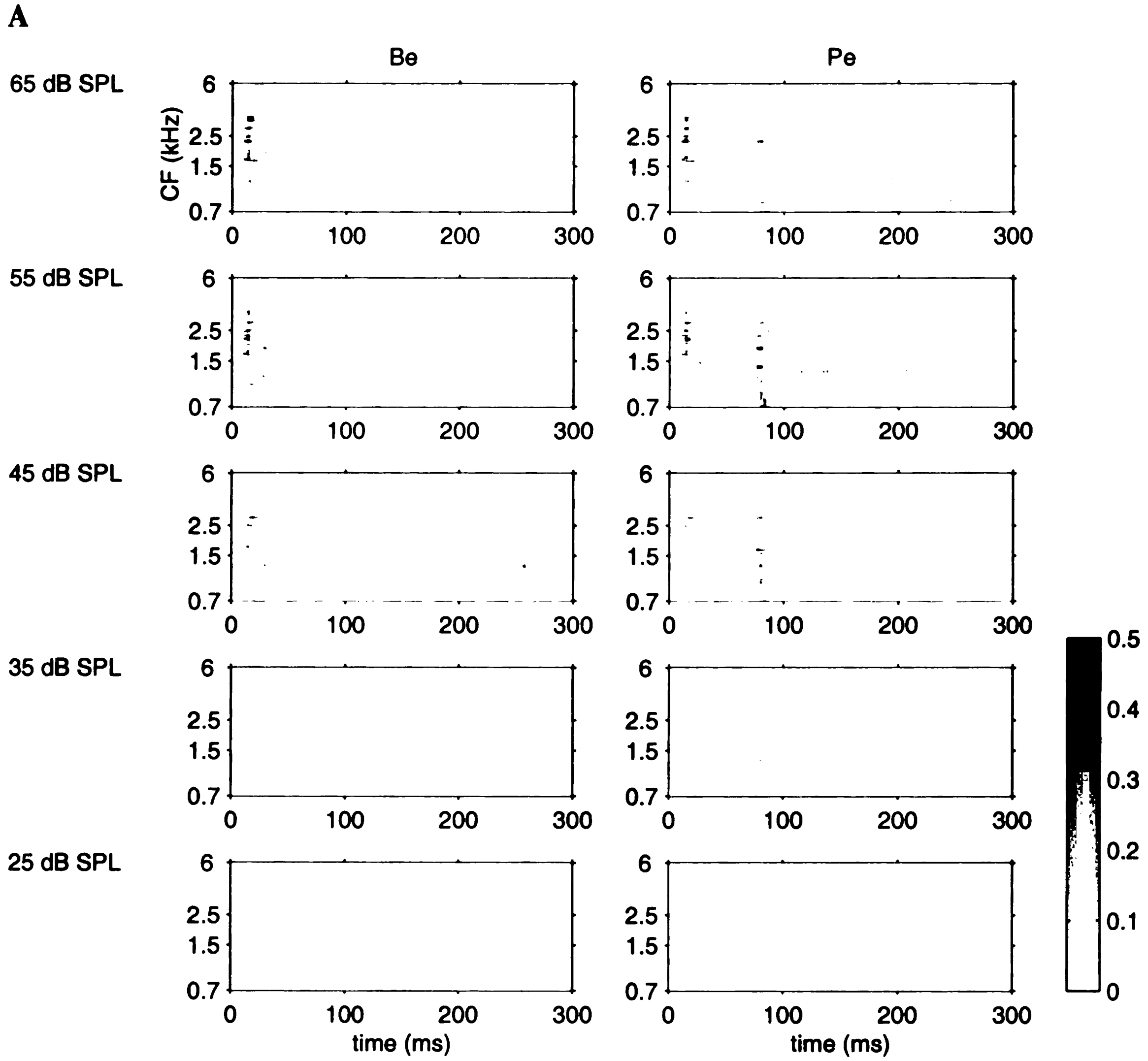


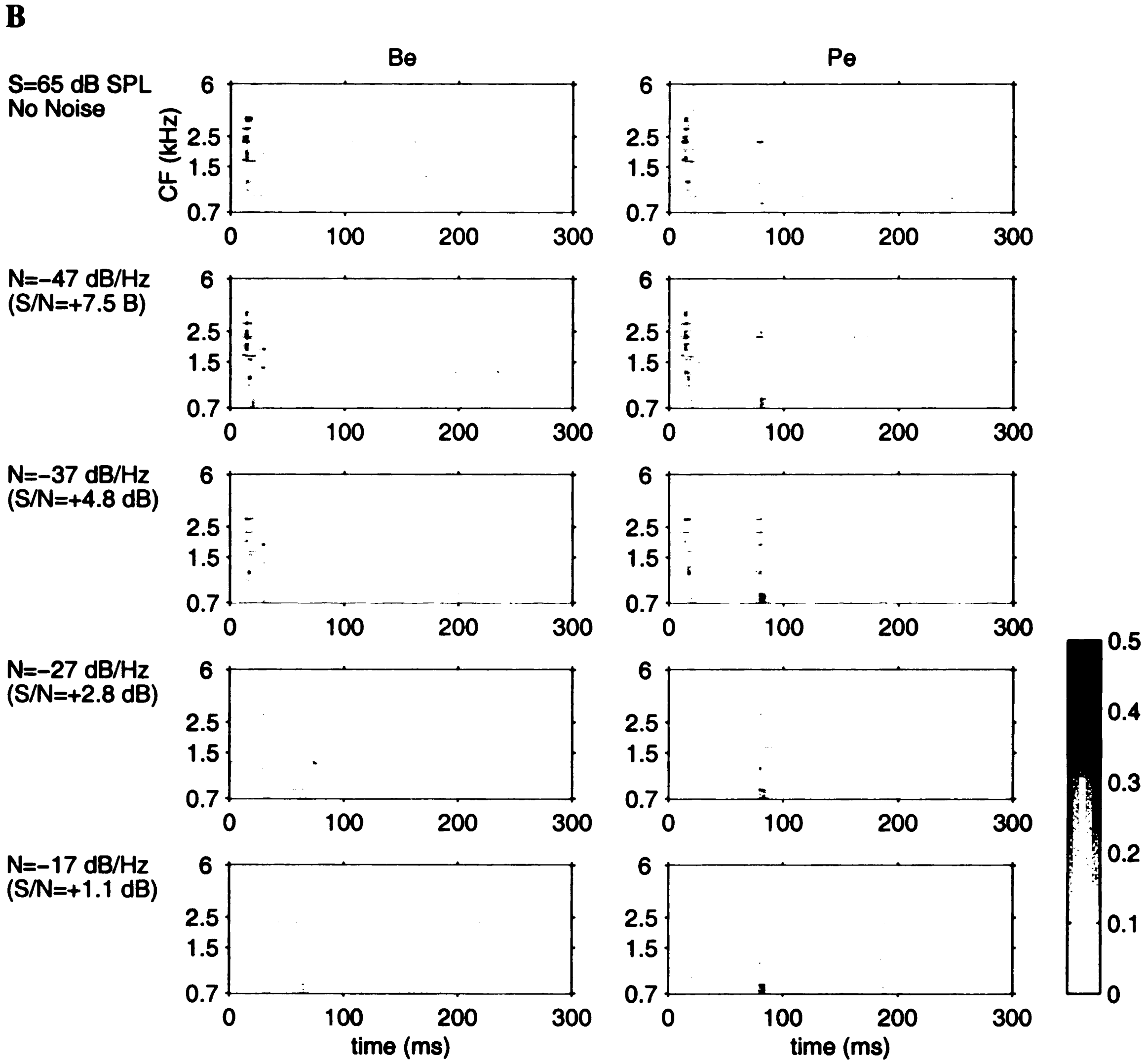
Figure 3.7 Mean Cortical Response and the Neurogram Representation

The mean PSTH response and the neurogram representation of /be/ and /pe/ are presented in **A** for a naive animal and in **B** for a trained animal at one intensity level (65 dB SPL). Descriptions of the mean PSTH response for /be/ and /pe/ were presented in figures 3.3A-D. They are presented here as reference for the temporal response manifest in the neurogram which was reconstructed from multiple-unit PSTH responses and organized by CF. In the neurogram, the darker pixels represent higher normalized spike activity (spike activity was normalized by the best firing rate to pure tones), while lighter ones represent lower normalized spike activity. Spike activity corresponding in time to the peaks of the mean PSTH was synchronized across CF, with some variability in onset which was observed to be longer at lower frequencies in **A**. Synchronization of spike activity to the onset of the consonant burst was more precise in **B**. The variability in the synchronization of spike activity to the onset of voicing, which was reflected in the peak width of the mean PSTH, did not appear dependent upon CF and was similar between either trained or untrained animal.

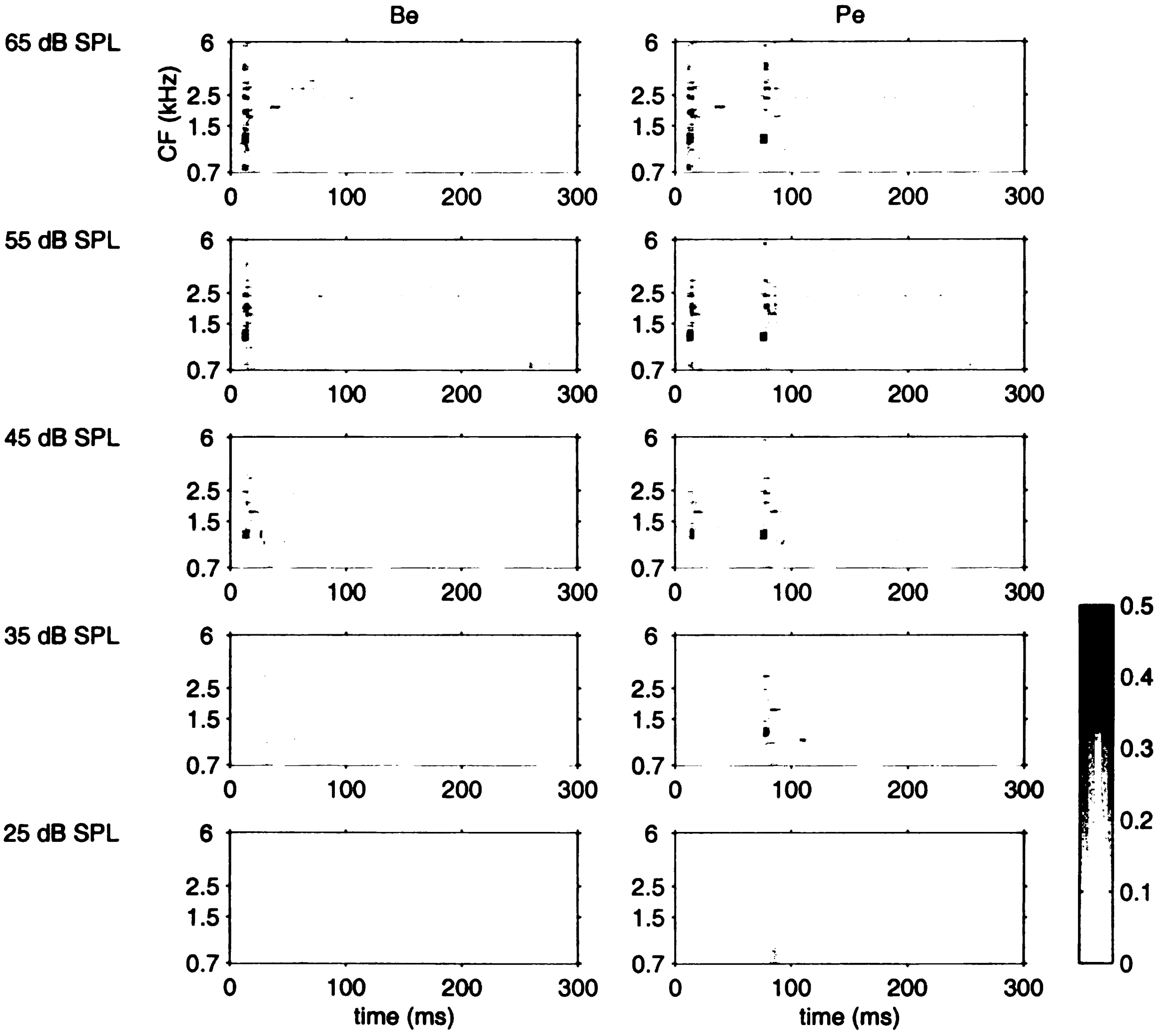
Figure 3.8A-D Neurogram of cortical activity to CV stimuli at different signal and signal-to-noise levels

The neurogram representation of /be/ and /pe/ were reconstructed for each signal and signal-to-noise level in the naive animal (seen in **A** and **B**, respectively) and in the trained animal (seen in **C** and **D**, respectively). In **A**, spike activity for /be/, which was strongest at the highest stimulus level (65 dB SPL), decreased in differing amounts across CF as stimulus level decreased. The onset latency of the synchronized response appeared to change as a function of CF in differing amounts with the decrease in stimulus level. For /pe/, spike activity to the consonant burst , which was strongest at the highest stimulus level, also decreased in differing amounts across CF as stimulus level decreased. The onset latency appeared to change as a function of CF with decreasing stimulus levels such that lower CF units showed more latency prolongation than higher CF units. This decrease in the temporal precision of synchronized responses was depicted in the width of the first peak. In contrast, the spike activity in the second peak response was stongest at a lower stimulus level (55 dB SPL). Neither changes in the onset latency nor in the temporal precision were observed with decreases in stimulus level. In **B**, spike activity for /be/, which was strongest in the absence of background noise, decreased with increasing background noise levels at differing amounts across the CF gradient. Latency and temporal precision of the synchronized spike activity did not appear to be as affected by changes in signal-to-noise level. For /pe/, the first peak response showed similar effects along dimensions of averaged firing rate, onset latency, and temporal precision as for the response to /be/. The second response for /pe/ showed decreasing averaged spike activity of varying amounts across the CF axis with increasing noise level. Responses from high CF regions decreased faster than from low CF regions. The onset latency and temporal precision was less affected by increases in background noise level. In **C**, level-dependency of averaged firing rate for /be/ and /pe/ was similar to that found in the naive animal which was described in **A**. Although the onset of the synchronized response was more temporally precise than in the naive animal, the effect of signal level on peak latency and temporal precision was similar to that found in the naive animal. In **D**, the effect of signal-to-noise level changes for /be/ and /pe/ was similar to that found in the naive animal which was described in **B**.

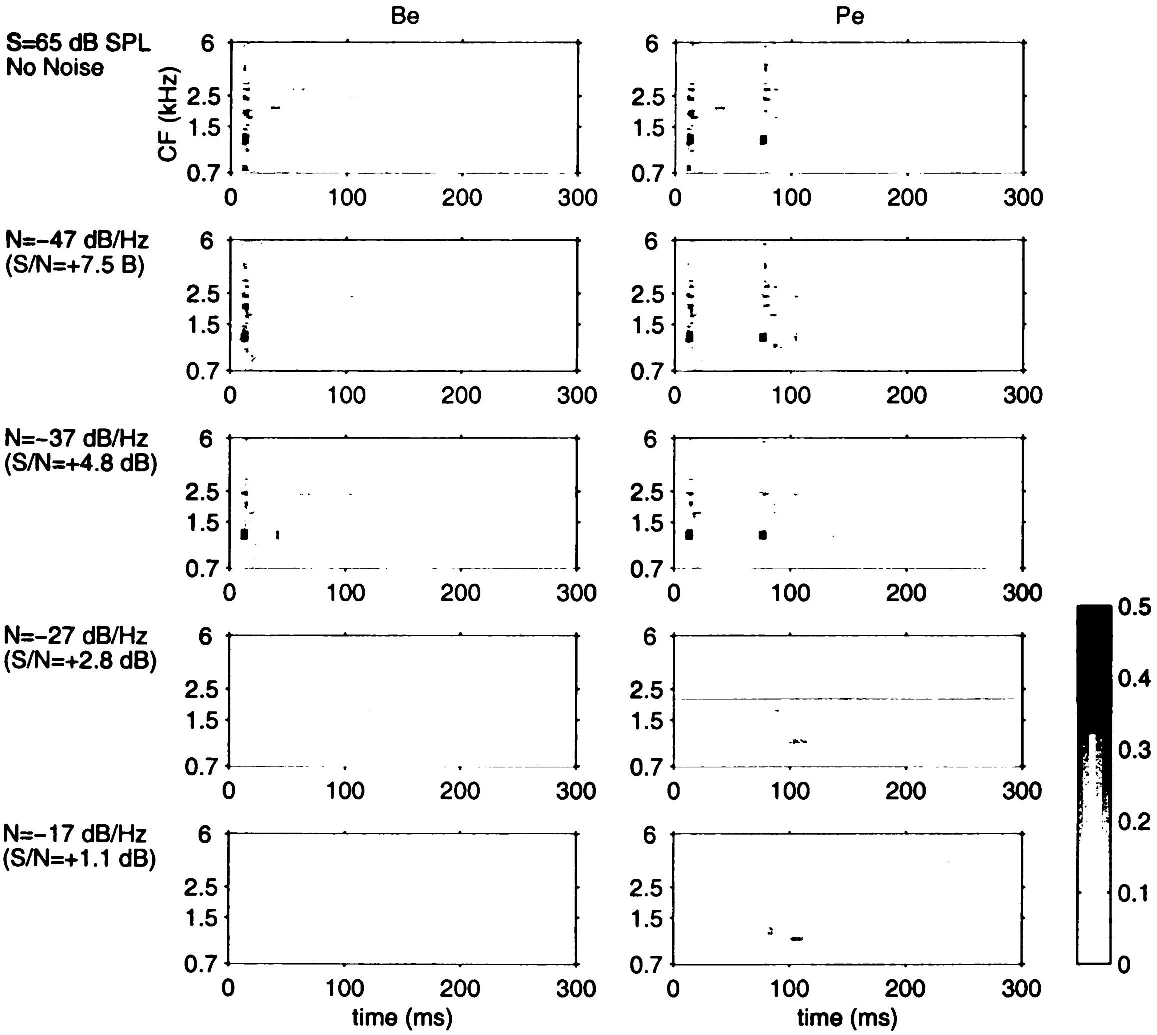

A
Be
Pe
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
CF (kHz)
6
2.5
1.5
0.7
0
100
200
300
time (ms)
0.5
0.4
0.3
0.2
0.1
0


B
Be
Pe
S=65 dB SPL
No Noise
N=−47 dB/Hz
(S/N=+7.5 B)
N=−37 dB/Hz
(S/N=+4.8 dB)
N=−27 dB/Hz
(S/N=+2.8 dB)
N=−17 dB/Hz
(S/N=+1.1 dB)
CF (kHz)
6
2.5
1.5
0.7
0
100
200
300
time (ms)
0.5
0.4
0.3
0.2
0.1
0

C


Be
Pe
65 dB SPL
55 dB SPL
45 dB SPL
35 dB SPL
25 dB SPL
CF (kHz)
6
2.5
1.5
0.7
0
100
200
300
time (ms)
0.5
0.4
0.3
0.2
0.1
0

D


Be
Pe
S=65 dB SPL
No Noise
N=-47 dB/Hz
(S/N=+7.5 B)
N=-37 dB/Hz
(S/N=+4.8 dB)
N=-27 dB/Hz
(S/N=+2.8 dB)
N=-17 dB/Hz
(S/N=+1.1 dB)
CF (kHz)
6
2.5
1.5
0.7
0
100
200
300
time (ms)
0.5
0.4
0.3
0.2
0.1
0

background noise levels. For both figures, the temporal precision of the onset response as depicted by the dark bands for either the first or second peak was most notable as continuous background noise was present during speech presentation and as the noise level increased. The magnitude of either primary or secondary band (in the case for /pe/) eventually decreased with subsequent increases of background noise. At the highest background noise levels, only traces of the secondary band were visible for /pe/. In figure 3.8D, the time width of this secondary band had also broadened with increasing amounts of background noise.

In summary, the neurogram representation of the speech stimuli presented at multiple stimulus levels and multiple levels of continuous background noise clearly revealed a reduced amount of neural information across cortical neurons necessary to represent speech. Any subtle differences among subgroups of cells as distinguished by their response profile to changes in signal-to-noise ratio were not clearly observable in the neurogram.

3.2.6 Cortical maps

Displaying the neural responses to speech topographically across the mapped cortical area was the final method used to represent the collected neuronal responses. By plotting the neural responses along cortical space, one can determine whether or not neural responses to speech are activated in a spatially-distributed manner and whether these activation patterns correspond to any previously identified functional topography. Figures 3.9A-D display cortical maps to speech stimuli in the first four columns (from left to right) and to pure tone stimuli in the fifth column.

Figures 3.9A and 3.9B show 1) the neural activation patterns for /be/ and /pe/ at two time slices corresponding to the peak activity of the temporal response waveform for each stimulation condition and 2) three functional topographies representing threshold, Q10, and onset latency (which were presented previously in Figures 2.5F,

2.9A, and 2.13A, respectively) in a naive animal. For the neural maps for speech stimulation, the regions marked in yellowish-red reflect higher magnitude of neural activity than those characterized in purple, which represents areas of minimal stimulus-evoked activity. For the spectral parameter maps, yellowish-red regions indicate larger values for the corresponding metric while bluish-purple regions indicate smaller values. Three isofrequency contours at 1, 2, and 3 kHz, which were derived by the interpolated CF values from the sampled units (see Figure 2.5F), are represented by the black lines in the colorplots for cortical frequency reference. In figure 3.9A, the first four columns (left to right) of color plots represent the spatial distribution of magnitude for either the first or second peak response to speech presented at different stimulus levels. The far left column of color plots represent neural activity for /be/ and /pe/ recorded at 15 ms and 80 ms for the loudest stimulus presentation level. At 15 ms, the neural activity for /be/ and /pe/, which show a strong resemblance, appeared to show a distinct pattern of activation with a number of highly activated sites interspersed with non-responsive locations. No clear relationship with the transitions and formant frequencies emerged. At 80 ms, there was little neural activity for /be/, although the trace amount of activity that was apparent was along portions of two lower CF lines. Three distinctions were found in the neural activity for /pe/ at 80 ms: 1) neural activity was concentrated in regions were there was little or no neural activity at 15 ms, 2) regions which did respond at 15 ms displayed reduced activity at 80 ms, 3) neural activity along the isofrequency dimension appeared to coincide with the sharply-tuned area as indicated by the Q10 map, and 4) a strong focal activation near the first formant frequency was visible. At a stimulus level 10 dB lower (S=55 dB SPL), reduction in the firing strength and cortical space of neural activity was observed in general for /be/ and /pe/ at 15 ms and /pe/ at 80 ms. Spatial correlation of neural activity at 55 dB SPL and 65 dB SPL for each CV at their respective time slice, excluding /be/ at 80 ms, revealed more than 50% correlation in the spatially-distributed neural activity (/be/ at 15 ms: r=.53;

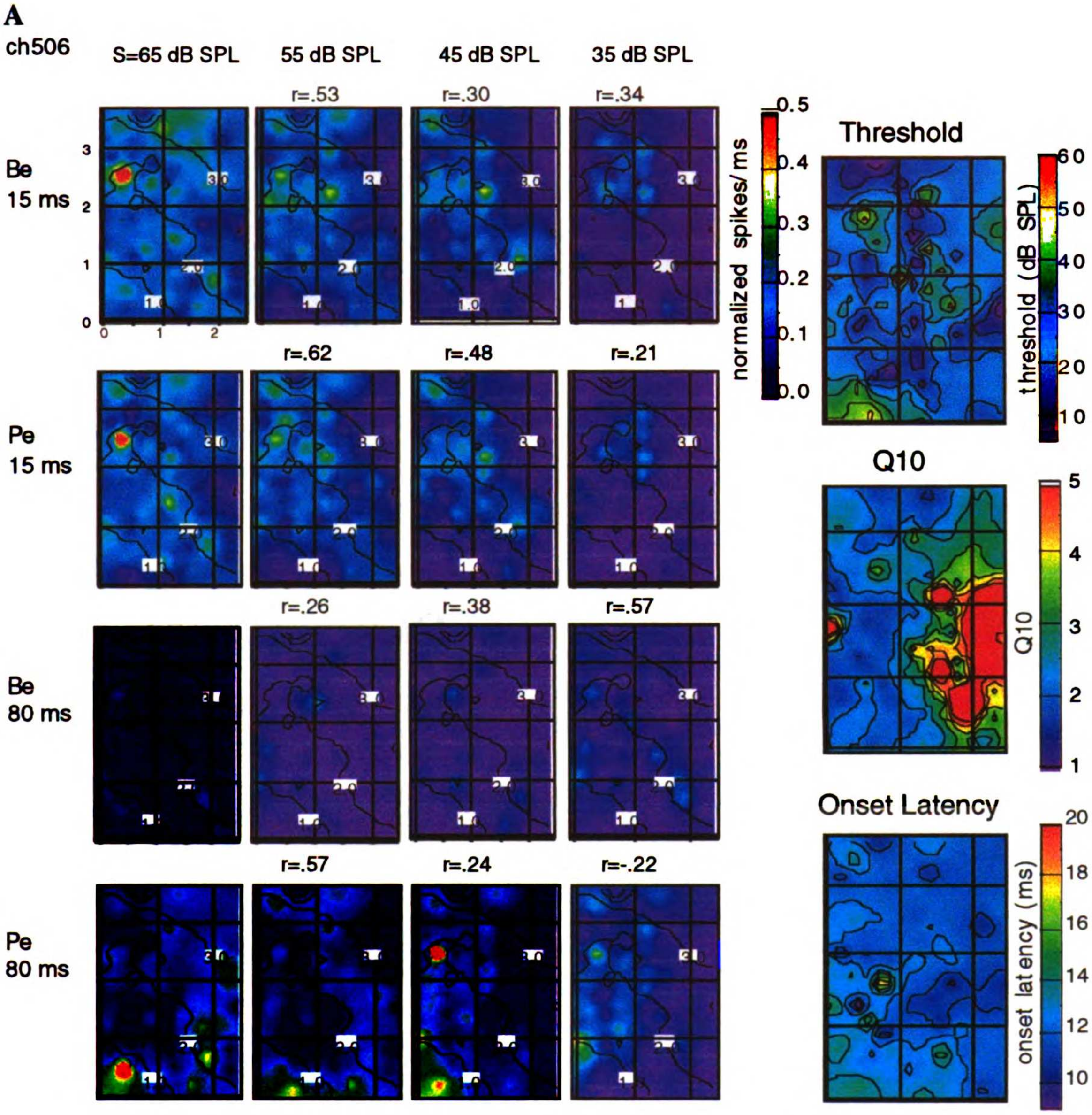


Figure 3.9A Cortical map of neural activity to CVs and functional maps in a naive animal

Spatial distribution of speech-evoked cortical activity for /be/ and /pe/ for four different signal levels at time intervals corresponding to peaks in the PSTH response are represented by temporally-synchronized neural activity (first four columns left to right). Yellow-red regions represent areas of high neural activity while bluish-purple regions low neural activity. The frequency gradient is represented by the black lines in each colorplot. Spectral parameter maps to basic stimuli are depicted in the last column. The degree of change in the spatially-distributed neural activity for the response to /be/ and /pe/ at each signal level with the highest stimulus level are reflected by the correlation coefficient above each cortical map to speech. A high correlation coefficient indicates similarity in spatial response pattern while a low coefficient very little resemblance. Level-dependent spatial-temporal patterns are reflected by the decreasing correlation coefficients with decreasing signal levels.

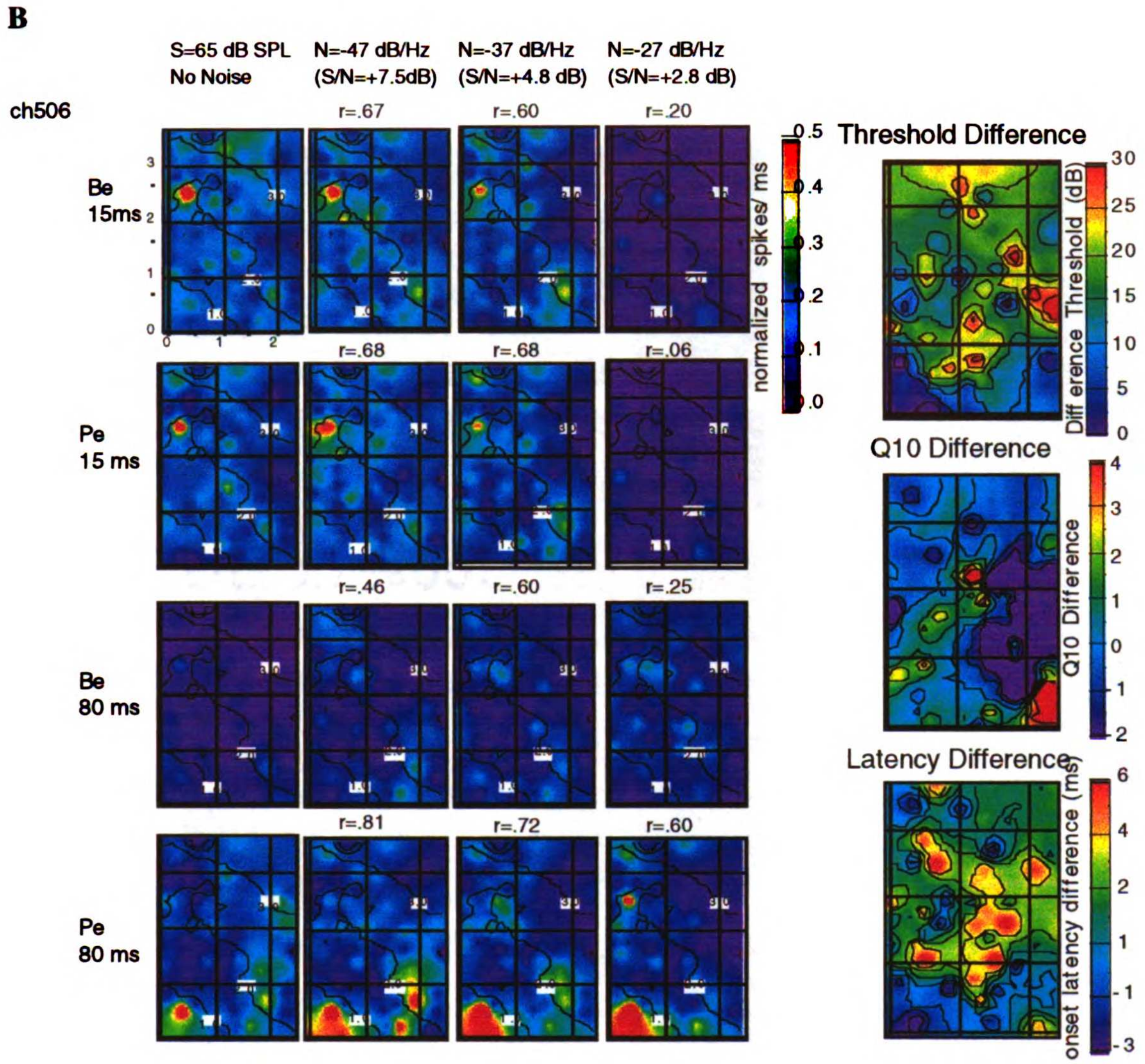


Figure 3.9B Cortical maps of neural activity to CV in background noise and functional difference maps

Spatial distributions of speech-evoked cortical activity at four different signal-to-noise levels are presented for /be/ and /pe/ at two time intervals corresponding to peaks in the temporal response. The color gradient representing the level of neural activity is the same as in figure 3.9A. Three spectral difference maps represented in the far right column depict the difference in response parameter in conditions with and without background noise. Yellow-red indicates an increase in response with noise while bluish-purple a decrease in response. The effect of decreasing the signal-to-noise ratio on spatial-temporal response pattern is depicted by the correlation coefficients which changed less with increases in background noise than with decreases in signal level (seen in figure 3.9A). That is, the synchronized locally-distributed spatial-temporal pattern was robust to changes in the signal-to-noise level.

C

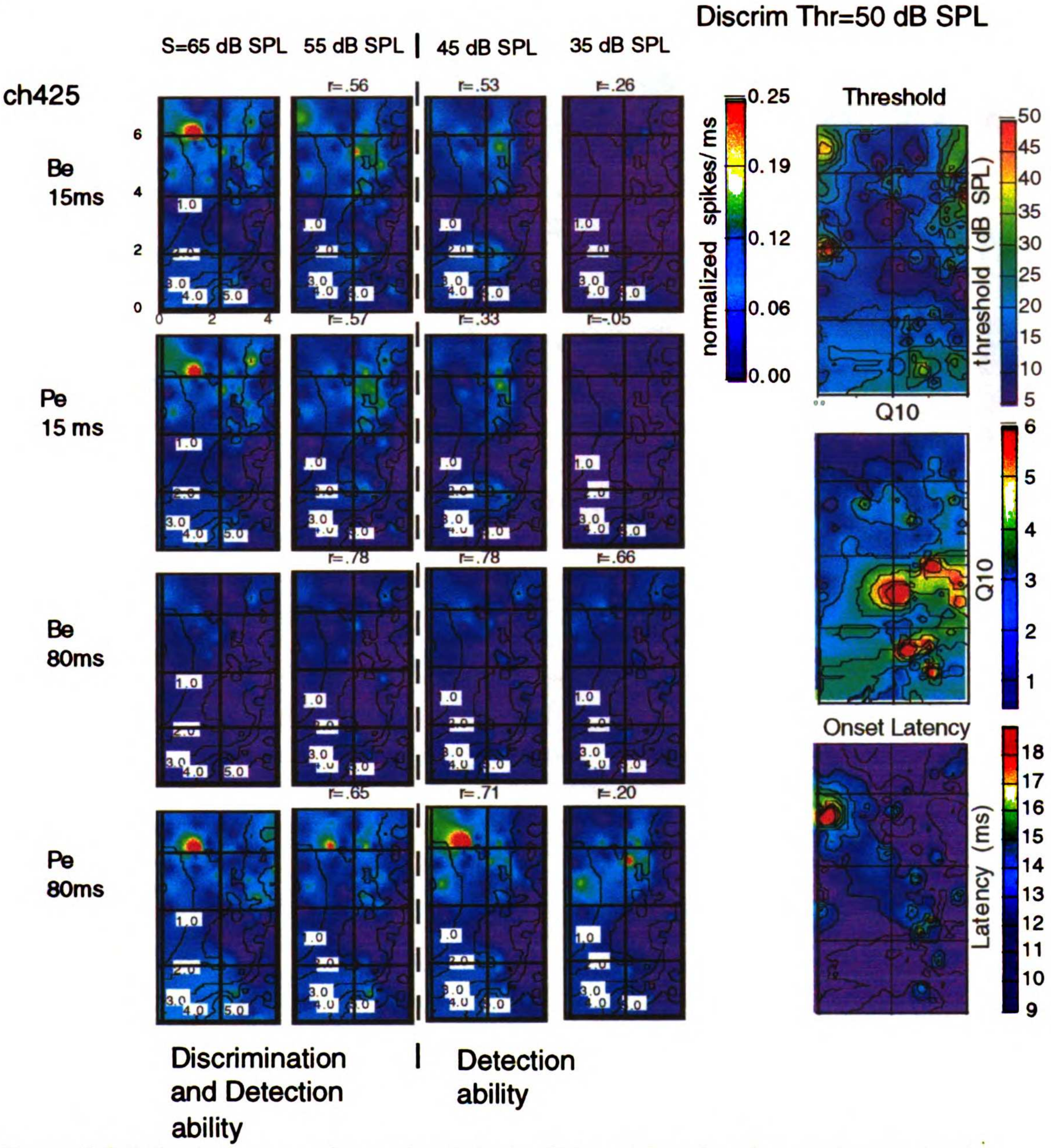


Figure 3.9C Cortical maps of neural activity to CV and functional maps for a trained animal

Spatial distributions of speech-evoked cortical activity at four different signal levels to /pe/ and /be/ at time intervals corresponding to peaks in the temporal response. Three spectral parameter maps are represented in the last column (far right). The color gradient for these plots are similar those described in figure 3.9A. The declining correlation coefficients, particularly at 15 ms, indicate level dependent synchronized, spatial-temporal neural activity. The decline in correlation coefficient at 15 ms coincides with the psychophysical boundary separating discrimination from mere detection abilities. Note the absence of synchronized neural activity in the centrally-located low-threshold, sharply-tuned region of the speech maps. Speech-evoked cortical activity was represented more by the broadly-tuned regions.

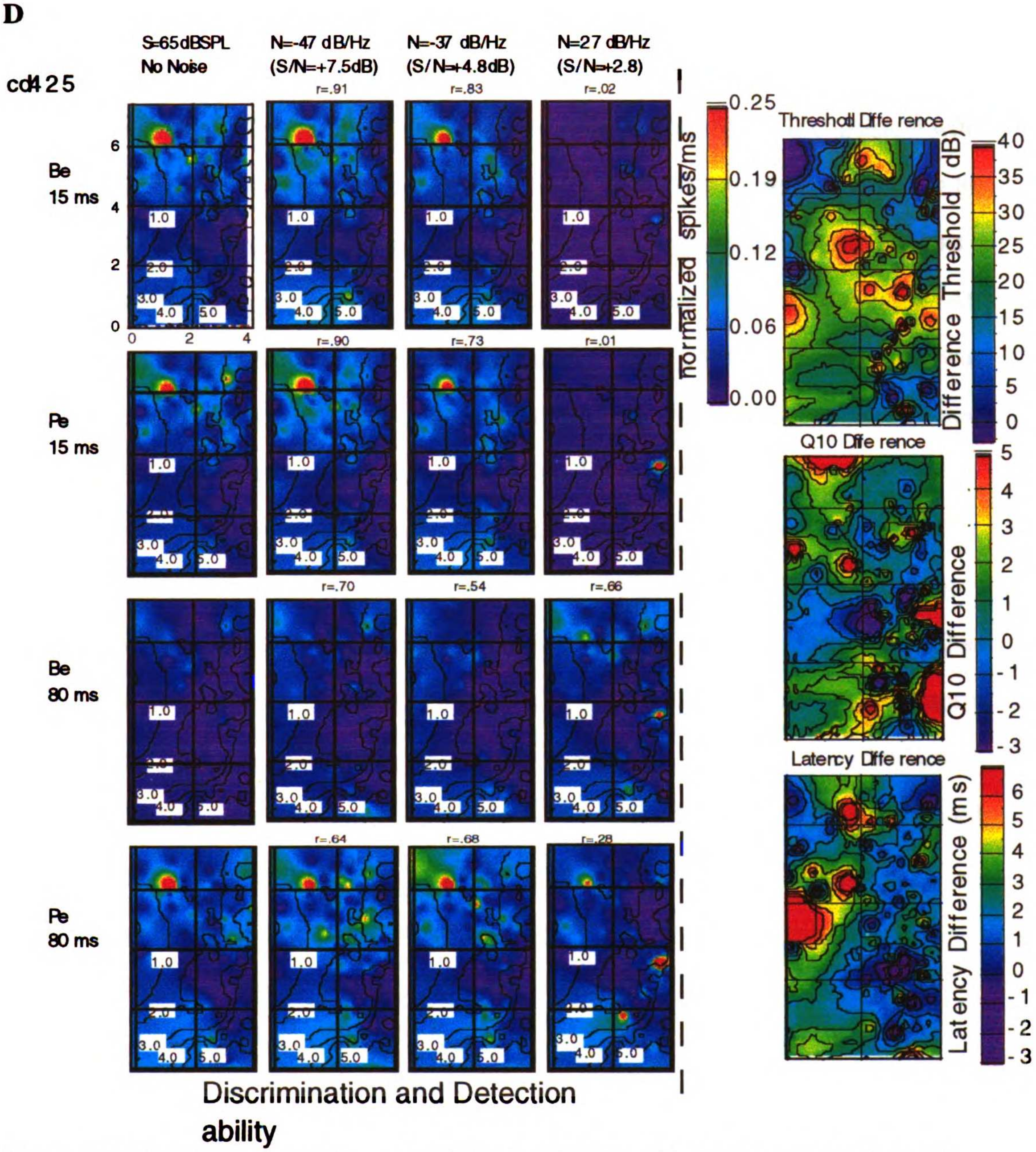


Figure 3.9D Cortical maps of neural activity to CV stimuli in background noise and functional difference maps for a trained animal

Spatial distributions of speech-evoked activity at four different signal-to-noise levels to /pe/ and /be/ at time intervals corresponding to peaks in the temporal response. The spectral difference maps are depicted in the far right column. The color gradients for these plots were similar to those in figure 3.9B. The spatial distributions, as indicated by the correlation coefficients, were robust to changes in signal-to-noise level at both time intervals, except at the lowest S/N level represented. The discrimination and detection ability the cat was represented by neural activity at all four of these S/N levels. Note the absence of synchronized neural activity in the centrally-located low-threshold, sharply-tuned region. Speech-evoked activity was represented more by the broadly-tuned regions.

/pe/ at 15 ms: r=.62; /pe/ at 80 ms: r=.57). This relatively high spatial correlation is indicative of similar spatial pattern of neural activity between this signal level and the highest signal level. The minimal activity for /be/ displayed at 80 ms is consistent with the one peak histogram response seen earlier. For /pe/, the second peak response at 80 ms showed a gradual increase of activity towards more dorsal, broadly-tuned regions of the isofrequency contours and a slight decrease in the central-to-ventral regions. At 45 dB SPL, further reduction in the magnitude and cortical space of neural activity was observed for /be/ and /pe/ at 15 ms. Spatial correlation for both CV syllables at 15 ms between 45 dB SPL and 65 dB SPL was greatly reduced (/be/ at 15 ms: r=.30; /pe/ at 15 ms: r=.48) indicative of less similarity in the spatial distribution pattern. Neural activity at this time frame appeared to be restricted to central-to-dorsal regions within isofrequency contours between 1 and 3 kHz. At 80 ms, the neural response to /pe/ has shifted towards more dorsally-located regions along the frequency axis. This is clearly demonstrated by the increased magnitude of neural activity at the 1 to 2 kHz frequency areas. At the softest stimulus level (S=35 dB SPL), further reductions in the magnitude and size of the cortical response were found for speech at 15 ms and at 80 ms.

In figure 3.9B, the first four columns of color plots represent neural activity for either the first or second peak response to a fixed speech stimulus level at different background noise levels while the last column of three color plots represent the difference values of the functional topographies with (N=-27 dB/Hz, S/N=+2.8 dB) and without continuous background noise for a naive animal (which were presented previously in Figures 2.5F, 2.9A, and 2.13A, respectively). Again, the isofrequency contours were represented by the black lines within the color plots as reference to the frequency gradient. The color gradient representing the speech-driven neural activity is similar to that found in figure 3.9A (i.e., yellowish-red indicates high neural activity while bluish-purple indicates low neural activity). The color gradient representing the spectral difference maps depict yellowish-red regions as increasing in metric value in

the presence of continuous background noise and bluish-purple regions as decreasing. The far left column of activity patterns depict the spatial response to the speech stimulus at 65 dB SPL in the absence of continuous background noise, which were presented also in the far left column of figure 3.9A. The next column to the right represents neural activation patterns for speech stimuli presented in the presence of low continuous background noise (i.e., N=-47 dB/Hz, S/N=+7.5 dB). Again, as seen in the previous representations of the data, the neural response for the speech stimuli at either time interval increases with low levels of noise, but only at restricted regions in cortical space. The locations of increased neural activity appeared to stretch across the frequency axis at three differently spaced points (lower right corner, center, and upper left corner of each plot in this column) along the isofrequency domain for the two time intervals displayed. These locations of increased neural activity coincide with places where Q10 changes in the presence of background noise. A cortical distance of approximately 1.5 mm appeared to separate the bands of increased neural activity which were found to be orthogonal to the frequency axis. The spatial correlation between neural activity at this noise level and in the absence of background noise revealed high levels of correlation indicating only a small difference in the neural pattern of activity (/be/ at 15 ms: r=.67; /pe/ at 15 ms: r=.68; /pe/ at 80 ms: r=.81). Although there was minimal spike activity in the case of /be/ at 80 ms, the spatial correlation (r=.47) indicated that the neural location in the sublevels of firing were still quite similar across the two stimulus levels. When the background noise level was increased by another 10 dB (i.e., N=-37 dB/Hz, S/N=+4.8 dB), a global decrease (with respect to the previous noise level condition) in spike activity to the speech stimuli was observed; however, the band of cells in the lower right corner of the two plots representing activity at 15 ms showed further increases in spike activity. Despite differential changes in the magnitude of spike activity, the pattern of synchronized activity was robust to the increase in background noise level as was indicated by the high spatial correlation values (/be/ at 15 ms: r=.60; /pe/ at 15 ms: r=.68). On the other hand, neural activity from these locations recorded at 80 ms showed decreased activity (with respect to the previous noise level condition); yet, the magnitude of activity was still higher than the condition where noise was absent. The high spatial correlation between this background noise level and the absence of noise, however, indicated that the cortical locations of spike activity were relatively the same between the two signal-to-noise levels (/be/ at 80 ms: r=.60; /pe/ at 80: r=.72). In addition, at 80 ms for the /pe/ stimulation condition, the neural activity in the lowest CF region (below 1 kHz) was noticeably higher. In the last noise level condition (N=-27 dB/Hz, S/N=+2.8 dB), neural activity for either CV stimulus at 15 ms was virtually eliminated. Hence, the spatial correlations were low (/be/ at 15 ms: r=.20; /pe/ at 15 ms: r=.06). In contrast, at 80 ms, neural activity was enhanced at spatially-segregated regions for /be/ and for selected regions for /pe/ compared to the condition without background noise. Areas where neural activity was enhanced for /pe/ included the CF region below 1 kHz and a banded region in the upper left corner.

Although direct comparison of spatially-distributed activity across test animals can be confounded by neural sampling and anatomical variance in cortices of the animals, one can examine the effect of speech training on spatially-distributed spike activity to speech sounds as a function of speech stimulus or background noise level. Figures 3.9C and 3.9D display the neural activity for the presentation of /be/ and /pe/ at four different speech stimulus levels (for figure 3.9C) and background noise levels (for figure 3.9D) and corresponding functional topographies for threshold, Q10, and onset latency for a trained cat. As in the other example case, the CF map is superimposed on the cortical map of neural activity to speech sound stimulation for reference to the cortical frequency organization of the animal. In addition, the vertical dotted line separating the speech topography columns represents the psychophysical boundary for speech discrimination ability as demonstrated by this trained animal. In figure 3.9C, the four color plots in the far left column represent neural activity for /be/ and /pe/ at the loudest stimulus presentation level recorded at 15 ms and 80 ms post-stimulus. Again, the spatial pattern of neural activity appeared to be similar for presentation of either speech stimulus at 15 ms. The pattern was distinguished by a central band of inactivity across the frequency axis which corresponded to the low-threshold, sharply-tuned region. At 80 ms, the neural activation pattern for /pe/ was similar to that found at 15 ms although there was a slight increase of activity in the ventral region along the isofrequency lines and more dorsally along the 3 kHz contour. As expected, there was minimal activity at 80 ms for /be/. At a stimulus level 10 dB lower, as shown in the next column to the right, the concentration of high activity in the spatial pattern for either /be/ or /pe/ at 15 ms shifted towards higher frequencies. At 80 ms, both the area of neural activity and magnitude of activity for /pe/ decreased while there appeared to be minimal amount of change for /be/. A further 10 dB decrease in stimulus level (S=45 dB SPL) reduced the area of neural activity to /be/ and /pe/ at 15 ms; hence, the spatial correlation decrease especially in the case of /pe/ (/be/ at 15 ms: r=.53; /pe/ at 15 ms: r=.33). However, at 80 ms for /pe/, despite the small reduction in the area of neural activity (r=.71), the strength in the response slightly above the 1 kHz isofrequency contour in the dorsal, broadly-tuned region increased. At the lowest stimulus level, neural activity was obliterated at 15 ms for both /be/ (r=.26) and /pe/ (r=-.05) whereas at 80 ms the response to /pe/ was still observable although the concentration of activity within the spatial pattern changed again (r=.20). Again, the lack of activity at 80 ms for /be/ stimulus presentation at any level corroborates with previous histogram plots displaying only one response peak to the voiced CV. The overall trend in the spatial activation of neural firing shown in this example, lends further support for level-dependency of spatially-distributed neural firing to speech stimuli.

In figure 3.9D, all four signal-to-noise levels that are represented correspond to levels at which the trained animal could both detect and discriminate speech syllables.

The last column of color plots represents the difference values for the functional topography in the presence of a background noise (N=-27 dB/Hz). The four speech plots in the far left column represent neural activity for /be/ and /pe/ stimulation in the absence of noise recorded at 15 ms and 80 ms post-stimulus. Again, these plots were depicted earlier in the far left column of figure 3.9C. In the column of plots to the right, the addition of low levels of background noise (N=-47 dB/Hz, S/N=+7.5 dB) increased the neural activity in both the dorsal and ventral regions that had originally displayed activity in the absence of noise at 15 ms; however, the spatial correlation indicated that the synchronous pattern of spatially-distributed neural activity to speech was relatively the same for this signal-to-noise level and the condition without background noise (/be/ at 15 ms: r=.91; /pe/ at 15 ms: r=.90). At 80 ms, the neural response expanded, by cortical dimensions, resulting in a spatial correlation that was lower than those found for the first peak response (/pe/ at 80 ms: r=.64). At the next noise level (N=-37 dB/Hz, S/N=+4.8 dB), the response decreased in the first 15 ms for either stimuli; however, the spatial pattern changed very little as expressed by a relatively high spatial correlation coefficient (/be/ at 15 ms: r=.83; /pe/ at 15 ms: r=.73). At 80 ms, the activity level changed differentially with respect to cortical location. For /pe/, neural activity was found to increase in the dorsal regions below 2 kHz while it decreased in other regions. That spatial correlation for the secondary peak response was relatively unchanged with the 10 dB increase in background noise (r=.68). Even for /be/, for which there was no recognizable second peak at 80 ms, neural activity increased and expanded in cortical space in the lower frequency regions. For the last noise level, neural activity was diminished greatly at 15 ms, except for a small location within the center of the 3 kHz isofrequency line; hence the low spatial correlations (/be/ at 15 ms: r=.02; /pe/ at 15 ms: r=.01). At 80 ms, further expansion of neural activity to /be/ was observed in the dorsal low-frequency areas while differential changes in neural activity to /pe/ were observed in the central band across the frequency axis and dorsal regions.

In summary, synchronous, spatially-distributed neural activity to speech syllables was found to be level dependent and more robust to changes in signal-to-noise levels. The effects of level and background noise level changes on the neural activation patterns in the primary auditory cortex appeared similar between naive and trained animals. At this level of analysis, the only obvious distinctions in the synchronous neural activity that could be related to psychophysical detection and discrimination abilities were the reduced response magnitudes and the changes in the cortical excitation patterns elicited by the CV syllables. The relationship between the sharpness of tuning corresponding to the areas of synchronous neural activity for the different signal levels, however, did not appear to be consistent between the trained and untrained animals. In the naive cat, synchronous neural activity was prevalent across the isofrequency contours without preference given to bandwidth tuning characteristics of the neurons. In contrast, for the trained cat, synchronous neural activity was most prevalent in the broadly-tuned neurons and was notably absent in the sharply-tuned neurons, except in the lowest signal-to-noise condition (S/N=+2.8). Further examination of the relationship between sharpness of tuning and spike rate activity to speech stimulation will be covered in the next section.

3.2.7 Statistical summaries of the cortical response to speech

A statistical review of the cortical activity to the speech syllables /be/ and /pe/ was completed to gain insight to the relationships among factors including 1) sharpness of tuning, 2) behavioral training, 3) temporal response peaks (i.e., first versus second peak response), 4) signal levels, 5) signal-to-noise levels. Either factor could account for the variances found in the neural responses as quantified by maximum firing rate, its latency, and response peak width (measured at 50% of the peak response). As an overview of the statistical trends, a factorial analysis of variance was determined from the mean values calculated for each animal.

3.2.8 Maximum spike rate

Due to the large amount of statistical data obtained, only those which were of any statistical significance are reported. Figure 3.10A-E display the significant statistical results from the ANOVA-factorial statistic testing the null hypothesis that there are no differences in maximum spike rate attributable to the factors mentioned in the above section for data collected when signal level alone was altered. As a confirmation of the earlier observations, figure 3.10A reveals that voiced and unvoiced CV syllables, /be/ and /pe/, differed in maximum firing rate only in the secondary peak response (labeled in the graph as 2) and not in the first peak response (labeled as 1) (two-factor ANOVA; p<.0001). Furthermore, as manifest in figure 3.10B, level dependency of maximum firing was found to be statistically significant for the first peak response and not for the secondary peak response (two-factor ANOVA; p<.0001). To test variances attributable to training, the animals were divided into three groups based upon the degree of training: 1) naïve animals (N=3), 2) trained animals (trained on two different speech discrimination tasks including the perceptually more challenging task of /pe/ vs. /be/ distinction, N=2), 3) trained2 animals (trained only on the one easier speech discrimination task of /be/ vs. /ko/, N=2). As seen in figure 3.10C, both the naïve and trained2 group of animals showed statistical difference in maximum firing rate between the first and second peak response whereas the trained group showed no difference (two-factor ANOVA; p<.0001). A second look at level dependency, as seen in figure 3.10D, showed that signal level does influence firing rate at both first and second peak response times in the naïve and trained groups, but not the trained2 group (three-factor ANOVA; p<.0001). The trained2 group exhibited level dependency in spike firing for the first response, but not for the secondary response. With regards to the effect of

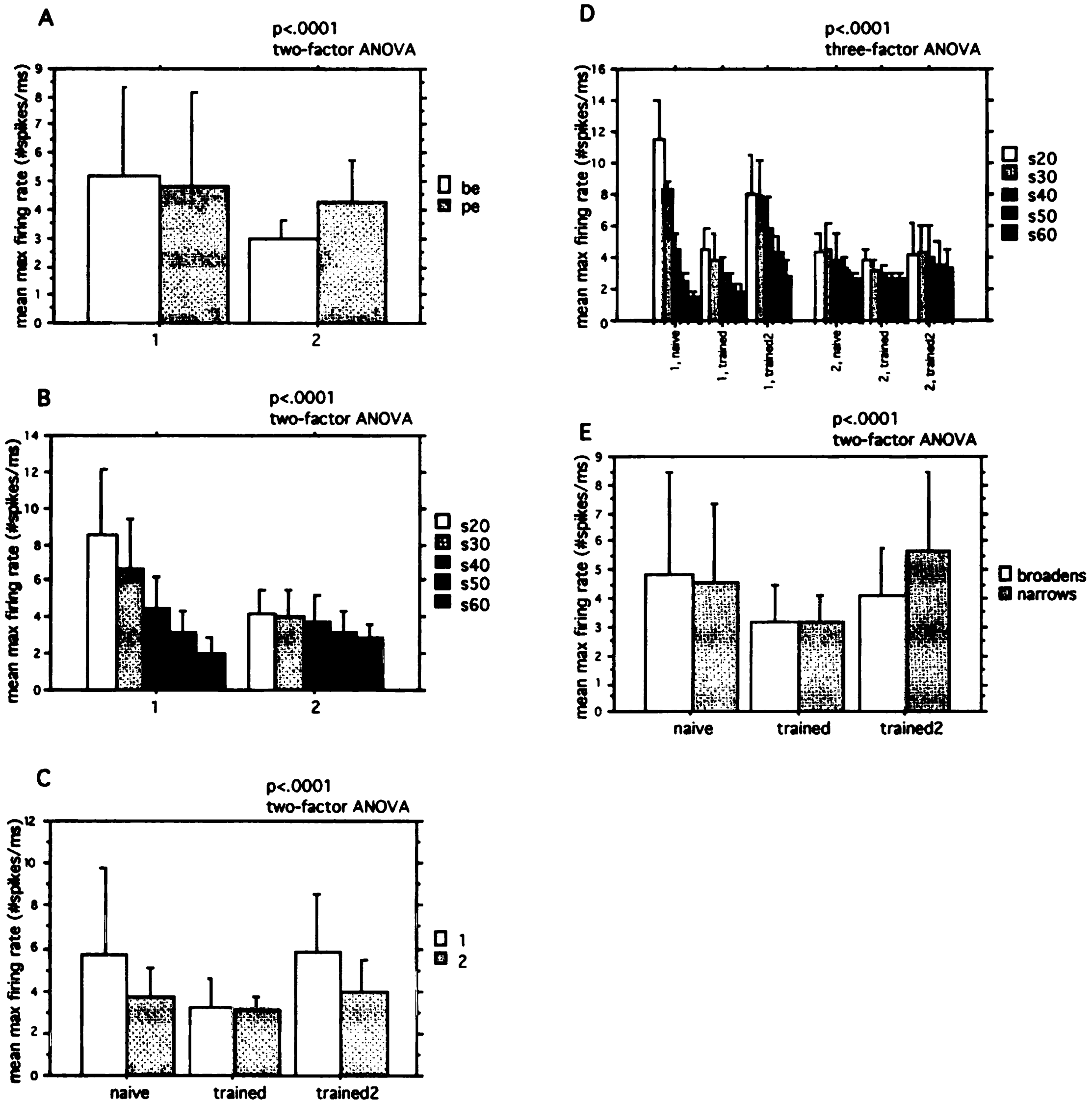


Figure 3.10 Summary statistics for mean maximum firing rate across all animals (N=7)

In **A**, mean firing rates for the first (1) and second (2) peak responses averaged across all stimulus levels were plotted for /be/ and /pe/. Unlike the first response, firing rate of the second response for /be/ was lower than for /pe/. In **B**, mean firing rates for the first and second responses averaged across both CV stimuli were plotted for each level (in dB SPL: s20=65; s30=55; s40=45; s50=35; s60=25). As signal level was lowered, the mean difference between first and second response decreased due to a decrease in the first peak mean firing rate. In **C**, mean firing rates for three groups of cats (naive: no training; trained: two discrimination tasks; trained2: one discrimination task) averaged across CVs at all stimulus levels were plotted for first and second peak responses. Unlike naive and trained2, the trained group showed a reduced first response that was larger than the second only at higher signal levels, as seen in **D**. In **E**, mean firing rates were plotted for subgroups of neurons categorized by their sharpness of tuning response (broadens: Q>2 whose tuning broadens with background noise; narrows: Q<2 whose tuning sharpens with background noise). Unlike in the naive and trained groups, broadly-tuned neurons in the trained2 group showed a higher mean firing rate than sharply-tuned neurons.

sharpness of tuning response characteristics on spike activity to speech sounds, figure 3.10E showed that there was no difference in firing rate between broadly- and sharply-tuned neurons for the naïve and trained groups (two-factor ANOVA; p<.0001). In contrast, broadly-tuned units responded with more spike activity than sharply-tuned units in the trained2 group.

Next, figure 3.11A-E display the significant statistical results from the ANOVA-factorial statistic testing the null hypothesis that there are no differences in maximum spike rate attributable to the factors mentioned in the above section for data collected when background noise level alone was altered, thus changing the signal-to-noise ratio. Again, as seen in figure 3.11A, maximum firing rate for the speech syllables /be/ and /pe/ differed only in the second peak response (two-factor ANOVA; p<.0001). When the first and second response were split by signal-to-noise levels as displayed in figure 3.11B, firing rate was higher in the secondary response at low signal-to-noise levels and higher in the first response at higher signal-to-noise levels (two-factor ANOVA; p<.0001). When looking at the effects of training, figure 3.11C revealed a significant difference in firing between the first and second response for the naïve group and none for either trained groups (two-factor ANOVA; p<.0001). However, taking signal-to-noise level effects into account, which is shown in figure 3.11D, it appeared that no differences in firing rate between the first and second response for the naïve group was apparent only at high signal-to-noise levels. At low signal-to-noise levels for the naïve group, spike activity was greater in the secondary response than for the first. For the trained2 group, no difference in the firing rate was found only for moderate signal-to-noise levels (n30, S/N=+4.8 dB). At higher signal-to-noise levels, firing was greater in the first response while at lower signal-to-noise levels, firing was greater in the secondary response (three-factor ANOVA; p=.0002). With regards to sharpness of tuning effects on driven activity to speech sounds, there was minimal difference in firing rate between broadly- and sharply-tuned units in either the naïve or trained

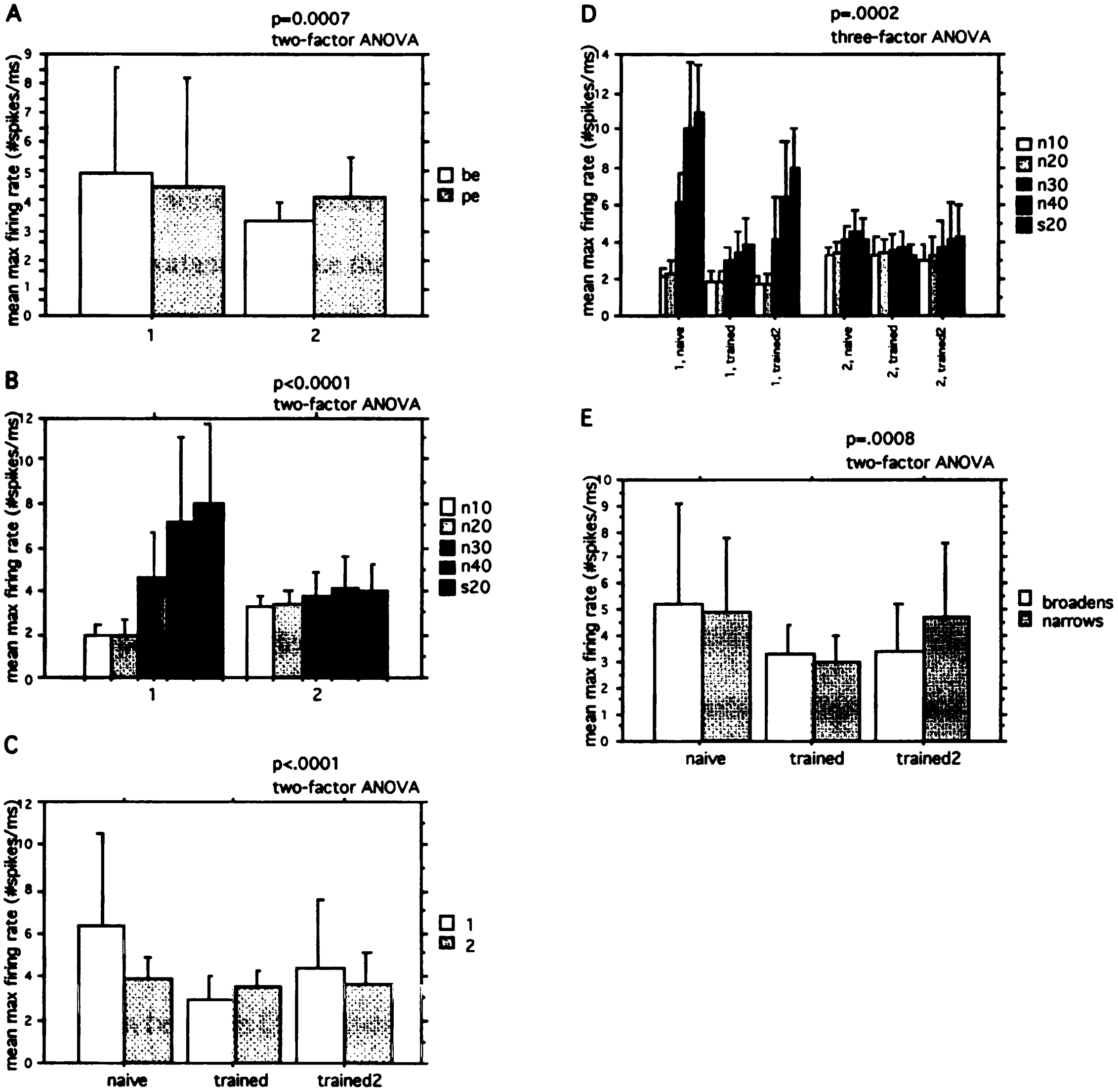


Figure 3.11 Summary of mean maximum firing rate across all cats (N=7) in noise

In **A**, mean firing rates for the first (1) and second (2) peak responses averaged across all signal-to-noise levels were plotted for /be/ and /pe/. Unlike the first response, firing rate of the second response for /be/ was lower than for /pe/. In **B**, mean firing rates for the first and second responses averaged across both CV stimuli were plotted for each level (in dB/Hz: n10=-17; n20=-27; n30=-37; n40=-47; s20=no noise). As signal-to-noise level increased, firing rate of the first response increased more than the second. In **C**, mean firing rates for three groups of cats (naive: no training; trained: two discrimination tasks; trained2: one discrimination task) averaged across CVs at all stimulus levels were plotted for first and second peak responses. Mean firing rate of the first peak response was lower in the trained group than either the naive or trained 2 groups, particularly in the higher signal-to-noise levels, as seen in **D**. In **E**, mean firing rates were plotted for subgroups of neurons categorized by their sharpness of tuning response (broadens: Q>2 whose tuning broadens with background noise; narrows: Q<2 whose tuning sharpens with background noise). Unlike in the naive and trained groups, broadly-tuned neurons in the trained2 group showed a higher mean firing rate than sharply-tuned neurons.

groups which was illustrated in figure 3.11E. For the trained2 group, broadly-tuned units exhibited more spike activity than sharply-tuned units. Sharply-tuned units in both trained groups showed an overall decrease in driven response in the presence of background noise (two-factor ANOVA; p=.0008).

It is important to notice that the overall changes in firing rate caused by level variations and by addition of noise were measurably similar. This is an important prerequisite to be able to compare spatial attributes of level effects and noise effects.

3.2.9 Peak latency

As a result of factorial analysis of variance testing, the mean first peak latency was found to be weakly affected by signal level (factorial ANOVA; p=.04) and signal-in-noise level (factorial ANOVA; p=.002). The mean difference values for first peak latency are displayed in table 3.1. Peak latency differences could also be attributed to the following factors that are displayed in figure 3.12A-D. In figure 3.12A, the only CV response combination that appeared to be affected by a change in level was the second response for /pe/ (factorial ANOVA; p=.0002). There was no significant interaction between CV syllables and signal levels (Sum of Squares (SS): 54.002; Mean Square (MS): 46:888; F-val: 2.353)/signal-to-noise levels (SS: 44.344; MS: 11.086; F-val: 1.353) for the first response nor was there any significant difference in mean second peak latency attributable to changes in signal-to-noise level (SS: 42946.982; MS: 10736.746; F-val: .540). Figure 3.12B illustrated that in all three animal groups the mean first peak latency to /be/ was higher than to /pe/ (one-factor ANOVA; p<.0001), but that there was no significant difference in the mean second peak latency between the CV syllables. That the variance in the values for the second response latency between the two CV syllables was similar is remarkable, especially since there was little observable second peak response for /be/. The variability in response latency for sublevels of neural activity to /be/ must be attributed to a release

Table 3.1 Mean differences for first peak latency

	Mean Diff.	Crit. Diff.	P-Value
s20,s30	-.560	1.328	.4033
s20, s40	-1.161	1.276	.0740
s20, s50	-2.239	1.328	.0012 S
s20, s60	-.336	1.276	.6016
s30, s40	-.601	1.328	.3704
s30, s50	-1.679	1.378	.0177 S
s30, s60	.224	1,328	.7372
s40, s50	-1.079	1.328	.1098
s50, s60	1.904	1.328	.0056 S
n10, n20	-1.361	1.522	.0791
n10, n30	-.304	1.522	.6925
n10, n40	.546	1.522	.4771
n10, s20	2.107	1.522	.0073 S
n20, n30	1.057	1.522	.1708
n20, n40	1.907	1.522	.0147 S
n20, s20	3.468	1.522	<.0001 S
n30, n40	.850	1.522	.2698
n30, n20	2.411	1.522	.0023 S
n40, s20	1.561	1.522	.0446 S

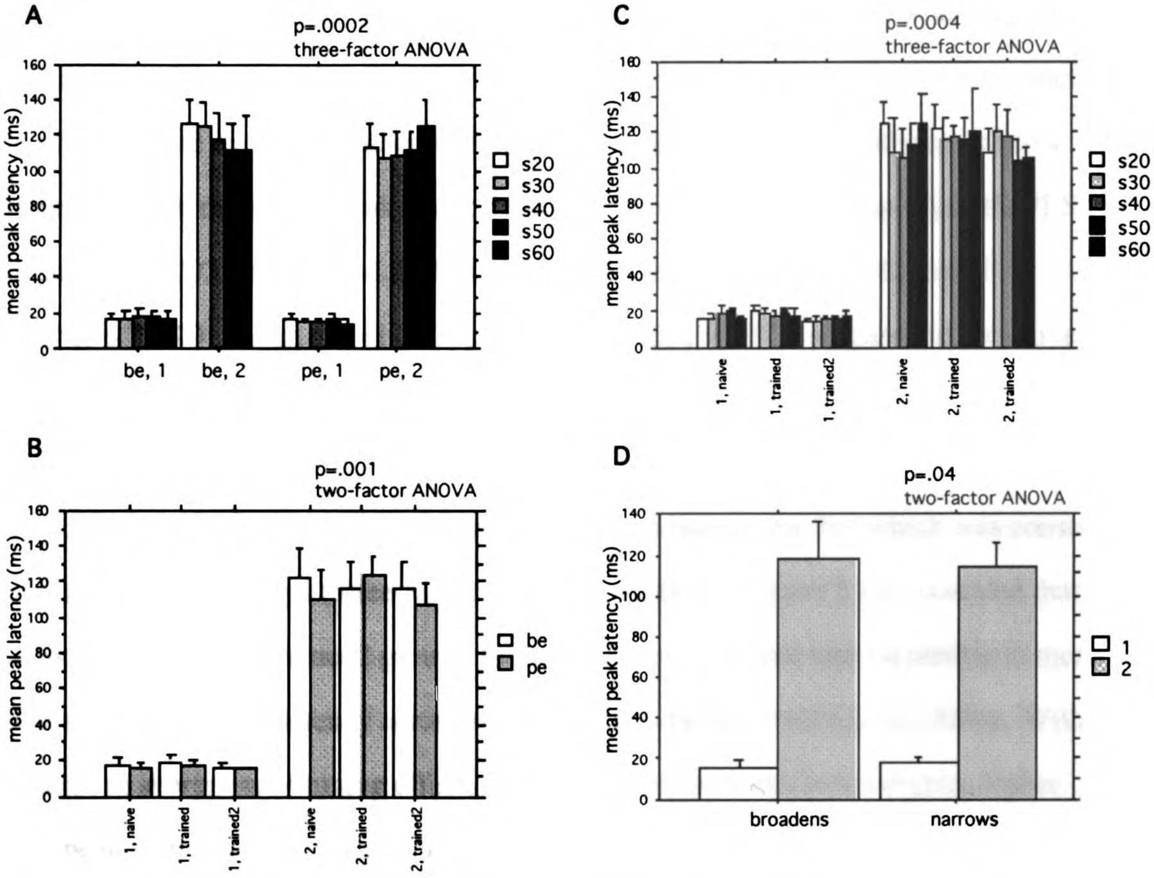


Figure 3.12 Summary statistics of mean peak latency across all animals (N=7)

In **A**, mean peak latencies for /be/ and /pe/ split by first and second responses were plotted for eaach signal level (in dB SPL: s20=65; s30=55; s40=45; s50=35; s60=25). Level differences in first and second peak responses for either CV stimuli were observed although there were very little differences between CVs. In **B**, mean peak latencies for first and second responses split by three cat groups (naive: no training; trained: two discrimination tasks; trained2: one discrimination task) were plotted for /be/ and /pe/. Mean latency of the second peak response to /pe/ was higher in the trained group than either the naive or trained2 groups. Furthermore, the trained group revealed fewer level-attributable differences in second peak latencies than either the naive or trained2 groups, as was seen in **C**. In **D**, mean peak latencies for units characterized by their sharpness of tuning response (broadens: Q10>2 whose bandwidth broadens with background noise; narrows: Q10<2 whose bandwidth sharpens with background noise) were plotted for first and second responses. There was very little difference in latency between the two groups.

from suppression of neural activity following the first peak response ; whereas, the occurrence of the second peak response with the onset of voicing to /pe/ must coincide in time with the excitatory rebound from the first peak response. With regards to a training effect, mean first peak latency was longer in the trained group (Fisher's PLSD: mean diff=1.703; crit. diff.=1.163; p=.005) and shorter in the trained2 group (Fisher's PLSD: mean diff=1.802; crit. diff.=1.163; p=.003) compared to the naive group (one-factor ANOVA; p<.0001). In addition, the second response to /pe/ obtained in the trained group was not only longer than either of the other two groups, but it was also longer than the random mean second latency for /be/ which was constant across all animal groups (three-factor ANOVA; p=.001). Figure 3.12C revealed that unlike the naive and trained2 groups, the trained group showed less variability in mean first and second peak latency across signal levels (3-factor ANOVA; p=.0004). With regards to sharpness of tuning effects, broadly-tuned units exhibited slightly higher first peak latencies on average than sharply-tuned units (one-factor ANOVA; p=.006), but revealed no significant difference in the mean second peak latency (3-factor ANOVA; p=.04).

In summary, the effect of signal level and signal-to-noise level changes on peak latency was small. Shifts up to 3 ms only were found between the loudest stimulus level and the lowest signal-to-noise level. Differences in peak latency were found among animals grouped by their level of auditory experience with the speech stimuli. Lastly, subtle differences in first peak latency were found between neuronal groups based on their sharpness of tuning.

3.2.10 Peak width

Temporal precision of the spike response measured for both the first and second response was represented by peak width. Figure 3.13A showed that with changes in signal level peak width was more restrictive for the first peak than the second and that

/pe/ showed better temporal precision than /be/ for the second response (two-factor ANOVA; p=.004), which lends further support that the variability found in the second response to /be/ is attributable to individual rebounds rather than a stimulus-evoked response . With regards to training effects, figure 3.13B showed that the trained2 group demonstrated more restrictive peak width for the first response than either the naive or trained groups (three-factor ANOVA; p<.0001).

When background noise level was altered, either responses became broader with lower signal-to-noise levels, with the exception of the second response to /be/ where there was sublevels of driven response and thus showed peak width values much higher than those for /pe/ (three-factor ANOVA; p=.02) as displayed in figure 3.14A. With regards to training effects, as noise level changed, peak widths were broader overall in the trained group than either the naive or trained2 groups (three-factor ANOVA; p=.02) as displayed in figure 3.14B. The relationship between sharpness of tuning and peak width is depicted in figure 3.14C where there was a small distinction between the peak widths of broadly- and sharply-tuned units (two-factor ANOVA; p=.004).

To summarize, changes in signal level did little to effect the temporal precision of the peak responses, while changes in signal-to-noise level revealed a broadening in the response width. Behavioral training, or exposure to the speech sound differences, resulted in slight changes in temporal precision compared to the naive animal. Lastly, subtle differences in temporal precision were found between two subpopulations of neurons classified by their sharpness of tuning characteristic.

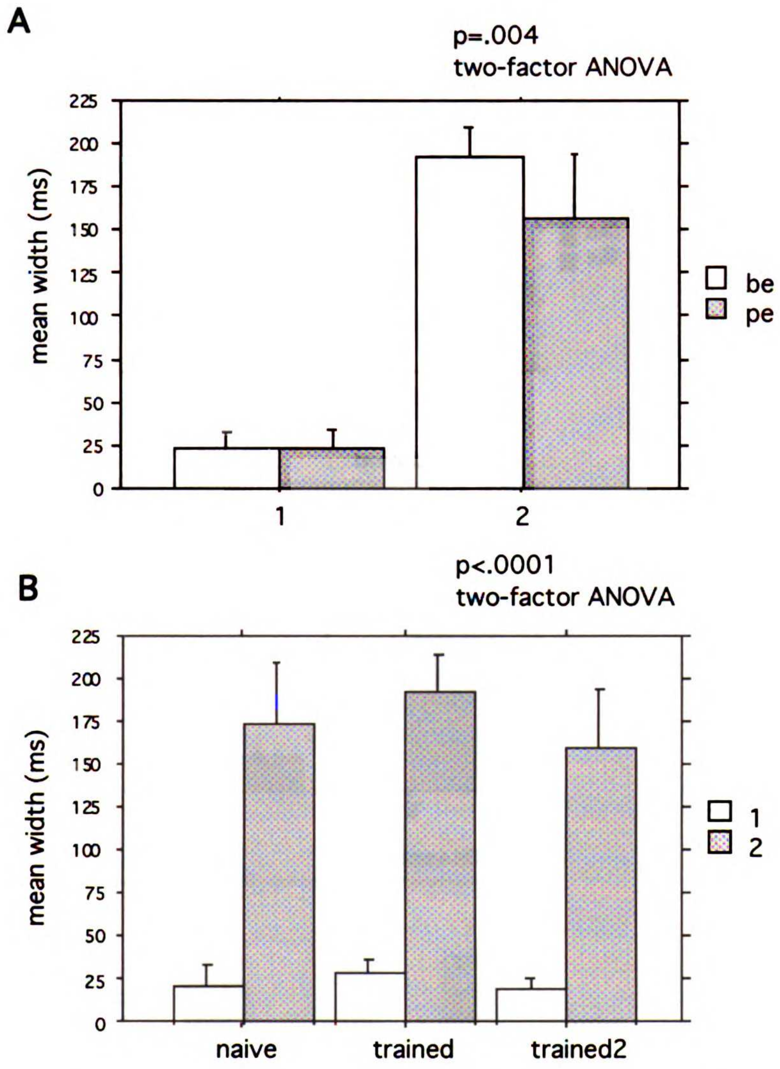


Figure 3.13 Summary statistics for mean width across all animals (N=7)

In **A**, mean widths for the first and second response averaged across all signal levels were plotted for /be/ and /pe/. The mean width of the second response was broader than the first. In **B**, mean widths for three groups of cats (naive: no training; trained: two discrimination tasks; trained2: one discrimination task) were plotted for the first and second response. Temporal precision of the first response was sharper in the trained2 group than the trained.

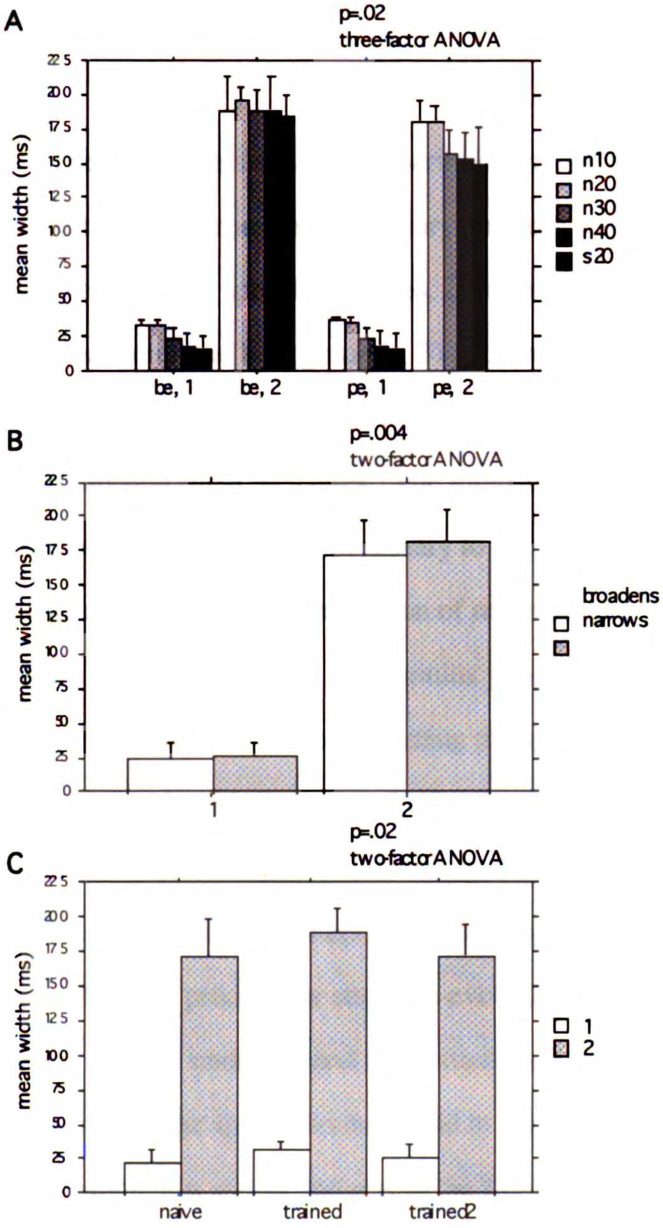


Figure 3.14 Summary statistics for mean width across all animals (N=7) in noise

In A, the mean widths of /be/ and /pe/ split by first and second response were plotted for each signal-to-noise level (in dB/Hz: n10=-17; n20=-27; n30=-37; n40=-47; s20=no noise). Mean widths were not only larger in the second response, but they were also decreased with increasing signal-to-noise levels for either CVs. In **B**, mean widths from three groups of cats (naive: no training; trained: two discrimination tasks; trained2: one discrimination task) were plotted for first and second responses. There was very little difference in mean width for either response across the groups of cats. In **C**, mean widths of the first and second response were plotted for units differing in their sharpness of tuning (broadens: Q10>2 whose bandwidth broadens with background noise; narrows: Q10<2 whose bandwidth sharpens with background noise). There was very little difference in mean width between the two types of units for either response.

3.3.0 Discussion

The goal in this series of experiments was three-fold: 1) To use receptive field properties of neurons in the primary auditory cortex as a basis for studying the cortical representation of speech syllables, 2) to test the robustness of the representation in various contextual backgrounds, 3) to compare the effect of behavioral training on the cortical representation, and 4) to relate behaviorally determined performance levels to neurophysiological measures. The multi-unit mapping approach that was used in these experiments was not only necessary to acquire an adequate cell sample per animal, but it also allowed for the comparison of response strength relative to cell position in cortical space for each of the stimulus presentation conditions. Satisfying this later condition was essential for analyzing the spatially-distributed response activity to speech stimuli with respect to the functional topographies found in the primary auditory cortex. Furthermore, since it was hypothesized that context-induced functional changes in receptive field parameters, such as sharpness of tuning, may be effective in serving a normalization process for stimulus-evoked activity, analyzing the spatial topography of neural activity under several contextual backgrounds was essential. Whether biological relevancy of the speech sounds used in these experiments alters the representation topographically within and across the frequency axis, which has been implicitly suggested by results from cortical studies of other sensory and motor systems, was a question that required the use of such neurophysiological methods as well.

The results are summarized as follows: 1) The spectro-temporal features of the speech stimuli were represented by a reduction in the temporal response pattern of cortical neurons to include only phasic responses to transitions in the stimulus representation. 2) The spectral content of the speech stimuli was represented spatially by focal activation of several locations along the isofrequency dimension of specific frequencies that do not necessarily correspond to the formant frequencies. 3) The cortical responses to speech sounds were strongly level dependent not only in terms of response strength but also with respect to the spatially-distributed response pattern elicited by the stimulus. 4) The magnitude of the cortical response to speech was strongly dependent on the level of the background noise. 5) The spatially-distributed response pattern was independent of changes in the background noise level, until neural activity was largely suppressed. 6) Behavioral detection and discrimination thresholds correspond to cortical excitation patterns of strongly reduced magnitude and spatial salience. 7) Differences in response selectivity between sharply- and broadly-tuned neurons may reflect cortical change attributed to behavioral training.

3.3.1 Relation to previous studies on the neural representation of speech sounds

Much of the work concerning the neural representation of speech sounds in the mammalian auditory system has been done at processing levels peripheral to the primary auditory cortex such as in the auditory nerve and the cochlear nucleus (Delgutte and Kiang, 1984; May, Prell, Sachs, 1998; Sachs, Voigt, and Young, 1983; Sachs and Young, 1979; Sinex and Geisler, 1983). The temporal response structure of the neurophysiological behavior from auditory nerve fibers and cells in the cochlear nucleus contrast to that found in the cortex due to a reduction in temporal response following capabilities to the carrier signal in favor of phase-locking to rapid transitions in the temporal envelope of more central neurons. This has been evident in the phasic responses found in the present study as well as in previous work recording cortical responses to speech stimuli (in cats: Eggermont, 1995; in monkeys: Steinschneider, Arezzo, and Vaughn, 1990). The increased neural firing to the onsets of the consonant burst and voicing that was present in primary auditory cortex lends further support to the proposition that perceptual features of stop consonant-vowel syllables may be represented in the temporal properties of cortical neurons as they appear to be in peripheral nerve fibers (Carney and Geisler, 1986).

In the awake monkey, it has been found that multiple unit activity in the primary auditory cortex to different consonant-vowel syllables vary in their responses to reflect spectral characteristics of the steady-state formants and formant transition onset frequencies (Steinschneider, Arezzo, and Vaughn, 1990). Furthermore, the phonetic feature of the voiced onset time (VOT) has been found in the awake monkeys (Steinschneider et al., 1994; Steinschneider et al., 1995) and in anesthetized cats (Eggermont, 1995) to be correlated with the temporal response patterns of cells in the primary auditory cortex. In the present study, the temporal activity corresponding to spectral events within the spectrogram were not only represented in the temporal response patterns of the multi-unit neuron cluster, but more importantly they were organized spatially by localized and synchronized neural activity along the frequency axis and across the isofrequency dimension. The synchronization of neural activity from noncontiguously-spaced subpopulations of neurons was a dominant feature found in the cortical response of cats to the speech syllables as has also been seen for species-specific vocalizations in the monkey (Wang, et al., 1995). More specifically, the neural pattern of synchronization corresponding to the consonant burst as represented tonotopically by the primary auditory cortex appeared to be similar for the voiced (/be/) and unvoiced (/pe/) syllable at frequencies corresponding to the second and third formant frequencies. The absence of low-frequency representation in the synchronized response to /pe/ could be attributed to acoustic properties of the unvoiced consonant burst. Based upon the observations from the reconstructed cortical maps of neural activity to the speech sounds, it appeared that detailed information about the spectra of the consonant-vowel syllables may be preserved in the synchronization of temporal activity of cortical neurons which has been proposed for speech-encoding in auditory-nerve fibers (Sinex and Geisler, 1983). That is, spatially-localized clusters of neurons responding in temporal synchrony to spectral components of speech sounds may be the method by which AI preserves spectro-temporal information of incoming signals.

3.3.2 The effects of stimulus level on rate encoding of speech

Overall, the cortical representation of speech sounds was altered with changes to stimulus intensity. This finding is not surprising given our knowledge of level-dependent cortical responses to pure tones (Brugge and Merzenich, 1973; Heil, Rajan, and Irvine, 1994; Phillips, et al., 1985; Sutter and Schreiner, 1995) and ripple stimuli (Schreiner and Calhoun, 1994) as reflected by changes in bandwidth and response strength. Level-dependence of average rate discharge encoding of speech sounds was clearly illustrated in the mean PST histograms, spectral profiles, and cortical maps. In response to /be/, level-dependent spike activity was reflected by the reduction in the magnitude of the mean peak response and the reduction in the spectral spread of spike activity across the CF representation with 10 dB decreases in stimulus intensity. The decrease in spatial correlation of spatially-distributed spike activity for the peak response to /be/ also exemplified level-dependence of rate encoding in spatial terms. In response to /pe/, level-dependent spike activity was reflected by a changing balance in the magnitude of spike activity between first and second peak response. That is, the reduction in spike rate for the first response to 10 dB decreases in stimulus intensity were complemented by a slight increase at localized CF locations for the second response which subsequently decreased as well further decreases in stimulus intensity. Likewise, the decrease in spatial correlation of spatially-distributed spike activity for the either peak response to /pe/ further substantiates the claim of level-dependent behavior in the average spike rate of cortical neurons.

3.3.3 The effects of background noise on rate encoding of speech

In general, the addition of continuous background noise resulted in a gradual reduction of response magnitudes and extension of response latency. Both effects are to be expected from previous studies of noise effects on cortical responses (Phillips, 1985; Phillips and Cynader, 1985). However, the suppressive effects of background noise on the average discharge rate of cortical neurons was found to be effective only in partial regions of the primary auditory cortex which may be related to differences in receptive field properties of the underlying cells. Although the spatial pattern of cortical activity in the presence of background noise did not change substantially in either the first or second response intervals (except at the highest noise level which drastically masked the first response), which was supported by high spatial correlation coefficients (of at least, r=.60), shifts in localized spike activity were observed along the isofrequency dimension. As reflected in the cortical maps for the first temporal response intervals, when the signal-to-noise ratio decreased from +7.5 dB to +4.8 dB, the synchronized response from a more dorsal region along the 2 kHz isofrequency contour decreased while another region more ventral along the same isofrequency contour increased in responsiveness; whereas, for the second teporal response interval, the synchronized response from the more ventral region along the 2 kHz isofrequency contour decreased while another region more dorsal along the same isofrequency increased. Since this reversal of response strength along the isofrequency dimension was also present in the first temporal response, factors related to release of inhibition such as a decrease in forward masking inhibition from the first spectral event can not adequately explain the increased neural activity with not only the addition of background noise, but also an increase in the level of background noise until the response is effectively masked.

Differential rate-encoded response to the effects of background noise was also revealed in the spectral profiles to each speech syllable. The spectral profile of the cortical response to the voiced and unvoiced speech syllable showed that at moderate background noise levels average discharge rate actually increased in cells with CF near the formant structure of the CV and decreased in cells with CF much above the third formant. At high background noise levels, the rate-encoded spectral information was degraded to near zero driven-rate levels due to suppression by the noise.

Furthermore, whereas the overall reduction in the mean first peak response amplitude with 10 dB decreases in the noise level was on the similar scale as the reduction seen with 10 dB decreases in signal level, the spectral spread of excitation, which was observed in the spectral profile and cortical maps, was maintained more with decreases in signal-to-noise level than in signal level. Specifically, as was depicted by the correlation coefficients calculated for the cortical maps, the spatial correlation decreased more rapidly for 10 dB reductions in stimulus level than in background noise level. Therefore, the decrease in spectral selectivity and sensitivity to softer signals reflects the reduction in neural recruitment a mechanism which accounts for growth in the loudness function at lower processing stages. The global change in spectral selectivity and sensitivity only at the highest background noise level when signal was fixed at one intensity level may be a reflection of critical band-like behavior in which neural activity becomes suppressed or swamped by the noise.

Studies in the auditory nerve and the ventral cochlear nucleus have shown that rate-place vowel representation in the presence of background noise is differentially encoded by subgroups of cells based on spontaneous rate (Sachs, Voigt, and Young, 1983) and cell types (May, Le Prell, and Sachs, 1998) for the respective processing levels. Sachs and colleagues (1983) reported that in the presence of background noise, 1) low spontaneous rate fibers, which provide better representation of vowel formant structure at moderate stimulus levels, were less able to maintain spectral peak information in their discharge rate than the high spontaneous rate fibers; 2) the suppression of noise-evoked rate activity was greater for low spontaneous fibers than for high; and 3) as a whole, the spectral profile changed drastically in the presence of background noise such that at signal-to-noise ratios of +9 dB, spectral peak responses near the second and third formant frequencies were diminished. In the cochlear nucleus of awake cats, May and colleagues (1998) showed that chopper units, which manifest degraded representations of the temporal content of the vowel compared to the auditory nerve or primary-like units in the cochlear nucleus (Blackburn and Sachs, 1990), were more robust to background noise effects on average discharge rate than the primary-like units. Moreover, they reported that within the chopper units, there was a subgroup (ChS units) that provided the best representation of formant features across all vowel levels and in the presence of background noise.

Thus, divisions in the representation of physiological responses to speech with or without the presence of background noise has already been precedented in auditory processing levels peripheral to the primary auditory cortex. However, one can only speculate at this point to what degree differential rate processing of speech information at the cochlear nucleus is transformed and/or maintained in the differential processing that has been found to exist along the isofrequency dimension of primary auditory cortex with or without background noise.

3.3.4 Effects of background noise on temporal encoding of speech

The temporal response to the voiced and unvoiced speech syllables were found to be affected by high background noise levels with only minimal effect at moderate background noise levels. At high background noise levels, the temporal response as reflected by peak latency and peak width to the voiced CV was degraded by suppression from the noise, while the temporal response to the unvoiced CV was only partially maintained by the second peak response in CF units below the second formant frequency. At moderate background noise levels, occurrence of the evoked-temporal event(s) (two in the case for /pe/) was maintained with no remarkable shifts in latency as would have been observed in the response to simple stimuli in the presence of continuous background noise. The temporal synchrony of neural activity locally-distributed spatially in the primary auditory cortex was also robust in the presence of moderate background noise levels (S/N>4.8 dB), although subtle shifts in the average discharge rate along the isofrequency dimension were detectable. At low signal-to-noise ratios (S/N=2.8 dB), the temporal synchrony of the responding cortical regions to the onset of voicing in the unvoiced CV changed slightly in spatial pattern along the isofrequency dimension. The implications of these subtle shifts in average rate discharge will be discussed in a later section. In the auditory nerve, temporal-place representations (i.e., phase-locking properties and average discharge rate) of vowel formants (Delgutte and Kiang, 1984; Sachs, Voigt, and Young, 1983) and CV syllables (Geisler and Gamble, 1989) are minimally affected by background noise. Therefore, with regards to temporal processing of complex sounds, it appears that the temporal representation is robust in low to moderate levels of background noise, but changes drastically for high levels of background noise. The degradation in the temporal response at low signal-to-noise levels (or high noise levels) is interpreted as being a result of peripheral adaptation and a reduction of detectable signal energy.

3.3.5 The effect of speech discrimination training on the cortical representation of speech

The issue of using species-specific communication calls to study physiological behavior to complex stimuli in cortical neurons is relevant for discussion at this time. It has been argued that one of the advantages for using species-specific calls to study neurophysiological activity to complex sounds is the biological relevancy of the stimulus (Bullock, 1976; Wang, et al., 1995). That is, the neurons have been “trained” to recognize calls made by the species, and thus, may respond preferentially to certain aspects of those calls. However, parametric study of neural representations of the species-specific calls would require knowledge of the spectral or temporal cues that cause the neuron to respond preferentially to one or another. In addition, attempts at correlating phonemic content with neurophysiological behavior would be even more challenging due to the inaccessibility of ascribing phonemes to animal vocalizations. In order to circumvent these issues, animals were trained to recognize and to make discriminations of speech syllables, for which the phonemic content has been covered by extensive literature; and thus, have physiological data recorded with behaviorally-relevant stimuli.

The basis for training animals to make behavioral discriminations of speech syllables comes from arguments that reorganization of cortical representations is a lifelong process that reflects an individual's ability to acquire new or improve old skills and behaviors (Merzenich, et al., 1988, 1990). Modification of the tonotopic organization in AI of owl monkeys following an auditory frequency discrimination task has provided evidence for training-induced cortical plasticity (Recanzone, et al., 1993) in the auditory system. Furthermore, it has been shown that behavioral relevance of the stimulus is essential in inducing changes in the temporal contingency between neurons that modulate cortical plasticity in the adult monkey (Ahissar, et al., 1992).

In the current study, significant differences attributable to behavioral training were discovered for rate-encoding and temporal encoding of speech CVs. Most notably with respect to the naive animals, mean firing rates decreased, mean peak latencies lengthened, and mean temporal precision reduced in the group of animals trained to make two behavioral discriminations (first: /pe/ vs /ko/, then /pe/ vs /be/). In contrast, the second group of trained animals who was only trained on a less challenging behavioral discrimination task (/be/ vs /ko/) did not show an overall decrease in response magnitude and manifested shorter peak latencies . Additionally in the second group of trained animals, firing rate was selectively encoded by subgroups of neurons based upon their sharpness of tuning receptive field parameter such that broadly-tuned neurons responded more strongly than sharply-tuned neurons. Neither the naive animals nor animals trained on two discrimination tasks showed any selective response difference based upon sharpness of tuning spectral parameter. This was illustrated by the cortical maps from the trained animal depicting synchronized neural activity from regions other than the central sharply-tuned area. Unlike the broadly-tuned areas, the central, sharply-tuned neurons have not only been shown to respond poorly to transients (Schreiner and Mendelson, 1990), but they also have been found to prefer slow, and downward-directed FM sweeps (Mendelson, et al., 1992). The broadly-tuned areas have been found to respond best to broadband stimuli (Schreiner and Mendelson, 1990) and prefer fast or medium upward-directed FM sweeps (Mendelson, et al., 1992). Therefore, in the case of the animals trained to make behavioral discriminations between /be/ (characterized by a fast upward-directed FM sweep between the consonant and vowel formant frequencies) and /ko/ (characterized by a fast downward-directed FM sweep for the F2 formant frequency), response to complex sounds from specific subaggregates of neurons distinguished by their sharpness of tuning and response to FM sounds were enhanced by the training.

The diverse changes in the neurophysiological responses to CV syllables may be interpreted as reflecting functional reorganizations specific to the training paradigm differences between the two trained groups of animals. The decline of neural activity and temporal precision observed in the animals trained on two different phonetic discrimination tasks (the second one involving a discrimination of VOT) may reflect a neural habituation to phonetic contrasts. In the case of the animals trained on one phonetic discrimination task, neural enhancement of a subpopulation of neurons (e.g., broadly-tuned neurons) may be indicative of an intermediary stage prior to neural habituation. These results require further investigation before concrete statements about "language" learning and changes in spatial-temporal neural events in the primary auditory cortex may be made. However, these results are suggestive of analogous evidence implicating functional reorganizations in the developmental progression of language acquisition in infants (Kuhl, 1994 ; Stager and Werker, 1997; Werker and Tees, 1984). That is, the sharpening of perceptual boundaries necessary to make fine phonetic feature distinctions of rapid FM transitions, timing information of voicing, or relative spectral energy peaks within the spectra of the speech sounds is a cortical reorganization process common to both language-developing infants and behaviorally-trained animals. Evidence suggesting that changes in spatial-temporal cortical activity are reflective of either different or sharpened perceptual boundaries in animals may yield great insights to cortical processing of phonemic information in speech sounds.

3.3.6 Implications of generalizations in the cortical processing of CV syllables on speech discrimination

Representational generalizations in AI response properties to signals in background noise have been alluded to in a previous study finding critical band-like behavior in central and ventrally-located AI neurons (Ehret and Schreiner, 1997). Furthermore, it has been postulated that the central and ventral non-monotonic regions, which are sharply-tuned and low in threshold compared to other locations along the isofrequency domain, are specialized for detecting signals just above background noise (Sutter and Schreiner, 1995). Based upon experimental mapping results reported in chapter two, it was also postulated that subgroups of neurons categorized by their sharpness of tuning response property played a role in maintaining a normalization of cortical activity to signals in background noise.

Irrespective of the mechanism by which normalization of cortical responses may be achieved, generalization of synchronized, locally-distributed, spatial-temporal responses in the presence of background noise at signal-to-noise ratios corresponding to behavioral discrimination abilities were found to exist. That is, the cortical spatial-temporal representation of CV syllables was robust to changes in signal-to-noise ratios at levels comparable to speech discrimination behavior. It was not clear, however, whether the magnitude of spatial-temporally synchronized spike activity was reflective of either behavioral detection or discrimination thresholds since neural activity was still visible in the second temporal response at signal-to-noise levels below even the detection threshold measured. There is a strong possibility that the animal's ultimate sensitivity and spectral analysis capabilities were not reflected in the thresholds obtained during training; however, the goal here was to find a neurophysiologic correlate for behavioral discrimination. Thus, driving the animal to asymptotic discrimination performance was not critical in this study. Therefore, the interpretation of the cortical map findings suggests that speech discrimination in the various background contexts is enabled by specific spatial patterns of cortical activity rather than by level-dependent changes in response strength and latency. However, generalizations of speech sounds in various signal intensities are not as obviously depicted by maintenance of spatial patterns within the cortical map representation. Further study is needed to reveal normalization processes that allow for the generalization of speech sounds across various sound levels.

3.3.7 Closing remarks

The representation of speech sounds in AI neurons can be considered to be a result of complex interactions across many different receptive field properties that reflect auditory processing from subcortical levels and that are shaped further by different auditory experiences. The ability to extract phonetic features from speech sounds, which has been demonstrated in auditory stages as early as in the auditory nerve, is best represented spatial-temporally in the primary auditory cortex at least in different background contexts. The mechanisms by which generalizations of sound speech across different listening contexts are still at best speculative and require further study in the future.

Conclusions

This research has used tonal stimuli and synthesized speech sounds presented under varying stimulus intensities and background noise levels to investigate how changes in signal-to-noise levels are reflected by the representation of sound in the primary auditory cortex. The experimental studies examined in detail the effects of background noise on 1) the spectral receptive field properties of single units and multiple-units ; 2) the corresponding functional topographies; and 3) the spatial-temporal neural activity to speech syllables in AI. The main findings were that 1) FRA bandwidth changes inversely in the presence of background noise which was reflected in the single unit response and in spatially-localized regions in the topography of bandwidth and 2) spatial-temporally activated neural activity patterns evoked by speech syllables are robust to changes in background noise levels than in stimulus intensity. The implications of these findings are relevant in understanding how background contexts impact the cortical response activity that account for the perceptual invariances that we have for speech sounds.

The conduct of this research in animal models was essential in understanding the neurophysiological representations of speech on a cellular basis. The use of non-invasive imaging techniques, which have been used to study how the brain processes speech sounds in humans, falls short of providing a link between cellular response behavior and spectro-temporal aspects distinguishing different speech sounds. Research in animals has shown that categorical perception, which is based on the premise that perceptual boundaries exist that allows for the spectro-temporal discriminations among speech sounds, is not exclusively a human ability (Kluender et al., 1987; Kuhl, 1981). This finding may not be such a surprise, in part, since categorical perceptual abilities have been found using non-speech sounds in humans (Miller et al., 1976). The existence of perceptual boundaries, thus, can be thought of as being a general property of sensory behavior across species. Although the methodology used in behavioral training of this research project did not force animals to use their perceptual limits before making discriminative decisions of speech sounds, the neurophysiological results suggest that perceptual boundaries may be revealed by linking it to differences in spatial-temporally distributed cortical activity. The results of this study are the first to show how speech is represented spatial-temporally in primary auditory cortex at different sound intensities and in different background levels.

Our ability to make fine distinctions of phonetic contrasts in the presence of background noise is an important aspect to consider in the practical sense when talking about speech intelligibility in noisy backgrounds. Intelligible speech is what most people rely on, but take for granted, for normal communication. This study has shown that the spectral and temporal analysis of speech sounds that the auditory system performs for feature extraction is represented as spatial-temporally synchronized neural events in the primary auditory cortex. These neural events occur at the onsets of spectral transitions. Even in the presence of background noise, these spatial-temporal abstractions of speech sounds do not change significantly at signal-to-noise levels which allow for speech discrimination ability, except at high noise levels. Cortical encoding of speech based on spatially-distributed synchronized activity may prove to be a neurophysiologic correlate for measuring speech intelligibility in the primary auditory cortex. This may prove to be helpful in understanding how sensori-neural hearing impaired individuals have markedly reduced speech intelligibility in the presence of competing background sounds. Perhaps in sensori-neural hearing impaired listeners the robustness of spatial-temporal neural activity patterns is compromised by degradations in the neural processing of acoustic information at lower auditory stations. In humans with a whole other set of hearing impairments, namely those stemming from central auditory neuropathies, whose problems arise only when trying to discriminate and recognize words. The problem that these hearing impaired individuals are suffering from may originate from loss in the synchrony of activity bounded temporally in cortical space.

The cortical representations of speech sounds as it is manifested by synchronized spatial-temporal neural activity may help in developing instructional aids for people with language learning problems attributable to deficits in auditory processing. Deficits in temporal processing of speech sounds has been seen in language-impaired children (Tallal et al., 1993). It has been shown that specific auditory temporal training can significantly speed up auditory-processing rates in language learning impaired children and that improved processing rates are correlated with improved speech processing (Merzenich, et al., 1996; Tallal, et al., 1996). The sharpening in the timing of neural events to specific acoustic events has been thought to occur with increased exposure as would occur with auditory training. It was found in these studies though that auditory discrimination training requiring more difficult acoustic feature extraction such as a difference in voice onset time was not entirely indicative of increased temporal precision of neural activity to stimulus onsets, whereas discrimination tasks requiring easier phonetic difference distinctions manifested increased temporal precision. Further study of the data collected for the CV sounds which were used in behavior training, however, may clarify this unsettled point. An emphasis of temporal precision of only an elite set of neurons localized in spatial proximity to each other rather than the entire spatially-distributed cortical pattern may be the end result of intensive auditory discrimination training. Furthermore, the enhancement of only a subset of cortical neurons, as demonstrated through the sharpening of their temporal precision for example, may be the basis for identifying the perceptual limits required for making accurate discriminations of subtle phonetic differences in speech sounds.

In summary, the study of cortical representations of speech using animal models is still in its initial stages of discovering the complexities of how the brain operates to allow us the ability to listen before we act, a process commonly referred to as communication. The importance of studying the representations of tones and speech sounds in a variety of background noise levels can not be more obvious especially given the demands that we are faced with on a daily basis to perceive sounds in time-varying background contexts. By incorporating the aspect of behavioral training into the methodology, behavioral correlates of these cortical representations can then be made to draw conclusions that may eventually lead to a neurophysiological understanding of speech perception. Future research using this type of methodological approach will be effective in shedding more light on the cortical organization of speech which will offer a link to other disciplines that are also interested in the study of speech sounds.

References

Ahissar, E., Vaadia, E., Ahissar, M., Bergman, H. Arieli, A., and Abeles, M. (1992) Dependence of cortical plasticity on correlated activity of single neurons and on behavioral context. Science, 257(5075), 1412-5.

Andersen, R.A., Knight ,P.L., and Merzenich M.M. (1980) The thalamocortical and corticothalamic connections of AI, AII, and the anterior auditory field (AAF) in the cat: evidence for two largely segregated systems of connections. Journal of Comparative Neurology. 194(3), 663-701.

Benson, D.A., and Teas, D.C. (1976) Single unit study of binaural interaction in the auditory cortex of the chinchilla. Brain Research, 103(2), 313-38.

Blackburn, C.C., and Sachs, M.B. (1990) The representations of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. Journal of Neurophysiology, 63(5), 1191-212.

Brosch, M., and Schreiner, C.E. (1997) Time course of forward masking tuning curves in cat primary auditory cortex. Journal of Neurophysiology. 77, 923-943.

Brugge, J.F., Dubrovsky, N.A., Aitkin, L.M., and Anderson, D.J. (1969) Sensitivity of single neurons in auditory cortex of cat to binaural tonal stimulation: effects of varying interaural time and intensity. Journal of Neurophysiology. 36, 1138-1158.

Brugge, J.F., and Merzenich, M.M. (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. Journal of Neurophysiology. 32, 1136-1158.

Bullock, T. (1976) Recognition of Complex Acoustic Signals: Report of the Dahlem Workshop on Recognition of Complex Acoustic Signals. 279-306.

Caird, D.M., Palmer, A.R., and Rees, A. (1991) Binaural masking level difference effects in single units of the guinea pig inferior colliculus. Hearing Research. 57, 91-106.

Calford, M.B., and Semple, M.N. (1995) Monaural inhibition in cat auditory cortex. Journal of Neurophysiology. 73, 1876-1891.

Carney, L.H., and Geisler, C.D. (1986) A temporal analysis of auditory-nerve fiber responses to spoken stop consonant-vowel syllables. Journal of the Acoustical Society of America, 79(6), 1896-914.

Chan, J.C.K., Yin, T.C.T., and Musicant, A.D. (1987) Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. II. Responses to band-pass filtered noises. Journal of Neurophysiology. 58, 543-561.

Code, R.A., and Winer, J.A. (1985) Commissural neurons in layer III of cat primary auditory cortex (AI): pyramidal and non-pyramidal cell input. Journal of Comparative Neurology. 242(4), 485-510.

Code, R.A., and Winer, J.A. (1986) Columnar organization and reciprocity of commissural connections in cat primary auditory cortex (AI). Hearing Research, 23(3), 205-22.

Colburn, H.S. (1977) Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. Journal of the Acoustical Society of America, 61(2), 525-33.

Delgutte, B. (1990) Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers. Journal of the Acoustical Society of America, 87(2), 791-809.

Delgutte, B., and Kiang, N.Y. (1984) Speech coding in the auditory nerve: I. Vowel-like sounds. Journal of the Acoustical Society of America, 75(3):866-78.

Delgutte, B., and Kiang, N.Y. (1984) Speech coding in the auditory nerve: V. Vowels in background noise. Journal of the Acoustical Society of America, 75(3), 908-18.

Durlach, N.I., Gabriel, K.J., Colburn, H.S., and Trahiotis, C. (1986) Interaural correlation discrimination: II. Relation to binaural unmasking. Journal of the Acoustical Society of America, 79(5), 1548-57.

Dye, R.H. Jr, and Hafter, E.R. (1980) Just-noticeable differences of frequency for masked tones. Journal of the Acoustical Society of America, 67(5), 1746-53.

Eggermont, J.J. (1995) Representation of a voice onset time continuum in primary auditory cortex of the cat. Journal of the Acoustical Society of America, 98(2 Pt 1), 911-20.

Ehret, G., and Merzenich, M.M. (1988) Complex sound analysis (frequency resolution, filtering and spectral integration) by single units of the inferior colliculus of the cat. Brain Research, 472(2):139-63.

Ehret, G., and Schreiner, C.E. (1997) Frequency resolution and spectral integration (critical band analysis) in single units of the cat primary auditory cortex. Journal of Comparative Physiology. a. Sensory. Neural. and Behavioral Physiology, 181(6), 635-50.

Erulkar, S.D., Rose, J.E., and Davies, P.W. (1956) Single unit activity in the auditory cortex of the cat. Bulletin of Johns Hopkins Hospital. 99, 55-86.

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

Geisler, C.D., and Gamble, T. (1989) Responses of "high-spontaneous" auditory-nerve fibers to consonant-vowel syllables in noise. Journal of the Acoustical Society of America, 85(4), 1639-52.

Geisler, C.D., and Sinex, D.G. (1980) Responses of primary auditory fibers to combined noise and tonal stimuli. Hearing Research, 3(4), 317-34.

Goldberg, J.M., and Brown, P.B. (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. Journal of Neurophysiology, 32(4), 613-36.

Goldstein, D.P. and Stephens, S.D.G. (1975) Masking level difference: a measure of auditory processing capability. Audiology. 14, 354-367.

Hall, J.L. 2d, and Goldstein, M.H. Jr. (1968) Representation of binaural stimuli by single units in primary auditory cortex of unanesthetized cats. Journal of the Acoustical Society of America, 43(3), 456-61

Heil, P., Rajan, R., and Irvine, D.R.F. (1994) Topographic representation of tone intensity along the isofrequency axis of cat primary auditory cortex. Hearing Research. 76, 188-202.

Henning, G.B. (1974) Lateralization and the binaural masking-level difference. Journal of the Acoustical Society of America. 55, 1259-1262.

Henning, G.B., and Gaskell, H. (1981) Binaural masking level differences with a variety of waveforms. Hearing Research. 4, 175-184.

Hienz, R.D., Sachs, M.B., and Aleszczyk, C.M. (1993) Frequency discrimination in noise: comparison of cat performances with auditory-nerve models. Journal of the Acoustical Society of America, 93(1), 462-9.

Hirsch, I.J. (1948) The influence of interaural phase on interaural summation and inhibition. Journal of the Acoustical Society of America. 20, 536-544.

Houser, C.R., Vaughn, J.E., Hendry, S.H.C., Jones, E.G., and Peters, A. (1984) GABA neurons in the cerebral cortex. In: Cerebral Cortex, Functional Properties of Cortical Cells, eds. Jones, E.G. and Peters, A. New York: Plenum, vol.2, 63-89.

Imig, T.J. and Adrian, H.O. (1977) Binaural columns in the primary field (AI) of cat auditory cortex. Brain Research. 138, 241-257.

Imig, T.J., and Brugge, J.F. (1978) Sources and terminations of callosal axons related to binaural and frequency maps in primary auditory cortex of the cat. Journal of Comparative Neurology, 182(4), 637-60.

Imig, T.J., and Reale, R.A. (1980) Patterns of cortico-cortical connections related to tonotopic maps in cat auditory cortex. Journal of Comparative Neurology, 192(2), 293-332.

Imig, T.J., and Reale, R.A. (1981) Ipsilateral corticocortical projections related to binaural columns in cat primary auditory cortex. Journal of Comparative Neurology, 203(1), 1-14.

Jeffress, L.A. (1948) A place theory of sound localization. Journal of Comparative Physiology and Psychology. 41, 35-39.

Jeffress, L.A., Blodgett, H.C., and Deatherage, B.H. (1952) The masking of tones by white noise as a function of the interaural phases of both components. Journal of the Acoustical Society of America. 24, 523-527.

Jeffress, L.A., Blodgett, H.C., Sandel, T.T., and Wood, C.L. (1956) Masking of tonal signals. Journal of the Acoustical Society of America. 28 416-426.

Jiang, D., McAlpine, D., and Palmer, A.R. (1997) Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. Journal of Neurophysiology. 77, 3085-3106.

Jones, E.G. (1984) Identification and classification of intrinsic circuit elements in the neocortex. In: Dynamic aspect of neocortical function, eds. Edelman, G.M., Gall,W.E., Cowan, W.M. New York: Wiley, 7-40.

Kiang, N.Y-S, Neame, J.H., and Clark, L.F. (1961) Evoked cortical activity from auditory cortex in anesthetized and unanesthetized cats. Science. 133, 1927-1928.

Kluender, K.R., Diehl, R.L., and Killeen, P.R. (1987) Japanese quail can learn phonetic categories. Science, 237(4819):1195-7.

Kuhl, P.K. (1981) Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. Journal of the Acoustical Society of America. 70(2), 340-349.

Kuhl, P.K. (1994) Learning and representation in speech and language. Current Opinion in Neurobiology, 4(6), 812-22.

Kuwada, S., and Yin, T.C.T. (1983) Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and

repetition rate on interaural delay function. Journal of Neurophysiology. 50, 981-999.

Langford, T.L., and Jeffress, L.A. (1964) Effect of noise crosscorrelation on binaural signal detection. Journal of the Acoustical Society of America. 36, 1455-1458.

LeVay, S., and Gilbert , C.D. (1976) Laminar patterns of geniculocortical projection in the cat. Brain Research, 113(1), 1-19.

Licklider, J.C.R. (1948) Influence of interaural phase relations upon the masking of speech by white noise. Journal of the Acoustical Society of America. 20, 150-159.

May, B.J., Le Prell, G.S., and Sachs, M.B. (1998) Vowel representations in the ventral cochlea nucleus of the cat: effects of level, background noise, and behavioral state. Journal of Neurophysiology. 79(4), 1755-1767.

Mendelson, J.R., Schreiner, C.E., Sutter, M.L., and Grasse, K.L. (1993) Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps. Experimental Brain Research, 94(1):65-87.

Mendelson, J.R., Schreiner, C.E., and Sutter, M.L. (1997) Functional topography of cat primary auditory cortex: response latencies. Journal of Comparative Physiology A. 181, 615-633.

Merzenich, M.M., Jenkins, W.M., Johnston, P., Schreiner, C., Miller, S.L., Tallal, P. (1996) Temporal processing deficits of language-learning impaired children ameliorated by training. Science, 1996 Jan 5, 271(5245):77-81.

Merzenich, M.M., Recanzone, G., Jenkins, W.M., and Nudo, R.J. (1990) How the brain functionally rewires itself. In: Natural and Artificial Parallel Computations, eds. Arbib, M. and Robinson, J.A. Cambridge, MA: MIT Press, 177-210.

Merzenich, M.M., Recanzone, G., Jenkins, W.M., Allard, T.T., and Nudo, R.J. (1988) Cortical representational plasticity. In: Neurobiology of Neocortex, eds. Rakic, P. and Singer,W. New York: Wiley & Sons, 41-67.

Merzenich, M.M., Knight, P., and Roth, G.L. (1975) Representation of the cochlea within the primary auditory cortex in the cat. Journal of Neurophysiology. 38 (1), 231-249.

Metherate, R., and Weinberger, N.M. (1989) Acetylcholine produces stimulus-specific receptive field alterations in cat auditory cortex. Brain Research, 480(1-2),372-7.

McAlpine, D., Jiang, D., and Palmer, A.R. (1996a) Binaural masking level differences in the inferior colliculus. Journal of the Acoustical Society of America. 100, 490-503.

McKenna, T.M., Ashe, J.H., Hui, G.K., and Weinberger, N.M. (1988) Muscarinic agonists modulate spontaneous and evoked unit discharge in auditory cortex of cat. Synapse. 2(1), 54-68.

Middlebrooks, J.C., Dykes, R.W., and Merzenich, M.M. (1980) Binaural response-specific bands in primary auditory cortex (AI) of the cat: topographical organization orthogonal to isofrequency contours. Brain Research. 181, 31-48.

Miller, J.D., Wier, C.C., Pastore, R.E., Kelly, W.J., and Dooling, R.J. (1976) Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. Journal of the Acoustical Society of America. 60(2), 410-417.

Orman, S.S., and Phillips, D.P. (1984) Binaural interactions of single neurons in posterior field of cat auditory cortex. Journal of Neurophysiology, 51(5), 1028-39.

Osman, E. (1971) A correlation model of binaural masking level differences. Journal of the Acoustical Society of America. 50, 1494-1511.

Phillips, D.P. (1985) Temporal response features of cat auditory cortex neurons contributing to sensitivity to tones delivered in the presence of continuous noise. Hearing Research. 19, 253-268.

Phillips, D.P. (1988) Effect of tone-pulse rise time on rate-level functions of cat auditory cortex neurons: excitatory and inhibitory processes shaping responses to tone onset. Journal of Neurophysiology. 59, 1524-1539.

Phillips, D.P. (1990) Neural representation of sound amplitude in the auditory cortex: effects of noise masking. Behavioral Brain Research. 37, 197-214.

Phillips, D.P., and Cynader, M.S. (1985) Some neural mechanisms in the cat's auditory cortex underlying sensitivity to combined tone and wide-spectrum noise stimuli. Hearing Research. 18, 87-102.

Phillips, D.P., and Hall, S.E. (1992) Multiplicity of inputs in the afferent path to cat auditory cortex neurons revealed by tone-on-tone masking. Cerebral Cortex, 2(5), 425-33.

Phillips, D.P., and Hall, S.E. (1986) Spike-rate intensity functions of cat cortical neurons studied with combined tone-noise stimuli. Journal of the Acoustical Society of America. 80, 177-187.

Phillips, D.P., and Farmer, M.E. (1990) Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. Behavioral Brain Research, 40(2), 85-94.

Phillips, D.P., and Irvine, D.R.F. (1983) Some features of the binaural input to single neurons in physiologically defined area AI of cat cerebral cortex. Journal of Neurophysiology. 49, 383-395.

Phillips, D.P., and Kelly, J.B. (1992) Effects of continuous noise maskers on tone-evoked potentials in cat primary auditory cortex. Cerebral Cortex. 2(2), 134-140.

Phillips, D.P., Orman, S.S., Musicant, A.D., and Wilson, G.F. (1985) Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. Hearing Research. 18, 73-86.

Phillips, D.P, Semple, M.N., Calford, M.B., and Kitzes, L.M. (1994) Level-dependent representation of stimulus frequency in cat primary auditory cortex. Experimental Brain Research. 102, 210-226.

Reale, R.A., and Brugge, J.F. (1990) Auditory cortical neurons are sensitive to static and continuously changing interaural phase cues. Journal of Neurophysiology. 64, 1247-1260.

Reale, R.A., and Imig, T.J. (1980) Tonotopic organization in auditory cortex of the cat. Journal of Comparative Neurology. 192, 265-291.

Reale, R.A., and Kettner, R.E. (1986) Topography of binaural organization in primary auditory cortex of the cat: effects of changing interaural intensity. Journal of Neurophysiology. 36(3), 663-682.

Recanzone, G.H., Schreiner, C.E., and Merzenich, M.M. (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. Journal of Neuroscience, (1), 87-103.

Recanzone, G.H., Merzenich, M.M., and Schreiner, C.E. (1992) Changes in the distributed temporal response properties of SI cortical neurons reflect improvements in performance on a temporally based tactile discrimination task. Journal of Neurophysiology, 67(5), 1071-91.

Rhode, W.S., Geisler, C.D., and Kennedy, D.T. (1978) Auditory nerve fiber response to wide-band noise and tone combinations. Journal of Neurophysiology, 41(3), 692-704.

Rose, J.E., Gross, N.B., Geisler, C.D., and Hind, J.E. (1966) Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. Journal of Neurophysiology. 30, 769-793.

Rouiller, E.M., Simm, G.M., Villa, A.E., de Ribaupierre, Y., and de Ribaupierre F. (1991) Auditory corticocortical interconnections in the cat: evidence for parallel and hierarchical arrangement of the auditory cortical areas. Experimental Brain Research, 86(3), 483-505.

Sachs, M.B., Voigt, H.F., and Young, E.D. (1983) Auditory nerve representation of vowels in background noise. Journal of Neurophysiology, 50(1), 27-45.

Sachs, M.B., and Young, E.D. (1979) Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. Journal of the Acoustical Society of America, 66(2), 470-9.

Sayers, B. McA., and Cherry, E.C. (1957) Mechanism of binaural fusion in the hearing of speech. Journal of the Acoustical Society of America. 29, 973-987.

Schreiner, C.E. (1991) Functional topographies in the primary auditory cortex of the cat. Acta Otolaryngol (Stockh) Supplement. 491, 7-16.

Schreiner, C.E., and Langner, G. (1997) Laminar fine structure of frequency organization in auditory midbrain. Nature, 388(6640), 383-6.

Schreiner, C.E. and Mendelson, J.R. (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. Journal of Neurophysiology. 64, 1442-1460.

Schreiner, C.E., Mendelson, J.R., and Sutter, M.L. (1992) Functional topography of cat primary auditory cortex: representation of tone intensity. Experimental Brain Research. 92, 105-122.

Schreiner, C.E., and Sutter, M.L. (1992) Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. Journal of Neurophysiology. 68, 1487-1502.

Sever, J.C. Jr, and Small, A.M. Jr. (1979) Binaural critical masking bands. Journal of the Acoustical Society of America, 66(5), 1343-50.

Sharf, B. (1970) Critical bands. In J.V. Tobias (ed.) Foundations of modern auditory theory, Vol.1. New York: Academic Press.

Silkes, S.M., and Geisler, C.D. (1991) Responses of "lower-spontaneous-rate" auditory-nerve fibers to speech syllables presented in noise. I: General characteristics. Journal of the Acoustical Society of America, 90(6), 3122-39.

Sinex, D.G., and Geisler, C.D. (1983) Responses of auditory-nerve fibers to consonant-vowel syllables. Journal of the Acoustical Society of America, 73(2), 602-15.

Sinnott, J.M., and Brown, C.H. (1993) Effects of varying signal and noise levels on pure-tone frequency discrimination in humans and monkeys. Journal of the Acoustical Society of America, 93(3), 1535-40.

Smith, R.L. (1979) Adaptation, saturation, and physiological masking in single auditory-nerve fibers. Journal of the Acoustical Society of America, 65(1), 166-78.

Stager, CL, and Werker, JF. (1997) Infants listen for more phonetic detail in speech perception than in word-learning tasks. Nature, 388(6640):381-2.

Steinschneider, M., Arezzo, J., and Vaughan, H.G. Jr. (1982) Speech evoked activity in the auditory radiations and cortex of the awake monkey. Brain Research, 252(2), 353-65.

Steinschneider, M., Arezzo, J.C., Vaughan, H.G. Jr. (1990) Tonotopic features of speech-evoked activity in primate auditory cortex. Brain Research, 519(1-2), 158-68.

Sutter, M.L., and Schreiner, C.E. (1995) Topography of intensity tuning in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. Journal of Neurophysiology. 73(1), 190-204.

Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S.S., Schreiner, C., Jenkins, W.M., and Merzenich, M.M. (1996) Language comprehension in language-learning impaired children improved with acoustically modified speech. Science, 271(5245):81-4.

Tallal, P., Mier, S., Fitch, R.H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. Annual New York Academy of Science. 682, 27-47.

Teas, D.C. and Kiang, N.Y-S. (1964) Evoked responses from the auditory cortex. Experimental Neurology. 10, 91-119.

Tunturi, A.R. (1950) Physiological determination of the arrangement of the afferent connections to the middle ectosylvian auditory area in the dog. American Journal of Physiology. 162, 489-502.

Tunturi, A.R. (1952) A difference in the representation of auditory signals from the left and right ears in the isofrequency contours of the right middle ectosylvian auditory cortex of the dog. American Journal of Physiology. 168, 712-727.

van den Brink, W.A., and Houtgast, T. (1990) Spectro-temporal integration in signal detection. Journal of the Acoustical Society of America, 88(4), 1703-11.

Wang, X., Merzenich, M.M., Beitel, R., and Schreiner, C.E. (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. Journal of Neurophysiology, 74(6), 2685-706.

Webster, F.A. (1951) The influence of interaural phase on masked thresholds. I. The role of interaural time-deviation. Journal of the Acoustical Society of America. 23, 452-462.

Werker, JF, and Tees, RC. (1984) Phonemic and phonetic factors in adult cross-language speech perception. Journal of the Acoustical Society of America, 75(6):1866-78.

Yin, T.C.T., Chan, J.C.K., and Carney, L.H. (1987) Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. Journal of Neurophysiology. 58, 562-583.

Yin, T.C., and Chan, J.C. (1990) Interaural time sensitivity in medial superior olive of cat. Journal of Neurophysiology, 64(2), 465-88.

Yin, T.C., and Kuwada, S. (1983) Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. Journal of Neurophysiology, 50(4), 1020-42.

Young, E.D., and Sachs, M.B. (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. Journal of the Acoustical Society of America, 66(5), 1381-1403.

Zwicker, E. (1961) Subdivision of the audible frequency range into critical bands. Journal of the Acoustical Society of America. 33, 248.