Information Processing in the Interaural Time Difference Pathway of the Barn Owl

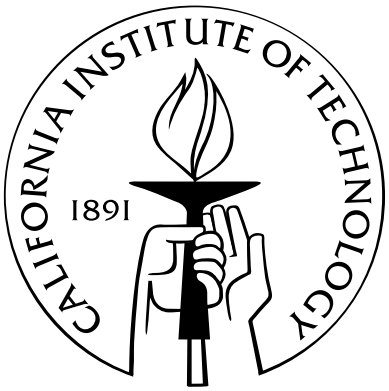
Thesis by

Gestur Björn Christianson

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2006

(Defended November 18, 2005)

To my parents and grandparents.

Acknowledgements

Thanks are due first and foremost to my advisor, Masakazu Konishi, without whom this work would never have been done. Also I must thank my collaborator, José Luis Peña, who recorded all of the NL data shown in this work. All members of the Konishi lab with whom I have had the pleasure of working have contributed in some manner to my research; I would like to particularly thank Ben Arthur, Brian Fischer, Sharad Shanbhag, and Teresa Nick, as well as Roian Egnor for a key piece of advice I wish I had heeded sooner.

I would also like to thank my committee members, Ralph Adolphs, Gilles Laurent, Jerry Pine, Shinsuke Shimojo, and Erik Winfree. Invaluable technical assistance was provided throughout this work by Chris Malek and Gene Akutagawa, as well as all the staff of the Office of Laboratory Animal Resources.

Peter Dayan, Peter Latham, Jennifer Linden, Richard Lyon, Ofer Mazor, Markus Meister, Peter Narins, and Maneesh Sahani have all provided me with helpful discussions on this material.

The work done in this research was supported by National Institutes of Health (grant DC00134) and by the National Sciences and Engineering Research Council of Canada (PGSA-208174-1998 and PGSB-208174-2000).

Abstract

The interaural time difference (ITD) is one of two primary binaural cues used to compute the position of a sound source in space. In the barn owl, the ITD is processed in a dedicated neural pathway that terminates at the core area of the central nucleus of the inferior colliculus (ICcc). The actual locus of the computation of the ITD is prior to ICcc in the nucleus laminaris (NL), and ICcc receives no feed-forward inputs carrying information that did not originate in NL. Here, we compare single-unit responses in these two nuclei. The neurons of both nuclei encode spectrotemporal properties of the stimulus with high resolution in both frequency and time domains, and their response to ITD cues has a spectral profile given by the square of their spectral tuning, a relationship consistent with the prediction that these neurons behave as if computing the cross-correlation of the signals at the two ears. However, the ITD tuning curves of ICcc neurons develop more rapidly than those of the neurons of NL, requiring as little as a single stimulus presentation per ITD to show coherent ITD tuning. ICcc neurons also display a greater dynamic range, with a maximal difference in firing rates due to ITD cues approximately double that seen in NL. These results suggest that ICcc neurons sum across a population of similarly tuned NL neurons to produce an averaged response with greater single-unit reliability.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The task of the nervous system can be grossly summarized as a three-step process: to encode information about the environment; to make decisions based on that encoded information about future actions; and to implement those actions. The first of these tasks is the duty of the sensory system, and the challenges it faces are remarkable. The world around us is vast and ever changing. Even taking into account the inherent limitations of our sensory organs, we are bombarded by enormous amounts of information. The sensory systems of the nervous system must first take in this information and then somehow reduce it to a manageable level, by efficient coding, discarding of irrelevant information, or combinations of the above. To study a sensory system is thus to attempt to answer the question of what information about a stimulus is encoded, and how.

Before this question can be tackled, however, the researcher must first be able to describe and analyze the stimuli he plans to use. From this perspective, hearing is a modality that offers considerable advantage. A sound is simply a time-dependent pressure waveform, and hence one-dimensional. Because of this low dimensionality, sounds can be characterized using the large body of mathematical work that has been done on one-dimensional signal analysis, permitting simple mathematical descriptions of complex stimuli. The ability to describe and manipulate complex properties of the signal in turn allows us to easily explore the dependence of neural response on even high-order characteristics of the signal.

Additionally, the tractability of auditory analysis makes it simpler to express precise mathematical models for the behavior of the related neural networks. While this does not address the question

of how neurons might implement these computations, or indeed whether the models are followed, and to what degree, the existence of a clear theoretical framework is an invaluable tool for the design of experiments.

There are a number of perceptual tasks in which audition plays a role. This research deals with the problem of sound localization. In particular, we focus on the neural processing of the interaural time difference between its initial computation and its convergence with the interaural level difference. Before that discussion begins, it is first necessary to provide background detail on the general nature of sound localization, as well as previous research on the neural mechanisms for its solution within the barn owl, the model used by this study.

1.1 Overview of sound localization

Audition is not an inherently spatial sense. Even though the responses of the primary receptors for both vision and audition are functions of stimulus intensity, the photoreceptors of the eye are arranged *spatiotopically*, in a two-dimensional map of space. Conversely, the hair cells of the ear are arranged *tonotopically*, arrayed across the cochlea as a map of frequency of the stimulus (von Bekesy 1960; Ruggero et al. 1997). Despite the lack of an explicit representation of space in the initial encoding of auditory information, animals and humans alike are capable of identifying the spatial source of a sound. The ability of animals to extract the position of a single source from auditory information represents a computational challenge that is simultaneously complex and well defined. The only inputs to the calculation are the two one-dimensional pressure waveforms that correspond to the sound at each ear, and the results of the computation can be reported using a single physical movement, usually either an eye or head saccade, or a pointing motion. The computations underlying sound localization have been studied for over two centuries using behavioral and physiological methods (Venturi 1796; Rayleigh 1876; Rayleigh 1907).

To understand some of the behavioral implications of sound localization, it is informative to compare it to echolocation. In echolocation, the animal emits a pulse of high frequency, and then analyzes the returning echoes to detect, characterize, and locate objects in the environment (Popper

and Fay 1995). Since it is an active system, information can be extracted by comparing the returned signal to the original signal. This means echolocation can make use of sophisticated cues such as the Doppler shift, where changes in the spectrum of the emitted pulse are caused by the relative velocities between animal and detected object (Müller and Schnitzler 1999). In addition, because the echolocating animal is the one that creates the sound, it can use a sound optimally suited to the environmental conditions (Schnitzler and Kalko 2001). Conversely, sound localization is a passive system, with no prior knowledge of or control over the sound to be localized. As a result, it is theoretically less precise than echolocation. From an ethological perspective, however, passive sound localization offers some advantages. Echolocation is energetically expensive (Speakman, Anderson and Racey 1989), while sound localization requires no additional energy expenditure during the search for prey. Additionally, prey can develop countermeasures to try to counteract echolocation (Greenfield and Baker 2003; Waters 2003), while there is no way to avoid passive sound localization other than the relatively unhelpful strategy of "don't make noise."

1.2 Sound localization cues

There are several cues present that can be used to determine the position of a sound (Blauert 1997). Primary among these are the binaural cues, arising from the differences in the sounds at the two ears (fig. 1.1), the use of which were first demonstrated by Rayleigh (1907).

The interaural level difference (ILD; sometimes used interchangeably with the interaural intensity difference, or IID, though the two terms are technically different) arises due to occlusion of the sound by the head at the far ear. For a sound source to the right of the animal, the head will cast an acoustic "shadow" over the left ear, reducing the sound intensity there (fig. 1.1). This effect is not uniform across frequencies. For low frequencies, diffraction causes the incoming waveform to propagate past the head with relatively little reduction in intensity (Shaw 1974). As a consequence, ILD cues are theoretically better suited to high-frequency signals.

The interaural time difference (ITD) cue arises because of the difference in path lengths from the sound source to the two ears. A sound to the right of the listener will arrive at the right ear before

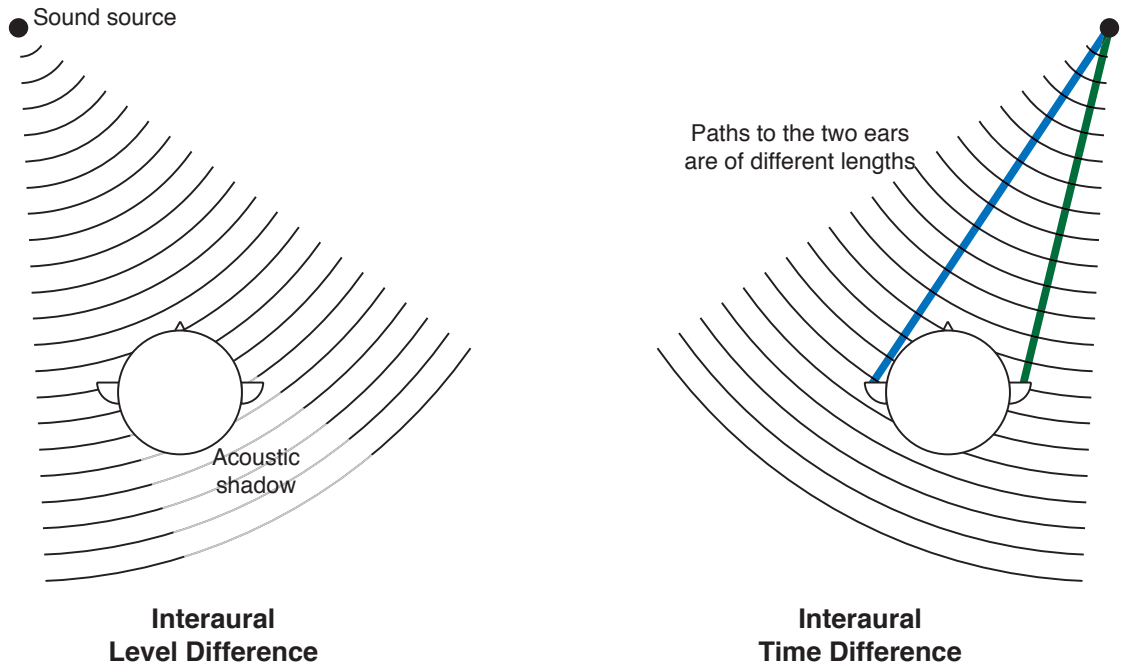


Figure 1.1. An illustration of binaural cues for the purposes of sound localization.

it arrives at the left (fig. 1.1), resulting in an ongoing temporal disparity for that sound between the copies at the two ears. The maximum ITD an animal can perceive is limited by the size of the head, as the maximal difference in path lengths from sound source to ears is given by the interaural separation. In general, this maximum ITD is quite small, on the order of about 630 μs for the human (Blauert 1997) and 180 μs for the owl (Moiseff and Konishi 1981). Thus, to use the ITD as a sound localization cue, the animal must be able to detect interaural disparities on a very short time scale. For reasons we will discuss in detail in section 1.5, this means that the computation of ITD is generally easier at low frequencies.

In addition to these binaural cues, there are monaural cues. The pinna serves to reflect incoming sound towards the inner ear. This reflection will create interference, which will be dependent on the frequency of the stimulus, as well as its incoming direction, and hence location (Blauert 1997). In addition, many animals have motile outer ears, and presumably can use changes in the perceived sound due to movement of the ear to relay information relating to the sound position (Populin and

Yin 1998). However, the evidence is that motion-based cues from the movement of the head in humans (whose outer ears are fixed) contribute, at best, weakly to localization performance (Pollack and Rose 1967; Thurlow and Runge 1967).

In the majority of systems that have been studied, including all mammalian systems, the symmetry of the placement of the ears on the head causes both the ITD and the ILD to encode position along the horizontal plane, with ITDs being used for low frequencies and ILDs being used for high frequencies. This interaction is referred to as the duplex theory. In humans, psychophysics confirms that monaural cues are used to determine elevation (Middlebrooks and Green 1991).

1.3 The barn owl

The barn owl (*Tyto alba*) is a nocturnal predator. Despite a well-developed visual system (Pettigrew and Konishi 1976; van de Willigen, Frost and Wagner 1998), light levels in the wild often fall below the levels at which the owl can see (Dice 1945; Curtis 1952), indicating that owls must use some non-visual sense to hunt. It was hypothesized that owls might have vision extending into the infra-red spectrum (Vanderplank 1994), though this was rapidly disproven (Matthews and Matthews 1939; Hecht and Pirenne 1940). Dice (1945) noted that in darkness owls searching for prey would occasionally touch dead mice without apparently noticing it, arguing against a role for olfaction in hunting. Similarly, Curtis (1952) found a direct correlation between light intensity and the ability of an owl to navigate barriers in its flight path, which made it unlikely that the source of the owl's hunting ability lay in either echolocation or the ability to perceive infra-red radiation.

It was Payne and Drury (1958) who first demonstrated the ability of the barn owl to locate and catch prey using only passive auditory cues. The barn owl's sound localization ability is impressive. It can localize sounds which are as short as 10 ms in duration (Konishi 1973), considerably less than the latency for the initiation of the head saccade (Knudsen, Blasdel and Konishi 1979). Additionally, it can resolve targets with less than 2° error in both azimuth and elevation (Knudsen, Blasdel and Konishi 1979), compared to 2° error in humans (Middlebrooks and Green 1991) and 4° error in the macaque (Brown et al. 1980; Brown et al. 1982).

Motivated by these reports, researchers began to explore the neural mechanisms that underlie the barn owl's localization ability. It was found that the position of sounds are represented in an explicit, spatially organized map in the external nucleus of the inferior colliculus (ICx) within the owl's midbrain (Knudsen and Konishi 1978a, 1978b), providing a direct correlate between behavioral reports and neurophysiological responses. While the superior colliculus of cat (Middlebrooks and Knudsen 1984), ferret (King and Hutchings 1987), and guinea pig (Sterbing, Hartung and Hoffmann 2002) have been reported to have maps of auditory space in register with a visual map of space, it is only in the barn owl that an exclusively auditory map of space, such as that seen in ICx, has been reported. This map, unconfounded by visual cues and providing a neural correlate to the localization behaviors, makes the owl an excellent model for the study of the mechanisms of sound localization.

There are several crucial differences between the localization mechanisms of the barn owl and mammalian species, including the human. The range of frequencies that the owl can perceive extends only up to about 12 kHz (Konishi 1973), while the hearing range of humans extends to about 20 kHz. Within this range, the owl can localize well using frequencies from 1 to 8 kHz, and preferentially uses frequencies in the comparatively narrow range of 4 to 8 kHz (Knudsen and Konishi 1979); conversely, mammals appear to use the entirety of their range for localization. In the barn owl, an asymmetry in the owl's outer ears (Norberg 1977) results in ILD coding for elevation (Knudsen and Konishi 1979), while monaural cues are not used (Egnor 2000); the preference for high frequencies ensures that diffraction of the sound about the head does not have a major influence on the use of the ILD cue. As a consequence of ILD coding for elevation, the owl must use ITD to encode azimuth throughout its sensitive frequency range (Knudsen and Konishi 1979), while most other animals only rely on ITD for frequencies less than 2 kHz.

While these differences make it difficult to use the owl as a model of the human process of sound localization, they argue in favor of the owl as the ideal model for understanding the computational challenges involved in solving the sound localization problem in a biological system. The owl's resolution of passive auditory spatial cues is unmatched by any other reported terrestrial species (Knudsen, Blasdel and Konishi 1979), and a direct comparison can be made between the performance

of the neurons of the midbrain and the behavioral capabilities. In addition, the use of the two binaural cues to resolve orthogonal directions, the lack of use of the monaural cues, and the absence of mobile pinnae greatly reduces the number of variables to be considered in analyzing the system. With the understanding that comes from studying such a specialized system, it becomes easier to recognize optimizations and trade-offs that will be made by non-auditory specialist species.

1.4 Anatomy of the sound localization pathway

The organization of the neural structures that precede ICx with its spatial map have been studied extensively. A primary result of this exploration was the discovery of parallel pathways for the processing of ITD and ILD (fig. 1.2; Sullivan and Konishi 1984; Takahashi and Konishi 1988a, 1988b). As each auditory nerve fiber enters the brain, it bifurcates to terminate in both of the cochlear nuclei, nucleus magnocellularis (NM) and nucleus angularis (NA) (Carr and Boudreau 1991). Neurons of nucleus magnocellularis encode the phase of tonal stimuli for frequencies up to 9 kHz, and have reduced variation in firing rate as a function of stimulus intensity compared to the auditory nerve fibres (Sullivan and Konishi 1984; Köppl 1997b). Conversely, neurons of nucleus angularis do not phase-lock, and their dynamic ranges in response to changes in stimulus intensity are very large (Sullivan and Konishi 1984). Thus, the two nuclei are specialized to process phase or time information and intensity information, respectively.

NA projects contralaterally to the nucleus dorsal lemnisci lateralis pars posterior (LLDp; previously known as the nucleus ventralis lemnisci lateralis pars posterior, VLVp; Takahashi and Konishi 1988a). This projection is excitatory, and combined with an inhibitory input from the contralateral LLDp (Takahashi and Keller 1992) produces a sigmoidal tuning to ILD in LLDp, with a preference for contralaterally-dominated ILDs (Moiseff and Konishi 1983; Manley, Köppl and Konishi 1988). NA also has projections to the superior olive (SO; Takahashi and Konishi 1988a), the nucleus ventral lemnisci lateralis (LLv, not shown on fig. 1.2; Takahashi and Konishi 1988a), and the lateral shell of the central nucleus of the inferior colliculus (ICcl; Takahashi and Konishi 1988b).

NM projects bilaterally and solely to nucleus laminaris (NL; Takahashi and Konishi 1988a). NL

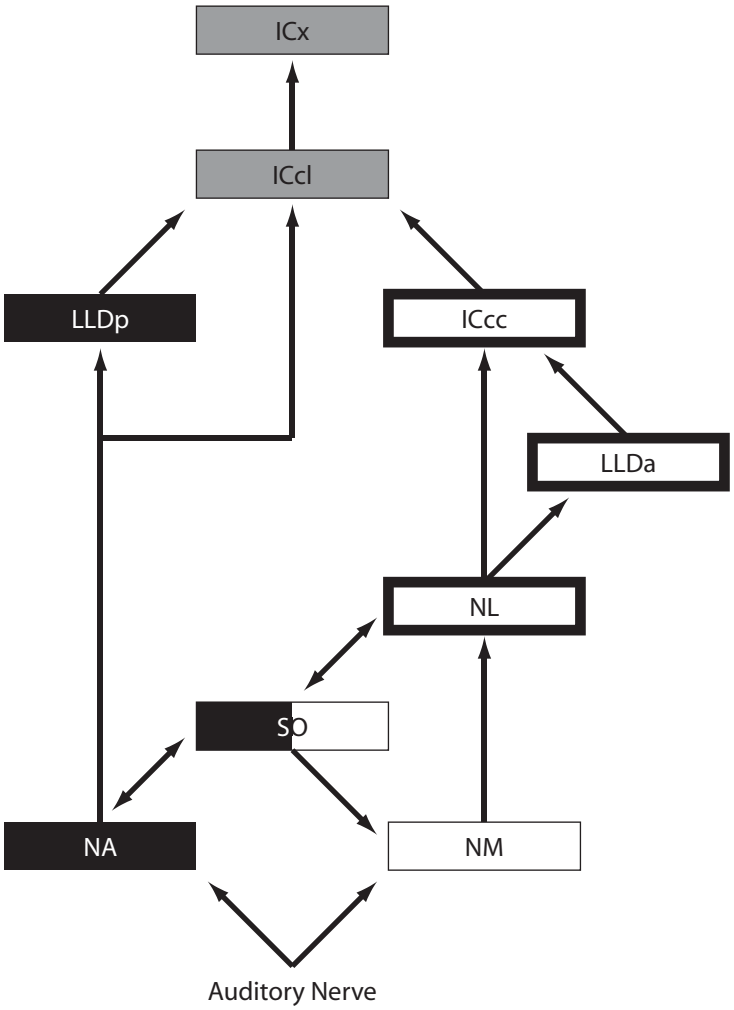


Figure 1.2. A simplified schematic of the sound localization pathway. The ILD pathway is shown in black, the ITD pathway in white, and the sites in which the cues converge are grey. The nuclei of the closed loop that is the focus of this thesis are emphasized with the heavy black border.

is sensitive to the ITD of the stimulus (Moiseff and Konishi 1983; Carr and Konishi 1990), organized by tonotopy and by ITD (Sullivan and Konishi 1986; Carr and Konishi 1990), and the tuning to ITD is largely unaffected by the ILD of the stimulus (Viete, Peña and Konishi 1997). At this stage, the ITD tuning is ambiguous: for any given firing rate, there are in general multiple ITDs that will result in that firing rate, regardless of the spectral properties of the stimulus (a point discussed in further detail in section 1.5). NL in turn has three projections: to SO (Takahashi and Konishi 1988a), to the nucleus dorsal lemnisci pars anterior (LLDa; previously known as the nucleus ventralis lemnisci lateralis pars anterior, VLVa; Takahashi and Konishi 1988a), and to the core region of the central

nucleus of the inferior colliculus (ICcc; Takahashi and Konishi 1988b).

SO receives excitatory projections from NA and NL (Takahashi and Konishi 1988b; Lachica, Rübsamen and Rubel 1994). The neurons of SO do not encode phase, and are relatively insensitive to binaural cues (Moiseff and Konishi 1983). GABAergic projections from SO to NL and to NM (Lachica, Rübsamen and Rubel 1994) have led to the hypothesis that SO serves to eliminate the effect of stimulus intensity on the tuning of NL neurons (Lachica, Rübsamen and Rubel 1994; Peña et al. 1996; Dasika et al. 2005).

Following NL in the ITD pathway, LLDa in turn projects to ICcc (Moiseff and Konishi 1983) and to the nucleus basalis (Wild, Kubke and Carr 2001), which is outside our consideration. There are no reports that LLDa receives inputs other than from NL. ICcc then projects to the ICcl (Takahashi, Wagner and Konishi 1989), as well as to nucleus ovoidalis in the thalamus (Proctor and Konishi 1997; Cohen, Miller and Knudsen 1998). Both LLDa (Moiseff and Konishi 1983; Albeck and Konishi 1995) and ICcc (Wagner, Takashi and Konishi 1987, 2002) are tuned to ITD, with a lack of sensitivity to intensity or ILD on a par with the neurons of NL, and their ITD responses are still ambiguous.

As can be seen from this discussion, ITD (fig. 1.2, white boxes) and ILD (fig. 1.2, black boxes) are processed independently in the first few stages of the ascending auditory pathway. ICcl, which receives inputs from both LLDp and ICcc, is the locus of the combination of the two cues (fig. 1.2. grey boxes). Neurons of the ICcl are sensitive to combinations of ILD and ITD, but are not necessarily space-specific (Feldman and Knudsen 1994; Mazer 1995; Keller and Takahashi 2000). ITD tuning may still be ambiguous, and ILD tuning may still be sigmoidal. Neurons throughout ICcl are believed to present a gradation which moves towards space-specific neurons, in which tuning to both ITD and ILD produces a single unambiguous peak (Mazer 1995). ICx is the target of an ICcl projection (Knudsen and Knudsen 1983), and this is the first truly space-specific nucleus in the auditory pathway (Knudsen and Konishi 1978b, 1978a). In ICx, neurons are not only space-sensitive, but organized spatiotopically. ICx in turn projects to the optic tectum (OT), where the auditory spatiotopic map is preserved and put into register with a visual map of space (Knudsen 1983; Knudsen and Knudsen 1985; Knudsen and Brainard 1991).

There is a second continuation of the localization pathway other than the tectal one described above, commonly referred to as the forebrain pathway. Both ICcl and ICcc project to NO, while ICx does not (Knudsen and Knudsen 1983; Proctor and Konishi 1997; Cohen, Miller and Knudsen 1998; Arthur 2005). From NO, there is a projection through Field L to the arcopallium (previously called archistriatum; Cohen, Miller and Knudsen 1998), where space-specific neurons can also be found (Cohen and Knudsen 1995). While the space-specific neurons of arcopallium do not have a spatiotopic map, microstimulation in either arcopallium or OT will result in head saccades (du Lac and Knudsen 1990; Masino and Knudsen 1990; Knudsen, Cohen and Masino 1995).

In this research, our interest lies in issues relating to the computation of ITD. As such, the nuclei of interest are those highlighted in figure 1.2. While the mammalian sound localization pathway lacks the clear segregation of that of the barn owl, and there is yet to be identified any equivalents to ICcl or the forebrain or tectal pathway, mammalian homologues for NM (the anteroventral cochlear nucleus, or AVCN), NL (the medial superior olive, or MSO), and ICcc (the central nucleus of the inferior colliculus, or ICc[1]), as well as a more tentative homology for LLDa (the dorsal nucleus of the lateral lemniscus, or DNLL) have been identified.

1.5 Computation of the interaural time difference

Jeffress (1948) was the first to formulate the coincidence detection model for the computation of ITDs. In the Jeffress model, a series of neurons are connected bilaterally to the two ears with calibrated delay lines. The input spikes to these neurons encode timing properties of the auditory stimulus, and the neurons fire only when they receive coincident inputs from both sides. In this arrangement, the only coincidence detector that will fire is the one where the difference in propagation times for the delay lines from either side precisely balances the ITD, resulting in a place code of ITD (fig. 1.3). Thus, there are three requirements that must be met for the auditory system to implement the Jeffress model: there must be encoding of the ongoing time (or, equivalently, phase)

[1] "Core" and "lateral shell" subdivisions of the central nucleus of the inferior colliculus have not been identified in mammalian models. However, from the perspective of the computation of ITD the entirety of the mammal's central nucleus (ICc) behaves in the same manner as the core area of the owl's central nucleus (ICcc), as there is no known locus of convergence of ILD and ITD cues in mammals.

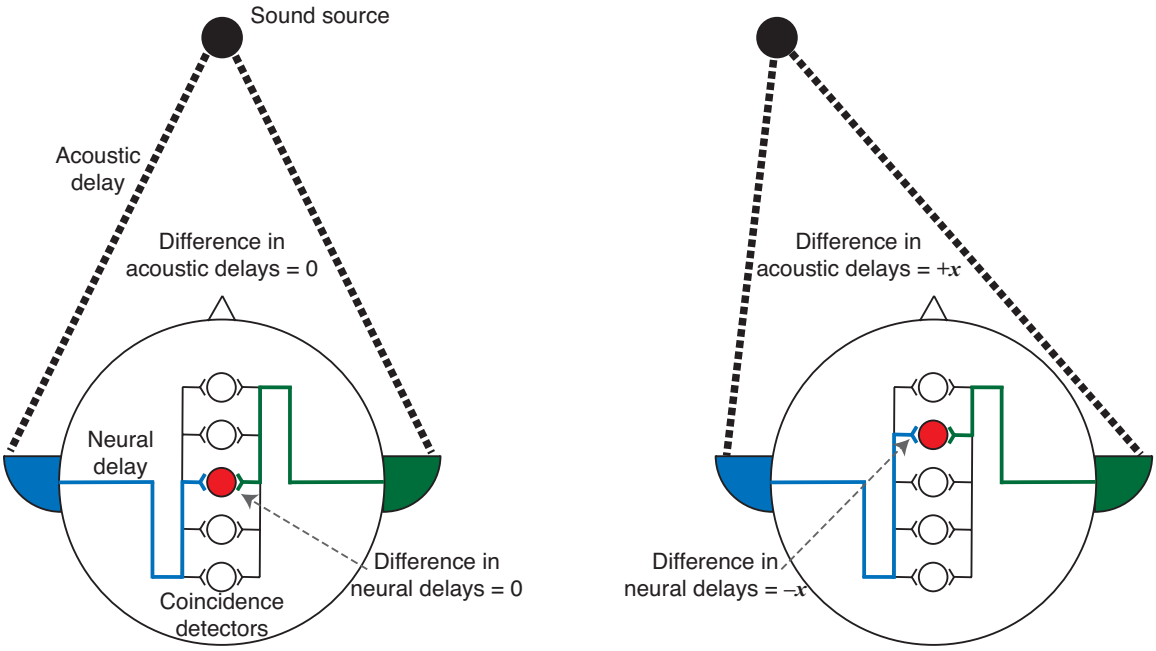


Figure 1.3. A schematic of the Jeffress model. Coincidence detectors are connected by a series of delay lines to the two ears; the delay lines are arranged such that the difference in the neural delays from either ear varies across the family of coincidence detectors. When the sound source is equidistant from the two ears, then only the coincidence detector whose neural delay lines are of the same length will receive inputs at the same time, and hence be active (*left*). When the sound source moves, so that there is an ITD of $+x$, only the coincidence detector whose neural delay lines have the same absolute difference but reversed in sign will receive coincident input (*right*).

of the stimulus, there must be neurons that can operate as coincidence detectors, and the inputs to the coincidence detecting neurons from the two sides must be delayed with respect to each other in a manner that is systematic across the neuronal population.

Recordings in the auditory nerve and NM reveal a tendency of these neurons, when stimulated with a tone, to fire preferentially for a particular phase of the input signal (fig. 1.4; Sullivan and Konishi 1984, Köppl 1997b). This provides in a straightforward manner an encoding of the ongoing time of the tone: each spike signals the reoccurrence of the preferred phase.

Two notes should be made in regards to phase-locking. Firstly, frequency and time are intrinsically related. For example, a jitter of 90° in phase with a stimulating frequency of 1 kHz means that spikes must occur within 250 μs of the true preferred phase. However, for a stimulating frequency of 5 kHz, that same phase jitter requires a temporal resolution of 50 μs. Neurons are generally thought to operate on timescales measured on the order of milliseconds. Thus, even encoding and

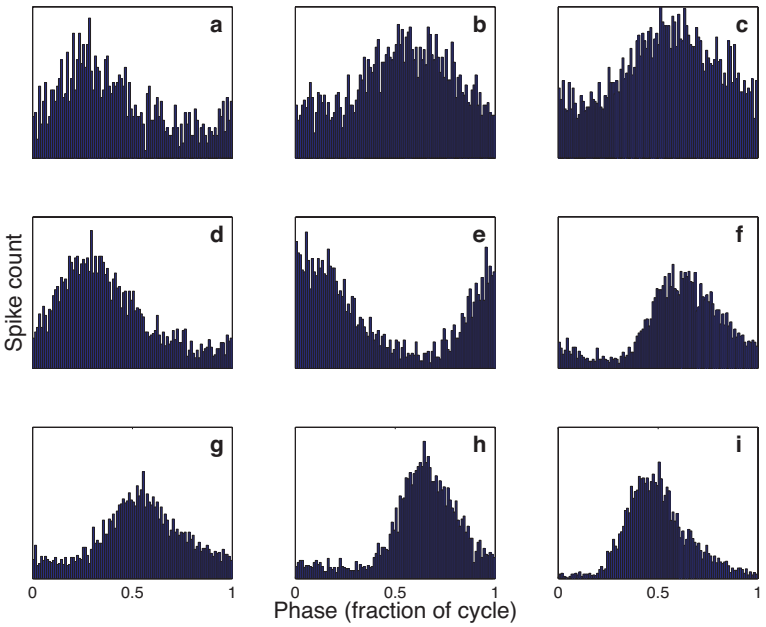


Figure 1.4. Examples of phase-locking in the auditory nerve of the barn owl. Each fiber was stimulated with a tone of its preferred frequency. Spike times were converted to phase relative to the stimulus, and then a histogram of these phases was plotted. Note that while a peak in the phase histogram is apparent in all cases, it is also true that spikes occur at all phases, with this being more pronounced for higher-frequency neurons (***a***: unit 650.01, 7,500 Hz; ***b***: unit 650.02, 7,750 Hz; ***c***: unit 650.03, 7,500 Hz; ***d***: unit 650.08, 5,750 Hz; ***e***: unit 650.12, 5,500 Hz; ***f***: unit 650.14, 4,750 Hz; ***g***: unit 650.18, 4,500 Hz; ***h***: unit 650.19, 3,500 Hz; ***i***: unit 650.20, 2,250 Hz).

detecting time differences of frequencies below 1 kHz push the limits of what can be done with neural hardware, and it is for this reason that most animals use ITDs only for low-frequency signals. The owl can phase-lock over an enormous frequency range, at frequencies up to 9 kHz (Köppl 1997b). At the same time, there is still a decline in the quality of phase-locking with increasing frequency (fig. 1.4).

We also observe that phase-locking is strictly only defined in the case of tonal stimulation. With more complex stimuli, "phase" is not a well-defined concept. Even for a stimulus as simple as the simultaneous presentation of two tones, "phase" can easily refer to the phase of either of the component tones, or the phase of the amplitude envelope of the sum. The analogue to phase-locking with complex stimuli is strictly speaking the existence of a spike-triggered average (also called a revcorr function; de Boer and de Jongh 1978). However, "phase-locking" is generally used to refer

to ongoing time encoding of stimuli regardless of the spectral properties of the stimuli. This is partly convention, but is validated by the bandpass properties of the cochlea. Neurons throughout the ITD pathway, and the auditory nerve and NM in particular, are responsive to only a very narrow range of frequencies. Even white noise is thus effectively narrowband noise from the perspective of a single unit, and narrowband noise shares many of the features of tones. In particular, periodic stimuli such as tones are *phase ambiguous*: that is, without reference to an absolute end-point, it is impossible to distinguish between any given periodic signal and that same signal shifted in time by integer multiples of the period. Narrowband noise is not periodic; however, for small time displacements, it can be ambiguous in a manner similar to that seen in tones (fig. 1.5).

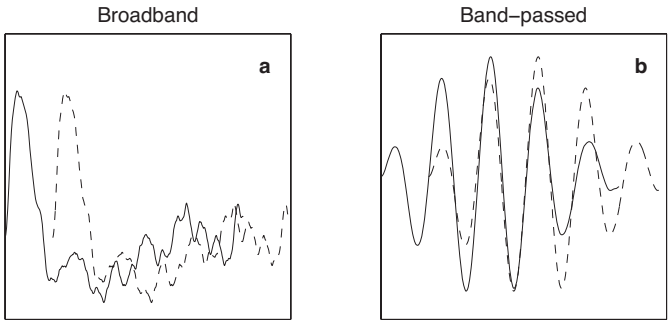


Figure 1.5. ***a***: A 1 ms segment of broadband noise (1–10 kHz) is plotted (solid line), and then shifted by 200 μs and plotted again (dashed line). Note that at any given time, the curves have very little relationship to each other. ***b***: The noise segment as used in ***a*** has been band-passed to have a 1 kHz bandwidth centered on 5 kHz, and plotted as before. The shift of 200 μs corresponds to the period of the center of the band-pass filter. It can be seen that there is a great deal of correspondence between the time-shifted versions now, though this correspondence will fall off as larger time shifts are used.

Having established that NM encodes time, it is then necessary to demonstrate that the projections of NM to NL form the requisite delay lines. Carr and Konishi (1988) showed that the ipsilateral and contralateral NM afferents enter from the dorsal and ventral surfaces of NL, respectively. They then pass perpendicular to these surfaces through NL, and interdigitate within the nucleus. By recording intracellularly from the NM fibers, they were able to show that there was an orderly progression of conduction delays within NL. The conclusion is that the magnocellular afferents act as the delay lines of the Jeffress model, setting up an orderly progression of conduction delay differences through the dorso-ventral extent of NL.

Finally, Carr and Konishi (1990) observed phase-locking in NL neurons. If NL neurons are coincidence detectors, this is expected: the output spikes of NL will be directly related to the timing of its phase-locked inputs, and hence will be phase-locked themselves. NL neurons are driven by monaural stimulation as well as binaural stimulation, and the monaural response is also phase-locked. The preferred phase in monaural stimulation is given by the modulus of the conduction delay by the period of the stimulating frequency. Thus, the difference in preferred phase between ipsilateral and contralateral stimulation is equal to the true difference in conduction delays (called the characteristic delay, or CD) up to some integer multiple of the period of the frequency. So therefore, in a coincidence detector, presenting tones with an ITD that is equal to the difference in monaural phases but reversed in sign should elicit a maximal response; using an ITD that corresponds to a 180° shift in phase should result in a minimal response. This turns out to be the case (Carr and Konishi 1990), confirming that NL neurons act as coincidence detectors. Thus, all elements are present, and the Jeffress model stands as the description for how ITD is computed in the barn owl's auditory pathway.

Jeffress' model was not entirely accurate, however. In his work, Jeffress assumed that there was a perfect encoding of the complete signal. As we have already discussed, however, the ITD pathway encodes the temporal properties of the signal within narrow frequency bands, rather than as a whole, and that this encoding contains a not insignificant amount of noise. The result of the noise, combined with the effects of phase ambiguity, is that individual NL neurons do not signal the presence of a particular ITD in a binary fashion; instead, they respond in a graded manner, going from maximal firing rate at CD to a minimal firing rate at ITDs which are out of phase. The result of the narrow frequency bands is that all stimuli behave essentially like tones, and are thus phase ambiguous. Given the range of ITDs the owl can experience, this means that for frequencies above 4 kHz any given firing rate for a particular NL neuron can be elicited by at least two distinct ITDs (fig. 4.1). The existence of a graded response is not a problem; in fact, it can be argued from theoretical grounds that it is actually advantageous (Harper and McAlpine 2004). The phase ambiguity, however, is a problem that the localization pathway must address.

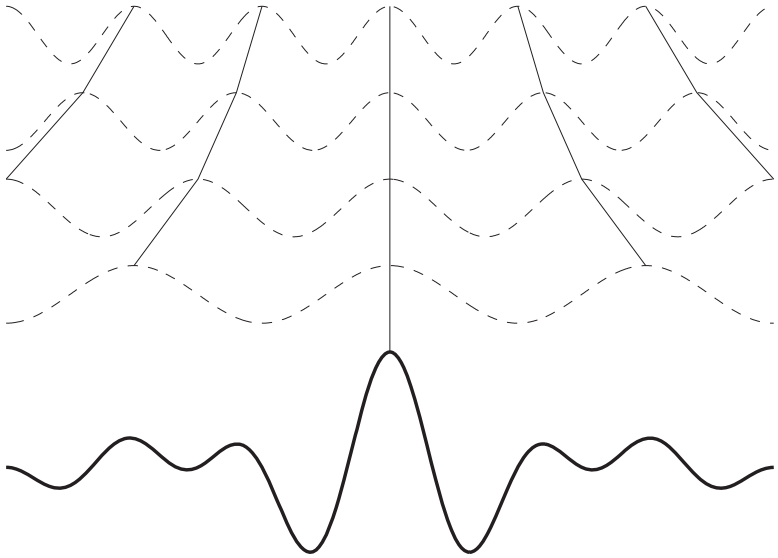


Figure 1.6. For a family of phase-ambiguous signals of different frequencies (*dashed lines*) but with one peak (the CD) in register, the side peaks will be out of phase with respect to each other (illustrated with the *thin solid lines*). The sum of these curves (*heavy solid line*) will have a maximal peak at the time of the CD alone.

The basic mechanism by which this ambiguity is resolved is frequency convergence (Mazer 1998; Saberi et al. 1999). A phase-ambiguous neuron stimulated with broadband will have a peak response at its CD, and at ITDs that correspond to the CD plus some integer multiple of the period of the center frequency of its associated band-pass filter (fig. 1.5). A neuron of the same CD but different preferred frequency will thus have its secondary peaks at different ITDs. By summing across a population of such neurons with different preferred frequencies, it is possible to reinforce the peak at CD while the side-peaks will destructively interfere with each other, resulting in a single unambiguous peak. In the barn owl, the process of frequency convergence has been shown to begin in ICcl (Wagner, Takashi and Konishi 1987; Mazer 1995, 1998).

To a certain extent, the story within the mammalian literature is not quite as clear. The neurons of MSO, the homologue to NL, have been shown to behave as coincidence detectors (Yin and Chan 1990; Batra, Kuwada and Fitzpatrick 1997), though there has recently been some argument that their coincidence detection is noisier than has previously been thought (Batra and Yin 2004). However, the existence of a systematic organization of delay lines such as the owl's magnocellular afferents

has not been unequivocally demonstrated (Smith, Joris and Yin 1993; Beckius, Batra and Oliver 1999). Related to this, McAlpine, Jiang and Palmer (2001) reported that in the guinea pig, CDs consistently fell outside the physiological range of ITDs as estimated by the interaural separation, though this work was done in ICc, and not MSO; Sterbing, Hartung and Hoffmann (2003) have argued that the observed CDs fall within the range predicted by the head-related transfer functions, as low frequency sounds distort about the head and pinnae, creating a longer path length. However, from the perspective of the computation of ITD, this is not a critical point: the distribution of CDs is of importance in the problem of decoding position from a range of ITD responses across a neuronal population (Harper and McAlpine 2004), and not an issue of how those ITD responses develop. Similarly, because mammals generally use ITD cues for sound localization only at frequencies below 2 kHz, phase ambiguity does not present a serious complication.

1.6 Role of the post-laminaris ITD pathway?

In the previous section, we described how the ITD was computed in NL, and how the problem of phase ambiguity was resolved in ICcl. However, this raises an interesting question. Between NL and ICcl lie the nuclei LLDa and ICcc (fig. 1.2). Both are exclusively tuned to ITD. In fact, the published data on their response properties suggests that their tuning is similar to that of NL (Moiseff and Konishi 1983; Wagner, Takashi and Konishi 1987; Albeck and Konishi 1995; Wagner, Mazer and von Campenhausen 2002). Since frequency convergence does not occur until ICcl (Wagner, Takashi and Konishi 1987; Mazer 1995, 1998), what role does the post-laminaris ITD pathway fill?

The simplest explanation is that it serves as a relay station, but we may dismiss this out of hand. This is not a single nucleus, but two; if the purpose was a relay, then presumably the NL–ICcc projection alone would suffice. In addition to this, ICcc and ICcl are subdivisions of the same nucleus, so that a direct projection from NL to the ICcl would be at best marginally longer, making it unlikely that a relay is required.

This redundancy is also present in the mammalian models (Yin and Chan 1990). In fact, because of the difficulties involved in recording from MSO (Guinan, Norris and Guinan 1972), the similarity

in response properties is regularly exploited to use ICc as a substitute for recording in MSO (for an example, see Yin, Chan and Carney 1987).

The goal of this research is to address this apparent redundancy. Since LLDa's only known contribution to the higher stages of the sound localization pathway is through ICcc, and because it receives no inputs other than those from NL, we can treat the post-laminaris ITD pathway as a black box, whose inputs are the activity of NL neurons and whose output is given by the activity of the neurons of ICcc. Thus, by comparing the response properties of NL and ICcc neurons in depth, we hope to understand the computational role filled by the post-laminaris ITD pathway. In chapter 3, we compare the spectral and temporal tuning of these nuclei using reverse correlation analysis; this is a hitherto unexplored aspect of the ITD pathway, and therefore a logical place to commence. In chapter 4, we use those results to explore the interaction between the spectral tuning and the response to ITD, which gives us a way to quantify and compare the shapes of the ITD tuning curves. Finally in chapter 5, the ITD responses of the two nuclei are contrasted in further depth. Our results not only provide an answer at least in part to the original question, but provide further insight into neural strategies for the coding of information.

Chapter 2

Methods

Methods limited to a particular experiment will be described in the appropriate chapter.

2.1 Surgery

The protocol for this study followed the NIH Guide for the Care and Use of Laboratory Animals and was approved by the Institutes Animal Care and Use Committee.

Data were obtained from 15 adult barn owls (*Tyto alba*) of both sexes. Owls were anesthetized by intramuscular injection of ketamine hydrochloride (20 mg/kg, Ketaject, Phoenix Pharmaceutical) and xylazine (2 mg/kg, Xyla-Ject, Phoenix Pharmaceutical), and injected subcutaneously with 10 ml of lactated ringer's solution (B. Braun medical). An adequate level of anesthesia was maintained by additional injections of both when needed as determined by movement and the blink reflex. In trimming feathers in preparation for surgery, care was taken to ensure that the facial ruff was left intact. During surgery and recording sessions, the owl was restrained with a soft leather jacket and body temperature was maintained using a water-based heating pad. During the first recording session, ear bars and beak holder were used to position the owl's head with the beak rotated 30° in the saggital plane, and the scalp was retracted. A head plate was implanted by removing the top layer of the skull and fixing the head plate onto the tracbeculae using dental cement, and a stainless steel reference post was implanted posterior to the head plate and similarly fixed with dental cement. Once the head plate was implanted, the ear bars and beak holder were removed and the head plate was used to fix the head from that point forward and in all following recording

sessions. The craniotomy was packed with gelfoam and sealed with dental cement, and the scalp was sutured closed. Following the surgery, analgesics (ketoprofen, 10 mg/kg, Ketofen, Merial) and antibiotics (oxytetracycline, 4 mg/kg, Maxim-200, Phoenix Pharmaceutical) were administered. The owl was kept in a heated cage until it recovered, at which point it was returned to the home cage. Weight was monitored from recording session to recording session, and the owl was given at least a week's rest between sessions.

2.2 Electrophysiology

Single neurons of NL were isolated and maintained by a loose patch method in which a glass patch electrode served as a suction electrode, allowing us to hold neurons for a long time. Electrodes were prepared from 1.0 mm borosilicate glass (World Precision Instruments) using a micropipette puller (Sutter Instruments P-87). Electrodes were filled with a patch solution (in mM: K-gluconate 100, EGTA 10, HEPES 40, $MgCl_2$ 5, Na-ATP 2.2, Na-GTP 0.3), and impedance varied from 4 to 10 MΩ. Neural signals were serially amplified by an Axoclamp-2A (Axon Instruments) in the conventional current-clamp bridge mode, and further amplified and filtered with a custom-made device (B.E.S. μM-200). NL neurons were identified stereotaxically and by their response properties: in the owl's brainstem, only NM and NL produce neurophonics, and of these, only NL has ITD tuning.

Single neurons of ICcc were isolated and maintained by extracellular recordings using 1 MΩ tungsten electrodes (A-M Systems, Inc.). Neural signals were amplified by a custom-made AC amplifier (B.E.S. μM-200). ICcc neurons were identified stereotaxically and by their response properties: only four nuclei in the owl's midbrain display phase-ambiguous tuning. LLDa and NL can be ruled out by stereotaxic arguments, and ICcl neurons are tuned to ILD, while ICcc neurons are not; additionally, ICcc neurons are tuned to ipsilateral space, while ICcl is tuned to contralateral positions. In the case of low-frequency neurons, where both phase ambiguity of the ITD response and ILD tuning could be difficult to determine, latency was estimated from the reverse correlation data. If latency could not be established, the neuron was discarded.

In both cases, a spike discriminator (SD1, Tucker Davis) converted neural impulses into TTL pulses for an event timer (ET1, Tucker Davis), which recorded the timing of the pulses. A computer running a custom software program (XDPHYS, written by J. A. Mazer, and modified by B. J. Arthur and C. Malek; available for download at `ftp://ftp.etho.caltech.edu/pub/xdphys`) was used for stimulus synthesis and online data analysis.

All recordings were done in a double-walled soundproof chamber (Industrial Acoustics Company, Inc.).

2.3 Acoustic stimulation

An earphone assembly consisting of a Knowles ED-1914 speaker, a Knowles BF-1743 damping device, and a Knowles EA-1939 microphone delivered sound stimuli. These components are encased in an aluminum cylinder that fits into the owl's ear canal. The gaps between the cylinder and the ear canal were filled with silicon impression material (Gold Velvet II, All American Laboratories). At the beginning of each experimental session, the earphone assemblies were automatically calibrated. The computer was programmed to equalize sound pressure level and phase for all frequencies within the frequency range relevant to the experiment (500–13,000 Hz).

Tonal and broadband stimuli 100 ms in duration and sampled at 48,077 Hz were presented at a rate of approximately twice per second. Broadband stimuli were bandpassed to contain signal only from 500–12,000 Hz or from 1,000–12,000 Hz, depending on the preferred frequency range of the unit under study. Signals were gated at rise and fall with a 5 ms linear ramp to prevent onset effects. We used PA4 digital attenuators (Tucker Davis) to vary stimulus sound levels.

2.4 Data collection

Long-range ITD curves were obtained by scanning in 10 to 30 μs steps from $\pm$1,500 to $\pm$2,500 μs with broadband signals (1–12,000 kHz) each repeated 5–10 times. During a recording session, the characteristic delay (CD) was estimated to be the ITD within the physiological range of response

that elicited the maximal response; when possible, the estimate was confirmed afterwards using a family of tonal ITD tuning curves.

Iso-intensity frequency tuning curves were obtained for sound levels 20 to 30 dB above threshold with randomized sequences of stimulus frequencies in steps of 100 Hz at the CD of the neuron. Frequency tuning curves were characterized by their width (W_{50}), center frequency (F_{50}) and best frequency (BF). W_{50} is the range of frequencies over which the cells discharge rate was equal to 50% of the difference between the maximal discharge rate and the spontaneous level. The frequency at the center of W_{50} was defined as F_{50}. BF is defined as the frequency that elicits the maximal discharge rate in an iso-intensity frequency-tuning curve.

In a single neuron, the sound intensities used for all protocols were the same.

2.5 Modeling

The spiking models in this paper are inhomegenous Poisson processes that follow the implementation given in Zhang et al. (2001). However, for NL and ICcc, we used a different choice of history function. NL neurons have a characteristic inter-spike interval histogram (fig. 2.1) which is not well matched by a sum of exponentials. Instead, a history function based on a hyperbolic tangent was used:

$$H(t) = \begin{cases} -\frac{1}{c_3}\tanh(\pi\frac{t-t_1-R_A-c_0}{c_1}+c_2) & (t-t_1)\geq R_A \\ 1.0 & (t-t_1)<R_A \end{cases}$$

where R_A is the absolute refractory period, t_1 is the time of the last spike, and the c_x are parameters.

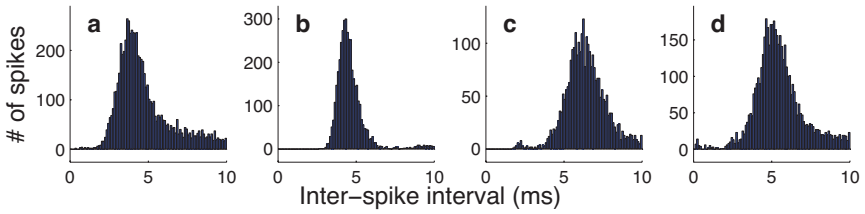


Figure 2.1. Four examples of ISIHs for NL neurons. Note the almost symmetric, gaussian peak.

For NM, a sum of exponentials was used:

$$H(t) = \begin{cases} c_0 e^{\frac{R_A-(t-t_1)}{c_1}} + c_2 e^{\frac{R_A-(t-t_1)}{c_3}} & (t-t_1) \geq R_A \\ 1.0 & (t-t_1) < R_A \end{cases}$$

though strictly speaking, c_2 was set to zero, and thus a single exponential was used.

The parameter values were chosen by hand to provide a reasonable approximation to the appropriate ISIH, though an inhomegenous Poisson process is incapable of accurately reproducing the ISIHs of NL. The values of the parameters used are given in table 2.1.

Table 2.1. Parameters used in the inhomogenous Poisson process

	NM	NL	ICcc
R_A	0.6 ms	1 ms	1 ms
c_0	1.0	0.003	0.0015
c_1	0.2	0.003	0.0015
c_2	0	1	1
c_3	12.5	2	2

Chapter 3

Spectrotemporal Receptive Fields

There is considerable evidence that the time-dependent structure of auditory signals is a major factor in the task of sound recognition. Fine temporal structure is a significant determinant in the discrimination and recognition of bird song (Brenowitz 1983), for example, and evidence suggests that intact temporal information permits the comprehension of speech even with degraded spectral cues (Drullman, Festen and Plomp 1994; Shannon et al. 1995; Wright et al. 1997). To process this information, the animal requires an encoding of the spectral properties of the stimulus as a function of time. When considered at the level of a single neuron, the combinations of frequencies and temporal profiles to which the neuron is responsive is referred to as the *spectrotemporal receptive field* (STRF).

The brainstem auditory nuclei of the barn owl have been primarily characterized in terms of their behavior in the context of sound localization. However, all auditory information, including that needed for sound recognition and discrimination, ascends through these nuclei. Keller and Takahashi (2000) characterized the spectrotemporal tuning features of the lateral shell of the central nucleus of the inferior colliculus (ICcl). ICcl receives inputs from both the ITD pathway (Takahashi and Konishi 1988b) as well as the ILD pathway (Takahashi, Wagner and Konishi 1989). Both cues are related to the task of sound localization, and it is not clear which, if any, of the two associated pathways would be better suited to encoding information for the purposes of sound recognition. Since none of ICcc, LLDp, and nucleus angularis (NA; the first cochlear nucleus in the ILD pathway, which also projects directly to ICcl) have been characterized in terms of STRF, the question remains

open as to which of these nuclei lead to the STRFs of ICcl, or if a combination of both is required. However, the initial nuclei of the ITD pathway, the auditory nerve (Köppl 1997b), NM (Köppl 1997b), and NL (Carr and Konishi 1990; Peña et al. 1996) phase-lock when presented with tonal stimuli: they fire preferentially at a particular phase of the stimulating frequency. This link between a property of the stimulus and spike timing in the case of tonal stimuli suggests that there might be significant spectrotemporal tuning in the case of complex stimuli. Indeed, a link between phase-locking and spectrotemporal tuning for complex stimuli has been demonstrated in the auditory nerve of mammals (de Boer and de Jongh 1978; Eggermont 1993; Kim and Young 1994; Lewis, Henry and Yamada 2002; Louage, van der Heijden and Joris 2004). Thus, the ITD pathway is a reasonable candidate for the source of spectrotemporal tuning in ICcl.

In addressing a possible role of ICcc in the context of the post-laminaris pathway, its spectrotemporal tuning is a logical place to commence. Unlike the auditory nerve, NM, and NL, ICcc does not phase-lock to tonal stimuli. At the same time, the temporal resolution required to produce phase-locking at the frequencies of interest in the barn owl is on the order of tens of microseconds (Köppl 1997b). It is entirely reasonable to believe that such high-quality temporal resolution can be degraded and still provide information about the instantaneous spectral properties of a stimulus with temporal resolution of use to the organism. In partial support of this hypothesis is the fact that ICcc projects directly to the auditory thalamus, which in turn projects to Field L, the avian homologue to the auditory cortex of mammals (Proctor and Konishi 1997; Cohen, Miller and Knudsen 1998). Conversely, it might be that there is no stimulus-locking present in any form in ICcc; this would mean that the firing rate of an ICcc neuron is a function solely of ITD, which would simplify the task of extracting the position of the sound source.

In the work presented here, we set out to describe the STRFs of nuclei in the ITD pathway, with the intent of determining if spectrotemporal information is preserved through to ICcc. Using reverse correlation methods, we describe the STRFs of neurons in NM, the cochlear nucleus that represents the start of the ITD pathway, in NL, and in ICcc. Though NM is not of direct relevance to the general topic of this dissertation, it provides an interesting point of comparison as the input

to nucleus laminaris. Our results demonstrate that the STRFs of neurons throughout this ascending pathway are similar, despite the loss of phase information in the transition from NL to ICcc. This suggests that phase is of relevance primarily for the computation of ITD, and is encoded in the output of NL as an artifact of the computation of ITD, but also that high temporal resolution information about the envelope of the stimulus is preserved throughout the ITD pathway.

3.1 Methods

See chapter 2 for general methods.

3.1.1 Data collection

The occurrence of a spike is assumed to be related to the occurrence of a stimulus feature to which the neuron is sensitive. To determine the stimulus features to which the neuron is sensitive we used reverse correlation (de Boer and de Jongh 1978). In this method, we first compute the *pre-event stimulus ensemble* (PESE): a matrix in which row n contains the segment of the stimulus that preceded spike n. By examining the statistical properties of this matrix it is possible to determine the stimulus features that precede, and presumably elicit, spikes.

Data for reverse correlation were obtained by presenting 100 ms broadband signals at the best estimate of the characteristic delay of each neuron until a large number of stimulus-evoked spikes (NM: 7,330 ±4,250; NL: 2,909 ±1,343; ICcc: 2,792 ±1,352) that occurred between 40 ms after stimulus onset and the end of the stimulus were collected. Onset transients in the neural response of these nuclei are generally gone within 20 ms of stimulus onset. The time window of the reverse correlation (e.g., the amount of stimulus preceding each spike that was considered in the analysis) was 20 ms. Hence, the first 40 ms of response was removed to prevent the onset transient from appearing within the stimulus window. For each stimulus presentation, the signal was synthesized *de novo* to avoid correlation artifacts, and stimuli using a second ITD (or intensity, in the case of NM) were interleaved during the collection to prevent any effects of habituation to ITD.

Data on the variability and reproducibility of neuronal responses were obtained using the same

protocol as reverse correlation, except that the stimuli were repeated presentations of the same broadband stimulus.

3.1.2 Analysis

Principal components analysis (PCA) of an $m \times m$ matrix M produces eigenvalues $\lambda_i, i = 1 \ldots m$ and a matrix of eigenvectors X such that $M = X\Lambda X^{-1}$, where $\Lambda(i,i) = \lambda_i$ and is zero elsewhere. The eigenvectors of M can be thought of as an optimal set of orthogonal axes for describing the data, and the corresponding eigenvalues describe the relative importance of that dimension. Smaller eigenvalues indicate that the corresponding eigenvector contributes little to the overall variance of the data, and can, if sufficiently small, be neglected.

PCA requires that the original data matrix be square. The singular value decomposition (SVD) can be thought of a generalization of PCA which removes this restriction. An $m \times n$ matrix M is represented by $M = U\Lambda V^T$, where U is $m \times m$, V is $n \times n$, and Λ is an $m \times n$ diagonal matrix whose entries are the *singular values* λ_i of M. The fractional energy of a singular value λ_i is given by $\lambda_i^2 / \sum_j \lambda_j^2$, and is a measure of the relative contribution of the associated singular vector pair to the reconstruction of the overall matrix. The fractional energy of λ_1 is equal to $1 - \alpha_{\mathrm{SVD}}$, where α_{SVD} is the degree of inseparability defined by Depireux et al. (2001). In this work, whenever we compute the SVD we first subtract off the mean of M; otherwise the first singular value is dominated by a constant component. Since we use PCA only for the analysis of the covariance matrix, which is guaranteed to have mean 0, this step is not required there.

To determine significant second-order spike-triggered effects, the covariance matrix of the PESE was computed. Covariance is a statistical measure which can be thought of as the two-dimensional equivalent of the variation. The covariance of two random variables x and y is given by:

$$\mathrm{Cov}(x,y) \quad = \quad E[xy] - E[x]E[y]$$

where $E[\cdot]$ is the expectation of the random variable (Ross 1994). In our case, the random variables

correspond to the values of the stimulus at a precedence of t. This gives a matrix $\mathbf{C}$, whose (i, j)th element is given by:

$$\mathbf{C}_{(i,j)} = \mathrm{Cov}(\mathrm{PESE}_i, \mathrm{PESE}_j)$$

where PESE_i is the random variable that describes the value of the stimulus at a precedence i.

To analyze the covariance matrix, we used PCA. This gives a family of eigenvectors, and we would like to demonstrate that we can safely disregard the majority of them. Fractional energy is an insufficient standard to make this argument except in the case of extreme values, as it does not provide a criterion for significance. Instead, we proceed by constructing a first-order model. By design, such a first-order model does not have second-order effects (i.e., it will have no significant covariance eigenvectors). If the distribution of eigenvalues in the data is significantly different from the distribution seen in the first-order model, then we can conclude that there are significant second-order effects present in the data. The first-order model assumes firing rate is related to the convolution of the spike-triggered average and the stimulus. A nonlinear weighting function of firing rate as a function of filter output was recovered using the method of Rust et al. (2004): a histogram of the inner products of all pre-event stimuli and the spike-triggered average was computed, and then piecewise normalized by the histogram of the inner products of all possible stimuli segments within the presented stimulus set and the spike-triggered average. The resulting histogram thus gives the probability of spiking as a function of the inner product of the stimulus and the spike-triggered average, and was fit with an asymmetric parabola of the form:

$$f(x) = \begin{cases} a_1 x^2 + b & x \leq 0 \\ a_2 x^2 + b & x > 0 \end{cases}$$

A spiking model was then implemented using an inhomogeneous Poisson process (Zhang et al. 2001) whose rate parameter was found by convolving the auditory stimulus with the spike-triggered average of that neuron, passing the result through the weighting function, and then recovering a scaling factor

through gradient descent to match the model's mean firing rate to the data (see section 2.5). The spiking model was used to generate 100 PESEs of the same size as the original data PESE, thus giving a family of eigenvalue distributions for their respective covariance matrices. An eigenvalue of the data covariance was considered significant if it and all eigenvalues greater in magnitude differed from the model distribution by at least two standard deviations.

Power spectral density (PSD) was estimated with the MATLAB implementation of Thomson's multi-taper method with the time-bandwidth product set to 5/2. We use three characterizations of the spectral properties of a signal. BW_{10} is the 10 dB bandwidth of the signal; that is, it is the width of the peak of the periodogram 10 dB below the peak. CF_{10} is the frequency on which the BW_{10} is centered, and PF is the peak frequency of the periodogram of the signal. Care was taken to ensure that multiple peaks did not confound this measure.

To compute the latencies of filters, we first computed the continuous wavelet transform of the filter using a fourth-order Daubechies over the set of scales that corresponded to the bandpass range of the stimuli (generally 1–12 kHz). After computing the mean and standard deviation of the transform over this range, we then computed the wavelet transform at a scale which corresponded to the PF of the filter (PWT). We identified all points of the PWT whose absolute value exceeded the mean plus two standard deviations of the full transform and treated them as significant. Adjacent singificant points were collected into segments, and segments which were separated by no more than half the period of the peak frequency were combined (i.e., the intermediate points were also treated as being significant). The segment that included the maximum of the absolute value of the PWT was treated as being the time interval of the filter, allowing us to identify minimum, maximum, and peak latencies. Example of the result of this method can be seen in figure 3.1.

Phase differences were calculated by fitting a gammatone (Lewis, Henry and Yamada 2002) to that portion of the filter between maximum and minimum latencies and then comparing the phase offsets of the cosine component of the fitted gammatone. A gammatone was chosen because of precedents set in the literature, and because it successfully managed to capture the phase structure of the filters. However, it should not be considered to be a truly valid description of the filters, as

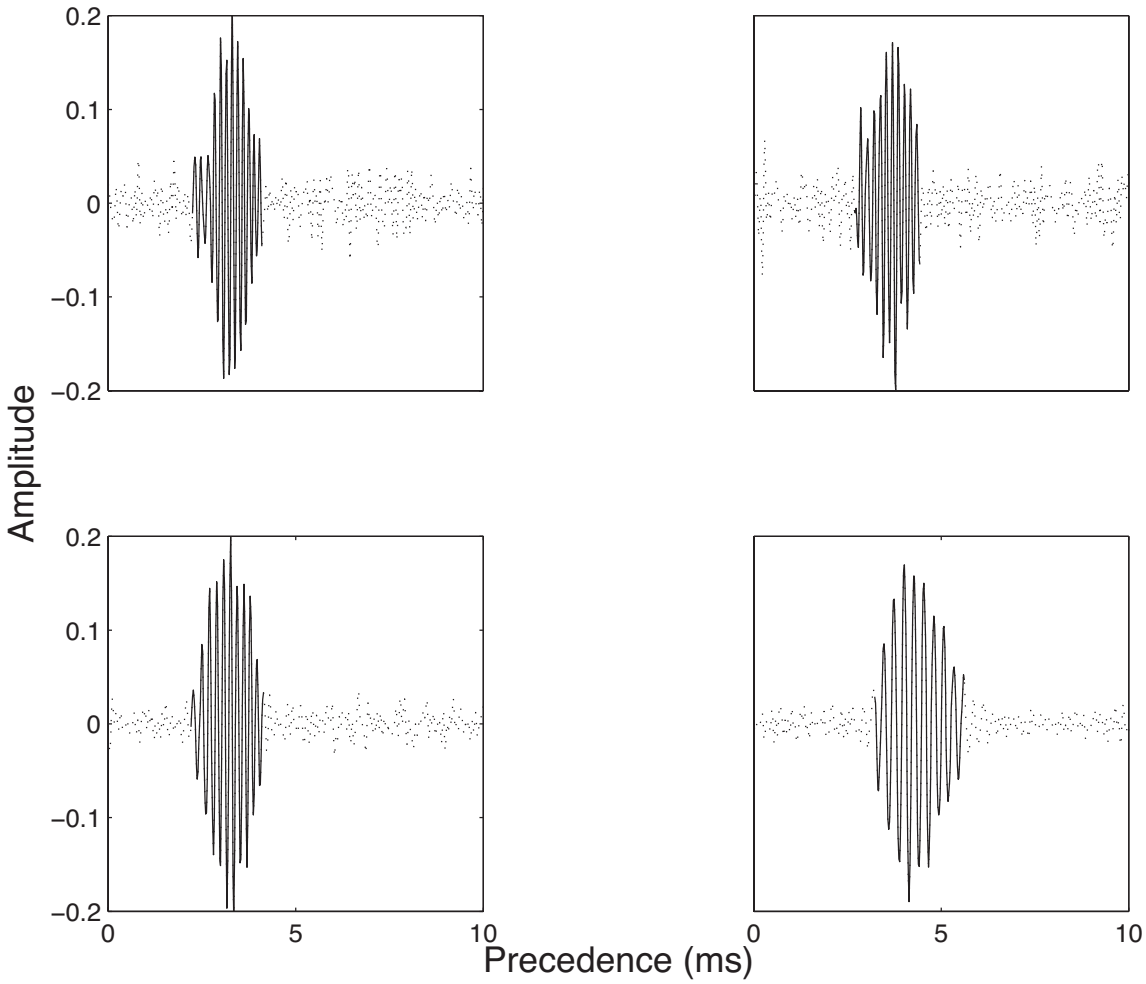


Figure 3.1. Shown are four results of the wavelet-based latency estimation method. The complete filter is plotted using the dotted line, and that region of the filter that was identified as being between the minimum and maximum latency is highlighted with a solid line. The method identifies the "non-noisy" portion of the signal as perceived by the eye well, and in an automated fashion. Peak latency would correspond to the time at which the maximum of the highlighted region occurs. In general, we found that both minimum (lowest value of precedence of the highlighted region) and maximum latencies were reliable, in the sense that they were systematic across the neurons of a given area. Peak latency tended to have more variability, and be a less reliable indicator of the neuronal latency.

the filters often displayed a temporal asymmetry about the peak.

Spectrograms were computed using a 100-sample Hamming window with an overlap of 75 samples. Color maps of spectrograms are in grayscale, with white corresponding to minimum values and black to maximum.

The shuffled autocorrelogram (SAC) is computed as in Joris (2003). We have N spike trains arising from presentations of the same broadband stimulus to a particular neuron, and begin by

choosing the first spike train. For each spike in that train, we compute the forward time intervals (that is, the difference in spike times between that spike and all spikes occurring after it relative to stimulus onset) between that spike and all spikes in the other spike trains, and then this procedure is repeated until all spikes in all spike trains have been used as the reference. Because there are no intra-train comparisons, effects of refractory period are eliminated. The SAC is guaranteed to be symmetric about the delay of zero (each forward time interval will reoccur as a backward time interval), and hence we need only consider the forward time intervals. A normalizing factor $N(N-1)r^2\Delta\tau D$, where r is the mean firing rate, $\Delta\tau$ is the bin-width of the correlogram, and D is stimulus duration, results in a unity baseline, where a spike train with Poisson statistics will have a flat SAC of height 1. We used a histogram of binwidth 50 μs, as in Louage, van der Heijden and Joris (2004).

3.2 Results

3.2.1 General properties

The spike-triggered average (STA) is computed by averaging the PESE across spikes, and gives the average stimulus preceding a spike. For auditory stimuli, the STA is only non-negligible in the case where there is precise synchronization to the stimulus; in particular, the neurons must display phase-locking (Eggermont 1993). Consistent with reports of phase-locking in NM (Köppl 1997b) and in NL (Carr and Konishi 1990; Peña et al. 1996), we saw coherent STAs in these nuclei for all neurons examined (fig. 3.2a–d). In ICcc, neurons do not phase-lock, and as anticipated, the STAs of ICcc neurons lacked coherent structure (fig. 3.2e).

The STA captures only the first-order relationship between stimulus and response, and is hence incomplete. As such, we also considered the spike-triggered covariance (STC), given by the covariance of the PESE. The covariance is a matrix whose interpretation is roughly analogous to cross-correlation: the covariance at (m, n) will be positive if the values of the stimulus at times m and n preceding the spike are likely to have the same sign, and negative if they are likely to have

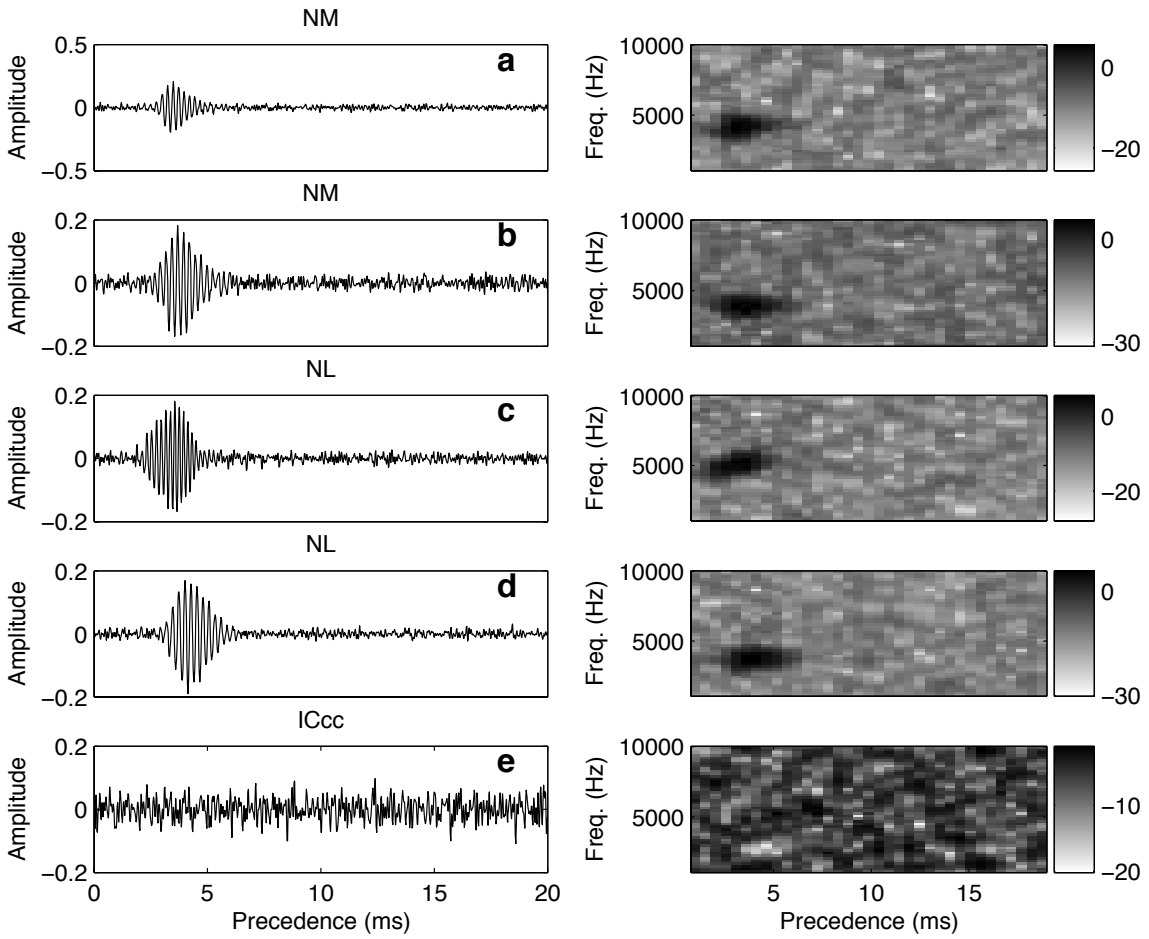


Figure 3.2. Shown are the spike-triggered averages of example NM (***a***–***b***, *left column*), NL (***c***–***d***), and ICcc (***e***) neurons, along with their spectrograms (*right column*). The abscissa gives the time preceding the spike; color bars are provided for reference, but in following figures will be omitted ***a***: 2003Oct31-863.06, 6,614 spikes; ***b***: 2003Nov28-863.07, 5,234 spikes; ***c***: 2005Feb18-854.02, 2,984 spikes; ***d***: 2005Mar11-813.02, 3,509 spikes; ***e***: 2005Jun08-842.03, 3,468 spikes.

opposite signs, with the magnitude of the covariance denoting the likelihood of that relationship. Because covariance looks at correlations within the signal, rather than absolute value, it does not require a precise synchrony to phase. Indeed, when we look at the spike-triggered covariance of an ICcc neuron, structure is apparent (fig. 3.3a). In this example, alternative bands of positive and negative covariance running parallel to the diagonal centered at a precedence of approximately 8 ms are present. For any given spike, the relationship of the phase of the stimulus at time t can be determined by taking a slice through the covariance matrix at t. Figure 3.3a tells us that for values of t near 8 ms, the nearby stimuli will vary in and out of phase with the value at t with a frequency

given by the alternation of those bands, regardless of the actual value of the stimulus at t,

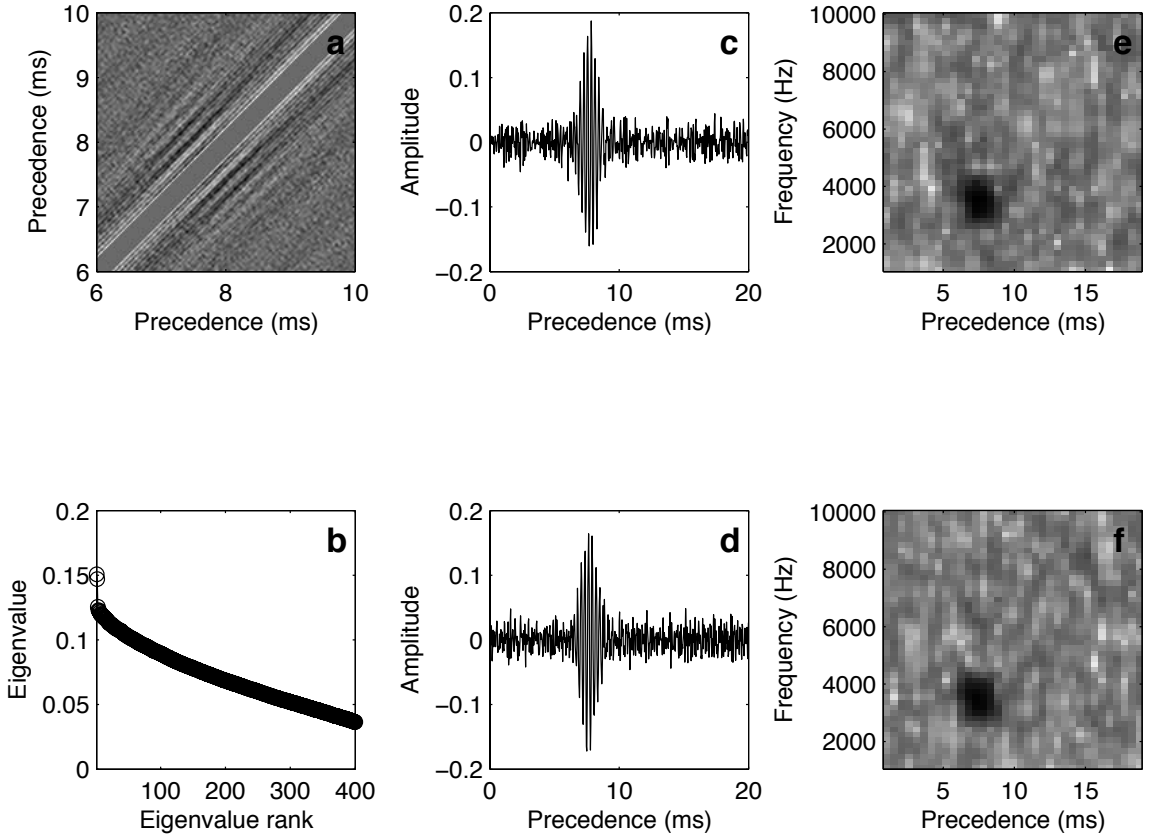


Figure 3.3. The spike-triggered covariance of the ICcc neuron shown in fig. 3.2e. ***a***: The covariance matrix itself is shown. The diagonal of the matrix gives the variance, which tends to dominate over the covariance. To emphasize the covariance structure, the central 40 μs along the diagonal have been zeroed out for this plot only (thus, the band of gray along the diagonal represents zero; black positive values, white negative values). The rippling pattern of the covariance is now visible. ***b***: The first four hundred eigenvalues of the covariance matrix, sorted by magnitude. It can be seen that the first two eigenvectors are distinct from the smooth progression seen in the remaining values. ***c*** and ***d***: The eigenvectors corresponding to the first and second eigenvalues, respectively. Note their structure is similar to the structure of the STAs in fig. 3.2. ***e*** and ***f***: The corresponding spectrograms of the eigenvectors.

The STC is difficult to visually interpret, and since the tuning is limited to a certain range of time, it contains an excess of data. To deal with this, we used principal component analysis, determined the significant eigenvectors of the matrix, and used only those eigenvectors in our analysis (see Methods; fig. 3.3b–f). Of the 27 ICcc neurons examined, 24 displayed at least one significant covariance eigenvector, and 18 of those had two.

The existence of a significant STA and STC are not mutually exclusive. Of the 16 NM neurons

examined, 3 had a single significant covariance eigenvector. Of the 19 NL neurons, 10 had at least one significant covariance eigenvector, and 6 of those had two. To understand why multiple filters (either the STA with STC eigenvectors, or STC eigenvectors alone) are present, we note that Yamada and Lewis (1999) demonstrated in the bullfrog's auditory nerve that non-phase-locked responses were described by pairs of filters. These pairs, called *quadrature pairs*, were identical in spectral profile but shifted ninety degrees in phase with respect to each other. A quadrature pair forms a basis with which all possible signals with the same spectral profile but different phases can be generated. Different weightings of the filters will produce different phase biases in the response. Even neurons with a high degree of phase-locking will respond to a wide range of phases (Johnson 1980; Carr and Konishi 1990; Köppl 1997b). Thus, depending on the nature and quality of phase-locking for a single neuron, a quadrature pair of an appropriate bias may be required to describe the response. When we compare the spectral profiles of the neurons with two filters (STA and the first STC eigenvector (STC_1) for NL and NM, and STC_1 and the second STC eigenvector (STC_2) for ICcc), we see that they are similar (fig. 3.4). Equally important is determining that their temporal profiles are similar, and the latencies of the appropriate filters are in fact well correlated (fig. 3.5). Estimations of the phase differences further confirm that the filters are separated by approximately ninety degrees (NM: $91 \pm 5.7°$; NL: $89 \pm 7.9°$; ICcc: $97 \pm 27°$), confirming that when only two filters are present, they approximately form a quadrature pair.

The quadrature pair hypothesis does not predict the possibility of a third filter. However, it is clear that while the CF_{10} of the respective filters are the same (fig. 3.4b), the BW_{10} (fig. 3.4e) and the latency (fig. 3.5b,e) do not show the same similarity as seen in the cases where there are only two filters. This point will be examined in further detail in the Discussion. Excepting this case, however, we observe that multiple filters are generally differentiated only by their phase. Since the majority of our analyses do not focus on phase, from this point forward we will primarily refer only to a single filter, and refer to that as the STRF.

An observation that springs from figure 3.5 is that the range of latencies seen in ICcc is much greater than in NM or in NL. ICcc latency is highly correlated with the preferred frequency of the

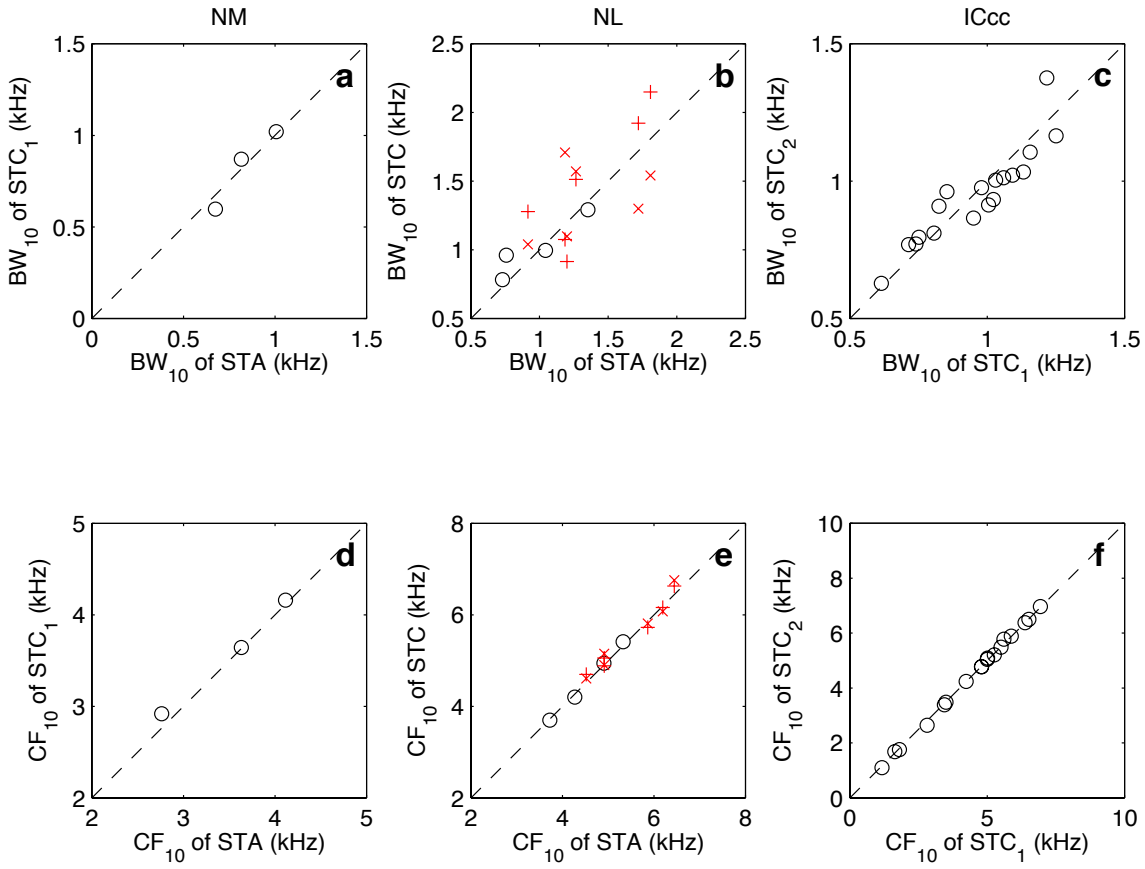


Figure 3.4. ***a***: When the BW_{10} for the STA and the first eigenvector of the STC (STC_1) are plotted against each other for NM, we see a generally linear relationship, though the paucity of data makes this conclusion tenuous. This is true when the same is plotted in NL for neurons with only STC_1 significant (***b***, ∘). The dispersion is greater for the NL neurons which had two significant eigenvectors (STA vs STC_1: ×; STA vs STC_2: +). ***c***: The BW_{10} of STC_1 and STC_2 for ICcc neurons with two significant covariance eigenvectors. Again, a good correspondence in the majority of cases is observed (regression: $r^2 = 0.83$, $p < 10^{-6}$. ***d***–***f***: The CF_{10} in all cases showed a high degree of correspondence between filters (symbols as in ***a***–***c***; regression for ICcc: $r^2 = 1.00$, $p < 10^{-6}$.). Regression values are not given for NM and NL due to low n.

neuron (fig. 3.6b; regression for maximum latency: $r^2 = 0.77$, $p < 10^{-6}$; regresssion for minimum latency: $r^2 = 0.66$, $p < 10^{-5}$). On the other hand, while the correlation is present in NL, it is weaker (fig. 3.6a; regression for maximum latency: $r^2 = 0.62$, $p < 10^{-4}$; regression for minimum latency: $r^2 = 0.26$, $p > 0.01$). Latencies in NM were entirely uncorrelated with frequency ($p > 0.01$ for both cases; data not shown).

In both NM and ICcc, linear regression did not reveal any significant dependence of BW_{10} on CF_{10} ($p > 0.05$) for the first filter (STA and STC_1, respectively; fig. 3.7a,c). In NL, however, a linear

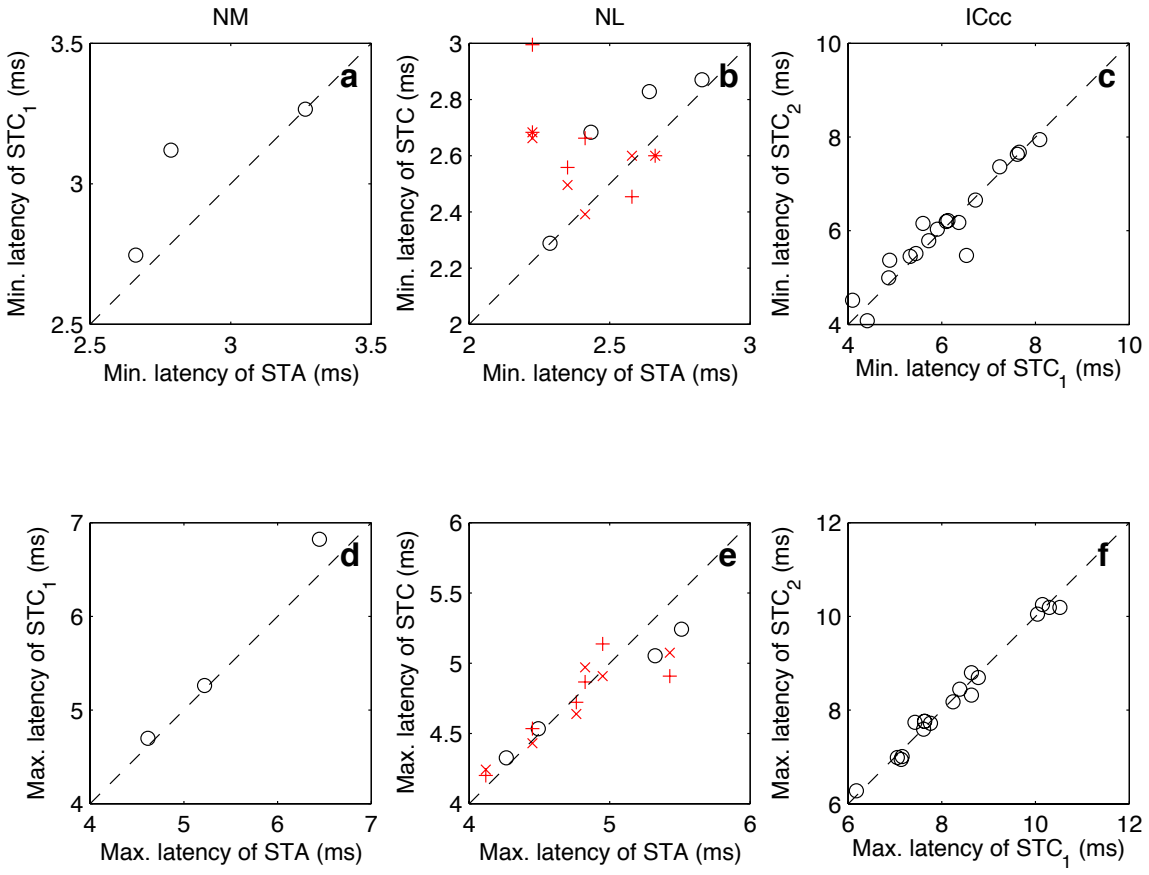


Figure 3.5. Comparison of minimum latency for paired filters in NM (***a***), NL (***b***), and ICcc (***c***; regression: $r^2 = 0.90$, $p < 10^{-6}$), with symbols as in figure 3.4. The two-filter cases are generally well correlated; though some deviations are seen in NM and NL, the overall range of values is quite small. The maximum latencies (***d***–***f***) were well correlated in all cases (regression for ICcc: $r^2 = 0.98$, $p < 10^{-6}$; again, regression values for NM and NL are not provided due to low n).

dependence was seen ($p < .001$) for the STA (fig. 3.7b). Segregating the NM and ICcc populations by the number of filters did not affect this result. This stands in contrast to the observation that in NM (Köppl 1997a) and in ICcc (fig. 3.8) estimates of frequency tuning bandwidths using single-tone methods reveal such a dependence.

The values of BW_{10} seen in NM were smaller than seen in NL (Kruskal-Wallis, $p < 0.001$) and in ICcc ($p < 0.005$; fig. 3.7d), while ICcc and NL saw a similar range in BW_{10} values ($p > 0.1$). This suggests a convergence across frequency channels from NM to NL, but that no such convergence takes place from NL to ICcc.

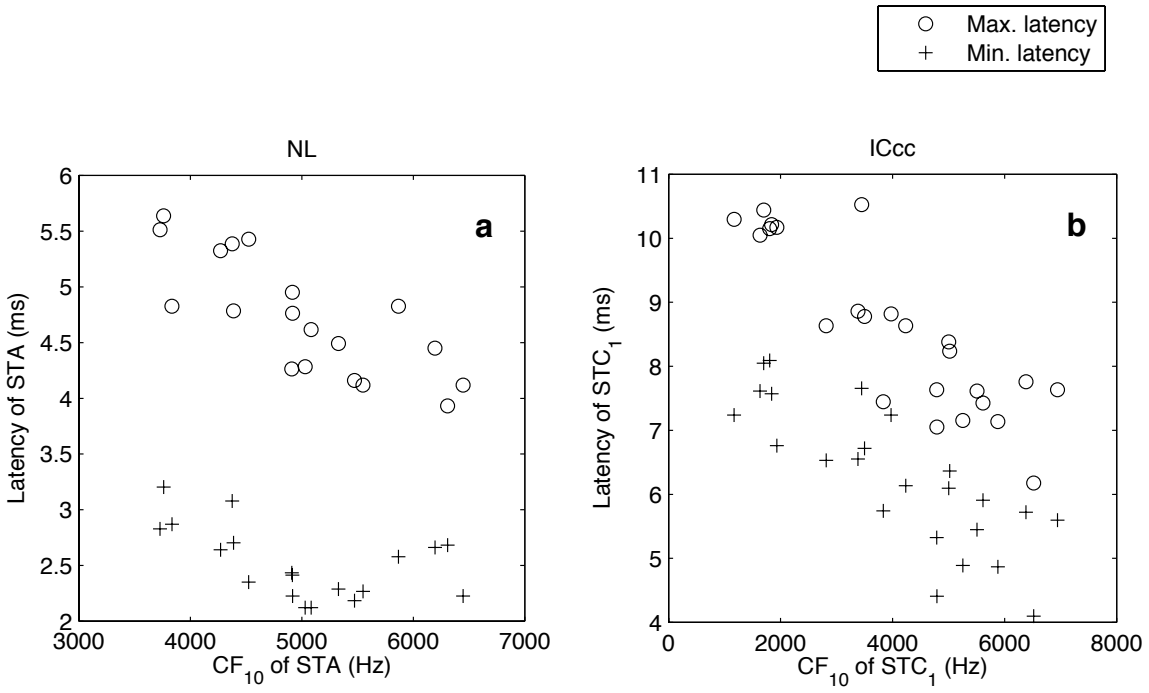


Figure 3.6. ***a***: In NL, a linear relationship between maximum latency (∘) and CF_{10} was observed ($r^2 = 0.62$, $p < 10^{-4}$). The correlation for minimum latencies was not significant (+; $p > 0.01$); as can be seen, the slope of the minimum latencies is almost zero. Both minimum and maximum latencies in ICcc showed a strong linear dependence on CF_{10} (***b***; maximum latency: $r^2 = 0.77$, $p < 10^{-6}$; minimum latency: $r^2 = 0.66$, $p < 10^{-5}$). No relationship was seen for latencies in NM (data not shown).

3.2.2 Effect of ITD on STRF

The neurons of both NL and ICcc are known to be tuned to ITD, which is a function of the position of the signal in space. The data on STRFs we have shown was collected at the CD alone. It is reasonable to ask what, if any, effect ITD has on the STRF properties. When we collected the reverse correlation, we interleaved trials using a second ITD to prevent habituation. These second ITDs were chosen so that across the entire set of neurons collected they covered a variety of different ITD conditions, including favorable (a peak on the ITD tuning curve other than CD), unfavorable (a trough on the ITD tuning curve), and intermediate values. In ICcc we rarely chose unfavorable ITDs, as they generally had firing rates that were too low to collect a sufficient number of spikes for reverse correlation.

Figure 3.9 compares the properties of the STRFs calculated with these non-CD ITDs to the

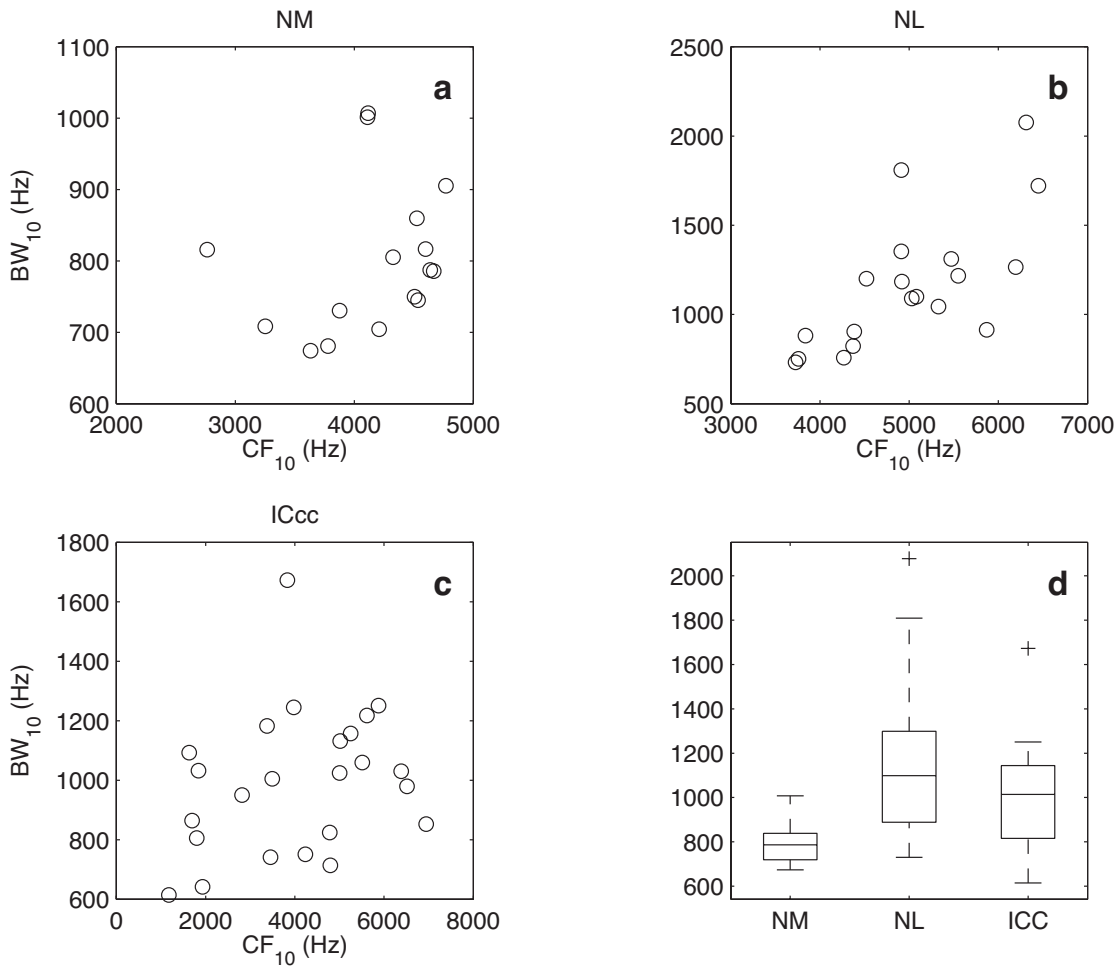


Figure 3.7. Plots of BW_{10} vs. CF_{10} for NM (***a***), NL (***b***), and ICcc (***c***). Both in NL and NM there is an dependence of BW_{10} on CF_{10}, while in ICcc no such relationship is apparent. ***d***: Distribution of BW_{10} in the three nuclei. The bandwidths observed in NM are significantly lower than seen in either NL or ICcc (see text). Removal of the three apparent outliers in ***a*** does introduce a linear relationship ($p < 0.01$), but examination of those neurons gives no grounds for exclusion. Removal of the outlier in ***c*** has no effect.

STRFs at CD of the corresponding neurons. For all the parameters examined, and in both NL and ICcc, a strong linear relationship was observed between CD and non-CD STRFs, and the 95% confidence intervals of the regression included the identity line in all cases. Note that we have not controlled for firing rate, so that some of the observed variability may be due to differences in spike counts for the various reverse correlations. Based on these data, there is no evidence to suggest that there is a significant dependence of the SRTF on the ITD of the stimulus.

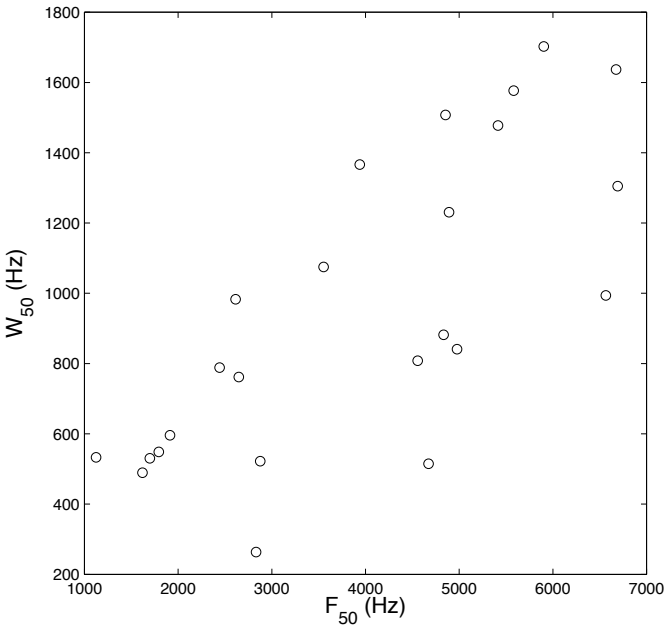


Figure 3.8. A linear relationship is observed between F_{50} and W_{50} when frequency is estimated using an iso-intensity frequency tuning curve in ICcc (regression: $0.18x + 237$ Hz, $r^2 = 0.55$, $p < 10^{-4}$).

3.2.3 Separability

The STRF gives the tuning of a neuron to two stimulus dimensions, frequency and time. A relevant question is to what degree these dimensions are separable: that is, can the STRF be considered the product of the separate tuning to time and frequency, or are there non-negligible interactions between the two? Using singular value decomposition (SVD), this question is equivalent to asking if a single singular vector pair is sufficient to reconstruct the STRF, or if more than one is required.

Figure 3.10 demonstrates the STRFs of three example units, one for each nuclei, as well as reconstructions of the STRF using one and two singular vector pairs. The STRFs do not in general display any obvious signs of inseparability (fig. 3.10a, d, g), such as multiple peaks, or angled peaks (which would indicate a change in preferred frequency with time). As expected from this, the reconstructions using only a single singular vector pair (fig. 3.10b, e, h) capture the primary features of the tuning as visible to the eye. Adding a second singular vector pair to the reconstruction (fig. 3.10c, f, i) adds some additional structure to the peaks; this is especially apparent in the asymmetric taper of the peak in either direction along the time axis. However, the net change is

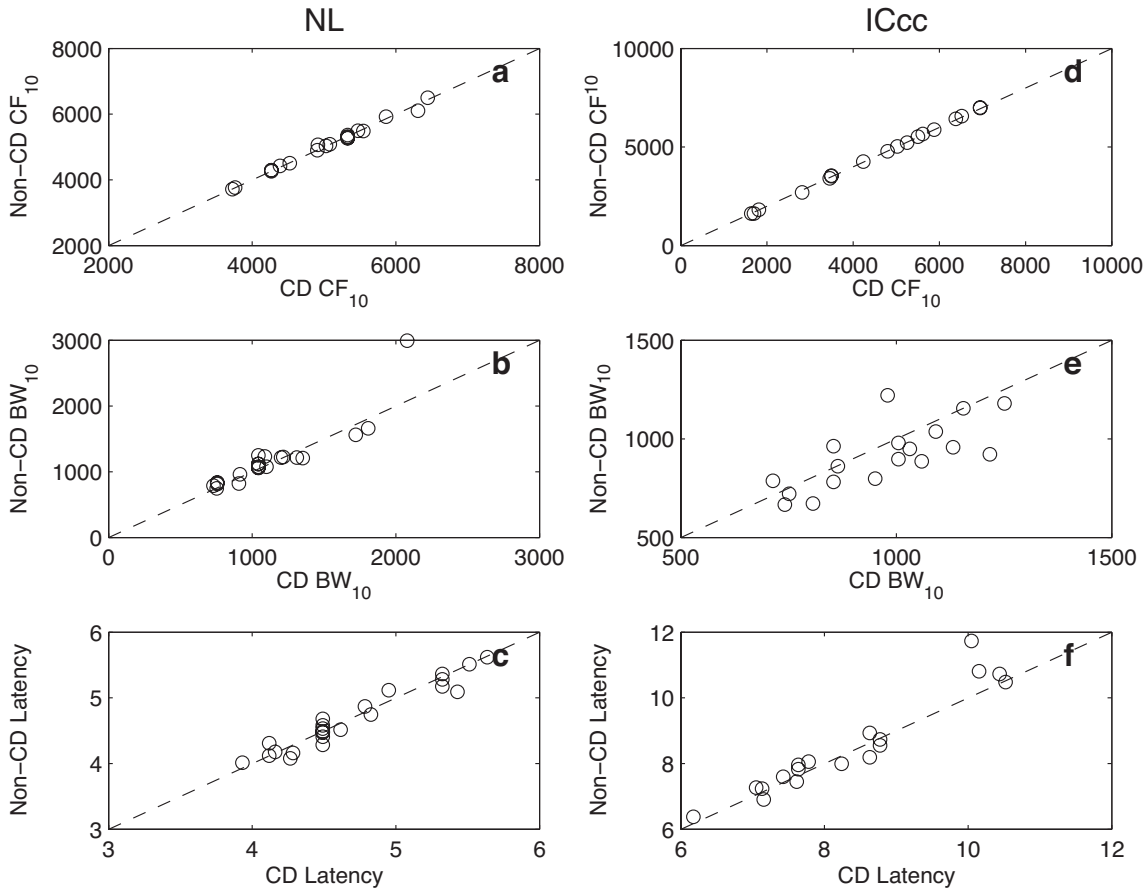


Figure 3.9. STRFs were estimated for all non-CD ITDs collected, and the CF_{10} (***a***, ***d***), BW_{10} (***b***, ***e***), and maximum latencies (***c***, ***f***) were estimated using those STRFs and then plotted against the corresponding values for the STRF of that neuron as estimated at CD. In NL (***a***–***c***), all parameters showed a strong linear dependence (***a***: $r^2 = 0.99$; ***b***: $r^2 = 0.82$; ***c***: $r^2 = 0.93$; $p < 10^{-6}$ in all cases). In ICcc, a correlation was also obvious, though weaker in the case of BW_{10} (***d***: $r^2 = 0.1$; ***e***: $r^2 = 0.54$; ***f***: $r^2 = 0.91$; $p < 10^{-6}$ for ***d*** and ***f***, $p < 0.001$ for ***e***).

apparently small, with most of the difference occurring in regions that seem to lack tuning.

When we look at the fractional energy of the singular values (fig. 3.11), we see that the first singular value clearly exceeds all other singular values in all cases. At the same time, the fractional energy of the first component in all cases represents less than half the total energy (NM: $41 \pm 4.8\%$; NL: $39 \pm 6.8\%$; ICcc: $30 \pm 4.1\%$), which indicates a high degree of inseparability (Depireux et al. 2001). However, this measure does not indicate the source of the inseparability. Specifically, figure 3.10 suggests that the higher-rank singular values may be used primarily to attempt to describe the noise. To address this, we note that based on considerations of neural latency, we expect tuning in

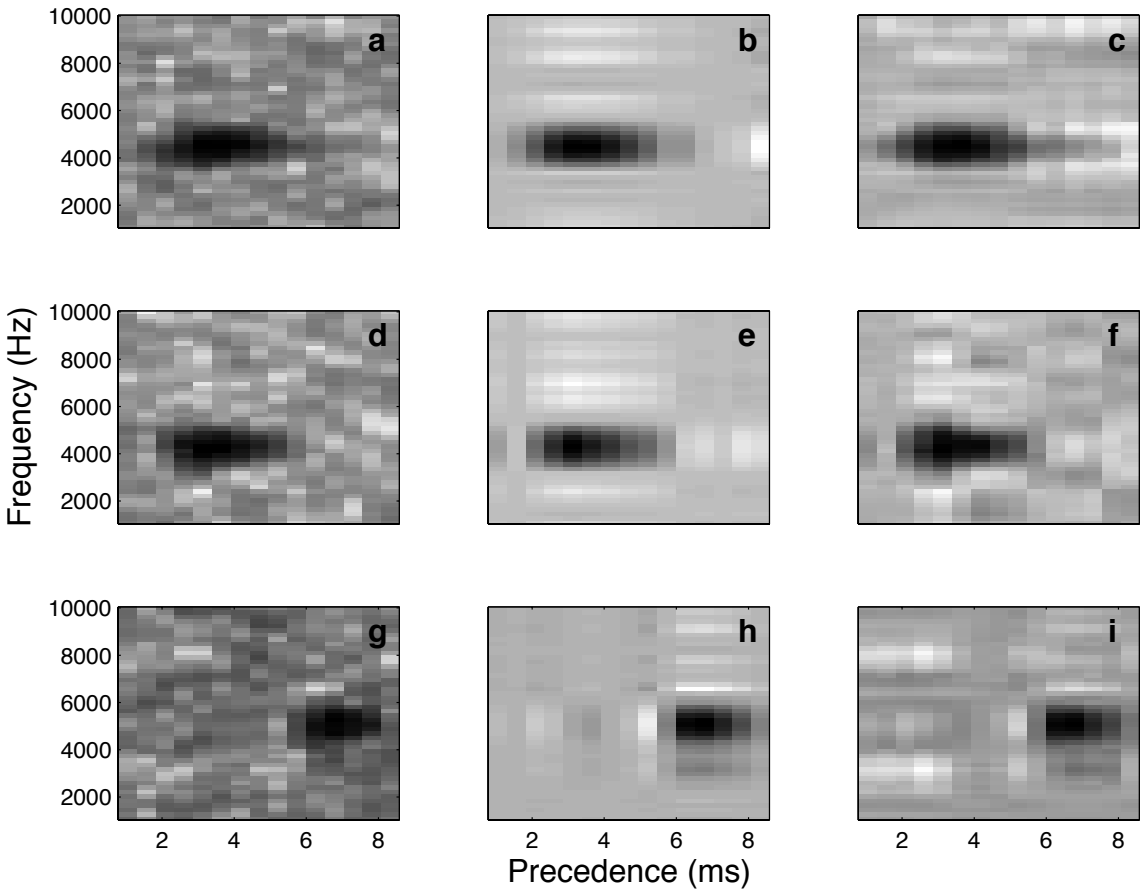


Figure 3.10. ***a***: Spectrogram of the STA for an example NM neuron. ***b***: Reconstruction of the STA's spectrogram using the first singular vector pair. ***c***: Reconstruction using the first two singular vector pairs. ***d***–***f*** and ***g***–***i*** are the same for the STA of a NL and the STC_1 of an ICcc neuron, respectively. The tapers of the tuning become more pronounced with the added singular vector pair.

all three nuclei to occur primarily at no more than 10 ms before spike time (figs. 3.2, 3.3). Thus, if we take the spectrogram of the STRF and divide it in half along the time axis, we expect that the majority of the power of the STRF (defined as the sum squared of the spectrogram) will occur in the 0–10 ms segment, rather than in the 10–20 ms segment. By extension of this, it is logical that if the spectrogram of a singular component of the STRF has the majority of its power in the first half, then that component contributes to the tuning itself. Conversely, a component that has power equally balanced between tuned or untuned regions, or that prefers the untuned region, is unlikely to be a major contributor to the actual tuning of the neuron. This analysis is shown in figure 3.12. As expected, the first singular value component is heavily biased towards the tuned region (fig. 3.12a,

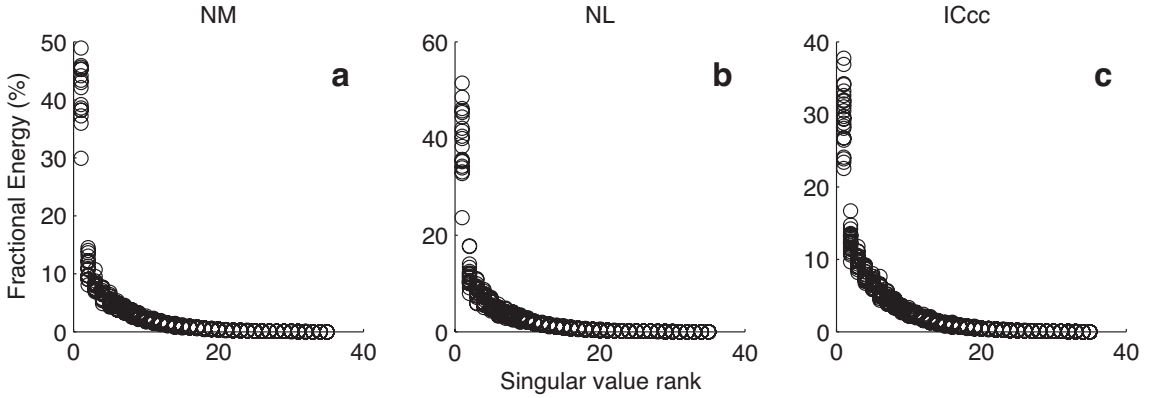


Figure 3.11. Fractional energy of the singular values of the STA (for NM (***a***) and NL (***b***)) and STC_1 (for ICcc; ***c***) for all neurons.

d, g). However, even the second component lacks this bias, with the power between the two regions having no significant difference (fig. 3.12b, e, h). Even when the summed activity of singular vectors other than the first is considered there is no indication of preferential action within the tuned region (fig. 3.12c, f, i). Taken all together, this suggests that there is no reason to believe that the observed inseparability reflects a significant interaction between frequency and tuning in the STRF.

3.2.4 Variability

The temporal profiles of the filters as shown in this research are narrow (figs. 3.2, 3.3). This indicates that the lag between the stimulus matching the STRF of the neuron and the spike time is relatively fixed across reoccurences of the preferred stimulus. Alternatively, we expect that the neurons of these nuclei would display low variability, and demonstrate similar responses for repeated presentations of the same stimulus. We tested this directly in 10 NL neurons and 22 ICcc neurons. As seen in figure 3.13a, b, the spike rasters are indicative of a fixed pattern of firing.

To quantify this, we used the shuffled auto-correlogram method (SAC; Joris 2003; Louage, van der Heijden and Joris 2004). The height of the SAC gives the likelihood that given a spike at time t in one presentation of a stimulus that another presentation of the same stimulus will elicit a spike at the same time; the width of the peak of the SAC gives a measure of how much jitter in spike timing is present (fig. 3.13).

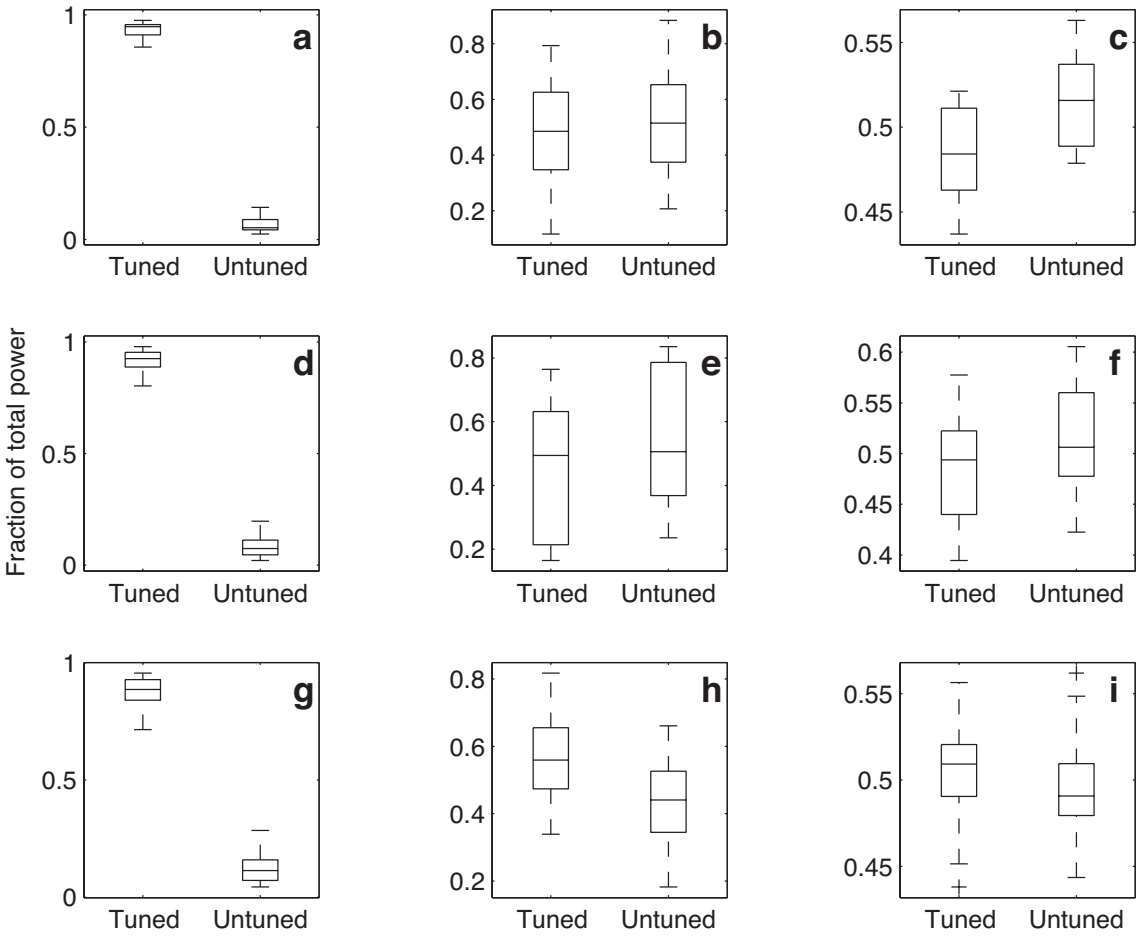


Figure 3.12. ***a***: Comparison of fractions of power in the first singular value of the STRFs for NM neurons in the tuned and untuned regions. ***b***: The same for the second singular value. ***c***: The comparison for the summed activity of all singular values but the first. Similarly for NL (***d***–***f***) and ICcc (***g***–***i***). Boxes extend from lower quartile to upper quartile of the sample, with the line marking the median. Outliers (+) are data points greater than 1.5 times the interquartile range of the sample.

Both NL and ICcc showed a low amount of jitter, as measured by the SAC half-height width (fig. 3.14a). The two populations were not significantly different (NL: mean 0.77 ± 0.37 ms; ICcc: mean 0.85 ± 0.29 ms; medians not different by Kruskal-Wallis, $p > 0.2$). The SAC heights also indicated a high probability of spikes reoccurring at fixed times, and there was no significant difference in this regard between the two nuclei (NL: mean 5.0 ± 4.9; ICcc mean 4.3 ± 2.6; medians not different by Kruskal-Wallis, $p > 0.5$). Thus, both NL and ICcc display jitter on the sub-millisecond scale, and this precision is maintained through the ascending pathway.

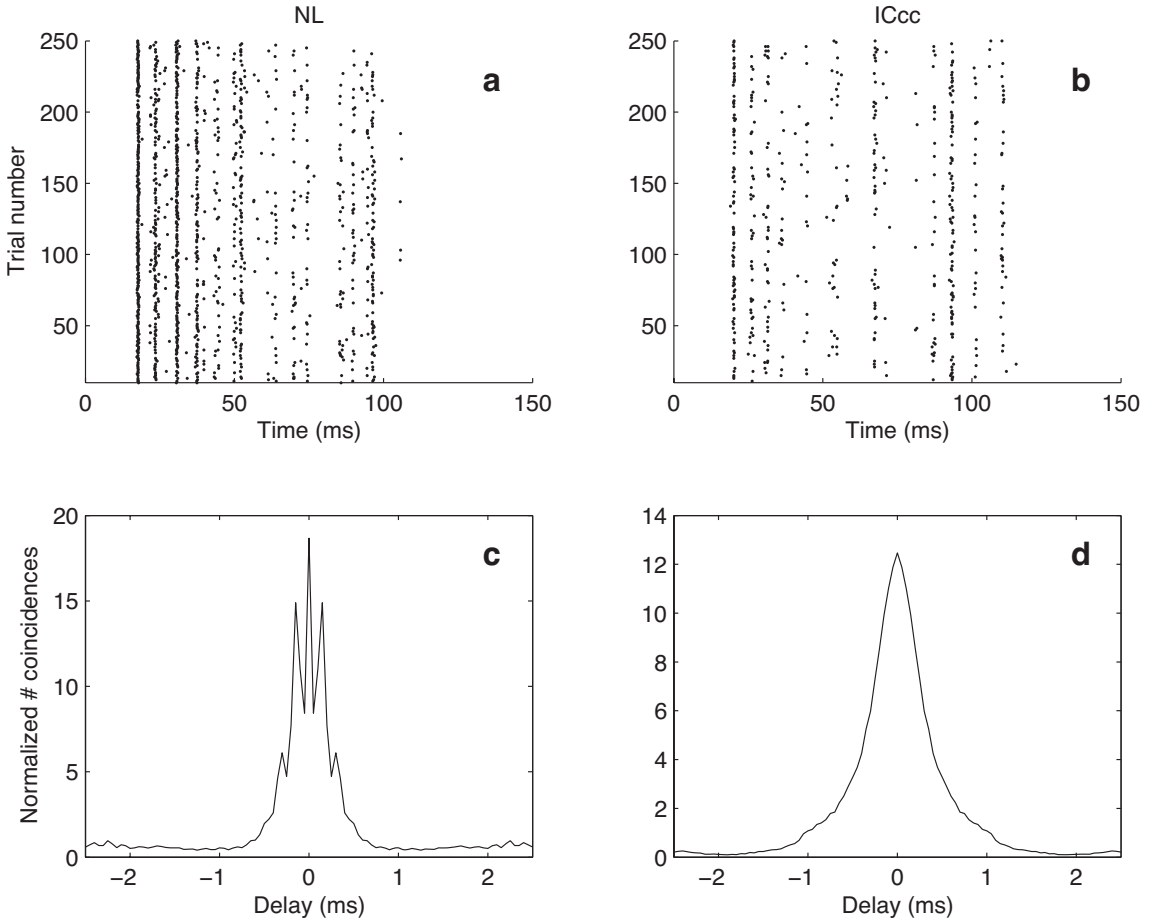


Figure 3.13. Example spike rasters for a NL (***a***) and ICcc (***b***) neuron, using repeated presentation of identical broadband stimuli. ***c*** and ***d***: The respective SACs for ***a*** and ***b***. Note the fine-time scale modulation of the NL SAC, which arises due to the phase-locking of the neuron, is absent in the ICcc SAC. Despite that, their outer half-height widths are similar.

3.3 Discussion

Previous studies in the barn owl (Keller and Takahashi 2000), as well as amphibian (Hermes et al. 1981), and mammalian models (Carney and Yin 1989), have demonstrated patterned (that is, stimulus-locked) responses in the midbrain using broadband noise. However, due to the segregation of sound localization cues in the barn owl, it is not immediately clear if these responses emerge from the ITD pathway or the ILD pathway. Our results indicate that the spectral properties of the stimulus are encoded with millisecond resolution throughout the ITD pathway. Keller and Takahashi (2000) observed that a continuum of STRF types in ICcl, ranging from simple, well-defined STRFs

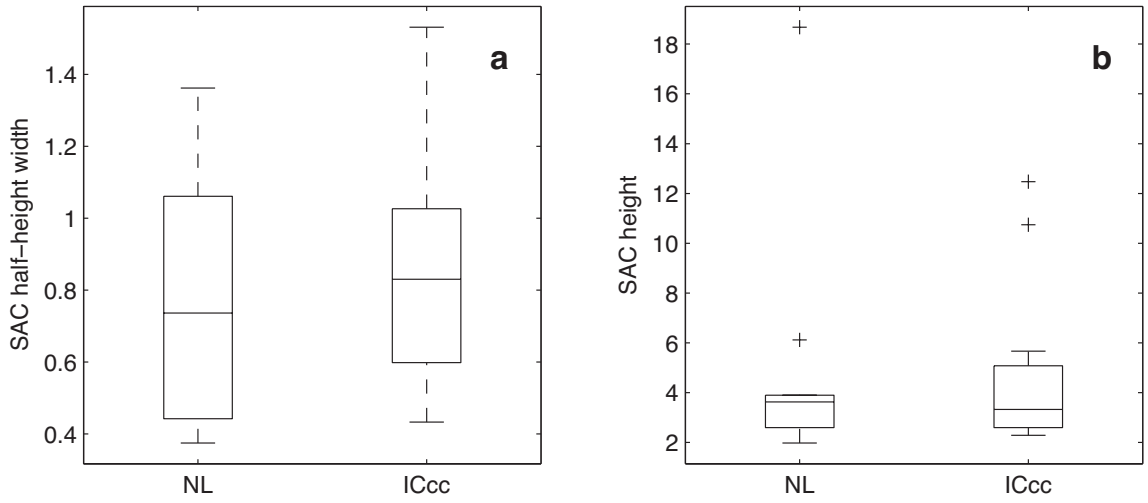


Figure 3.14. ***a***: The half-height widths of the SACs for NL and ICcc neurons. ***b***: The maximum height of the SACs for NL and ICcc neurons.

to more complex ones. Our study was done using different techniques, but the STRFs we describe in ICcc are similar to the simpler of the STRFs that they describe. While this does not rule out a contribution from the ILD pathway, it must be noted that STRFs have yet to be demonstrated within the ILD pathway.

The existence of three filters in some NL neurons is not immediately intuitive. The simplest explanation is the one alluded to in the results above: NL neurons do not phase-lock equally well to all frequencies to which they respond. Since elimination of phase-locked responses requires two filters, the presence of a frequency that elicits a non-phase-locked response necessitates the addition of a third filter. While we did not observe any such similar effect in NM, the fact that we see a broadening of STRF bandwidth from NM to NL accounts for this. Another possibility arises from the mechanics of an NL neuron. While this study treats the NL neurons as if their responses were produced by a single family of filters, in reality they are expected to be driven by two sets of filters, one arising from each ear. It may be that nonlinear interactions between the two sides result in the third filter.

Equally surprising are those ICcc neurons that demonstrated only a single significant covariance eigenvector. A strict statistical interpretation would lead to the conclusion that these neurons

responded preferentially to a particular phase and that phase shifted by 180°, but not to intermediate phases. A more likely explanation is that this is the result of poor spectrotemporal tuning. For both the neurons with one significant eigenvector, and those with none, there was no correlation seen with either number of spikes used in the analysis, shape of the peristimulus time histogram, preferred frequency, or any other tuning parameter. In general, these neurons could be very well tuned to ITD, suggesting that there was no deficiency in isolation or recording procedure. The conclusion is that some subset of ICcc neurons has at best very weak spectrotemporal tuning (in our study, 9 of 27). Because there was no correlation of the number of STC eigenvectors with any other studied response parameter, we can not conclude if this represents a distinct subpopulation of neurons or a continuum of tuning within the nucleus.

Another oddity is the lack of correlation between bandwidth and center frequency in NM and ICcc. In the case of NM, this may be a sampling problem. Köppl (1997a) collected many more units over a broader range of frequencies than we were able to collect. The dependence described in that data is somewhat weak, and would be difficult to establish over a smaller subset of frequencies. Explaining the lack of correlation in ICcc is more difficult. Iso-intensity frequency tuning curves reveal a dependence over the same frequency range and with the same number of units. An immediate concern is that this is due to noise in our ICcc data. However, as we will discuss in detail in chapter 4, there is a third method of estimating neuronal response bandwidth using the ITD response curve. For ICcc, this estimate of bandwidth is in good correlation with the STRF bandwidth (fig. 4.7). The conclusion is that iso-intensity frequency response is a poor predictor of a neuron's response to broadband noise, an observation that has also been made in the owl's ILD pathway (Spezio and Takahashi 2003).

Taken all together, our results suggest that the nuclei of the ITD pathway, and specifically ICcc, can be treated as a dynamic spectrum analyzer with a temporal resolution of approximately a millisecond. In humans, the time-dependent envelope of speech is known to be of importance: for example, when the envelope of speech is distorted by a low-pass filter, comprehension is degraded (Drullman, Festen and Plomp 1994). However, it remains unknown if the owl can detect time-

dependent properties of the stimulus on a scale such as observed here. It is known that the owl is capable of discriminating between nearby frequencies with ease (Quine and Konishi 1974), and complex noises that differ by as little as 2.5 dB in a single $\frac{1}{3}$ octave band (Konishi and Kenuk 1975), but the stimuli used in these studies were not characterized in terms of their temporal modulation. Nonetheless, behavior suggests that the owl is capable of subtle tasks of recognition: for example, the owl has been reported to be able to distinguish between the sounds of a mouse walking through dry leaves versus a mouse walking through chopped-up paper (Konishi and Kenuk 1975). Combined with the psychophysical work in humans on the relevant stimulus parameters for sound recognition, it is likely that the owl is capable of detecting spectrotemporal information on a relatively short timescale.

Both NL and ICcc are tuned to ITD as well as to the spectral properties of the signal. In their study of the ICcl, Keller and Takahashi (2000) found that firing patterns were independent of stimulus location, and Carney and Yin (1989) observed the same in the cat's inferior colliculus. While we did not characterize firing patterns directly, we observed that the STRFs estimated at different ITDs for the same neuron in general had near-identical bandwidths, center frequencies, and latencies. It seems clear from this that the STRF is independent of ITD. It is tempting to declare from this that the firing pattern is independent of ITD, but this statement requires a caveat. Specifically, the firing pattern is a function not only of the SRTF, but of the mean firing rate, and the mean firing rate is a function of ITD. Thus, for fixed firing rates, we can say the firing pattern of NL and ICcc neurons are independent of ITD. Similarly, ITD will only be independent of the spectrotemporal structure of the signal if the response is averaged over a sufficiently large period. With these details in mind, however, it seems clear that neurons of NL and ICcc appear to encode ITD with their firing and spectrotemporal structure with their firing patterns.

An interesting observation that arises from this involves the role of phase information. In the auditory nerve and NM, the SRTFs are sharp enough that they encode the phase of the stimulus. It is widely accepted that it is necessary that phase be encoded at this point in the sound localization pathway, as phase-locked inputs are a necessary requirement for the coincidence detection that is the

mechanism for the computation of ITD (Jeffress 1948; Carr and Konishi 1990; Carr 1993). NL also encodes phase, but this is a consequence of coincidence detection: if NL neurons spike only when they receive coincident impulses from NM inputs, then assuming there is no significant jitter in the time required to initiate the spike, the output spikes of the NL neurons will have the same general timing as the phase-locked input spikes, and hence be phase-locked. The fact that the STRFs are generally maintained in ICcc, however, shows that this now apparently superfluous phase information is not discarded. Instead, we see that in the ITD pathway phase encoding is used first to compute the ITD, and then, in a slightly less precise way, to provide information about the spectrotemporal properties of the stimulus, providing an elegant demonstration of encoding multiple stimulus parameters with a limited coding language.

Chapter 4

Relationship between Spectral and ITD Tuning

As discussed in the Introduction, both mammals and birds use the interaural time difference (ITD) to determine the horizontal position of a sound source in space (Heffner and Heffner 1992; Moiseff 1989). The accepted model to explain the computation of ITD uses a combination of coincidence detection and delay lines to produce neurons that fire maximally when the differences in arrival time of the sound at each ear and the neural conduction delays cancel each other exactly (Jeffress 1948). Consistent with this, the axons that project to nucleus laminaris (NL) of the barn owl act as delay lines converging on neurons that act as coincidence detectors (Carr and Konishi 1988, 1990). Even in mammals, where the exact method of encoding ITDs is under question (McAlpine, Jiang and Palmer 2001), the neurons of the medial superior olive (MSO), the mammalian homologue to NL, behave as coincidence detectors (Goldberg and Brown 1969; Yin and Chan 1990).

The coincidence detection model has been argued to be a special case of a more general cross-correlation algorithm (Licklider 1959), which expresses the degree of similarity between two time-dependent signals as a function of lag (how much the signals have been shifted in time with respect to each other). Experiments have verified that the neurons of NL and MSO do show some cross-correlation-like behavior (Carr and Konishi 1990; Yin and Chan 1990). The full extent to which cross-correlation is a description of the behavior of the neurons of NL remains open, and with it, the completeness of the current coincidence detection model.

Cross-correlation in the NL neurons may also be of perceptual relevance outside the context of

localization, as the interaural correlation will be also influenced by non-spatial features of the auditory environment, such as the presence of echoes and multiple sound sources. The sound localization of owls (Saberi et al. 1998) and humans (Blauert and Lindemann 1986) and the firing rate of higher order neurons of the owl's sound localization pathway (Albeck and Konishi 1995) strongly depend on interaural correlation. How the brain detects dissimilarities between the signals that reach each ear may thus involve computing an estimate of the correlation at fixed lag within narrowband frequency channels.

A property of cross-correlation is that it has linearity in the frequency domain: the cross-correlation between two broadband signals is given by the sum of the cross-correlation between the component non-overlapping frequency bands of the signals. This property enters into play in addressing the problem of phase ambiguity in sound localization. Due to the remarkable phase locking ability observed in the primary afferent fibers, barn owls can detect ITD in sounds of frequencies beyond 8 kHz (Sullivan and Konishi 1986; Köppl 1997b). For periodic signals, ITDs that differ by integer multiples of the stimulating signal will produce identical responses, a problem called phase ambiguity. Because of the small range of frequencies to which an NL neuron will respond, even a broadband signal is bandpassed sufficiently that the stimulus is effectively periodic (section 1.5), so that phase ambiguity is an issue for complex noises as well as pure tones. For much of the relevant range, the frequencies used by the owl are sufficiently high that phase ambiguity becomes an issue in the resolution of sound location, in that all ITDs are encoded ambiguously (Mazer 1998; Saberi et al. 1999). If the neurons of NL can be treated as cross-correlators, and hence have the property of frequency-domain linearity, then simple linear integration across frequency bands in the ascending sound localization pathway will result in what is effectively the cross-correlation of the two original binaural stimuli. However, a formal approach to the nature of the frequency interaction within NL has been elusive, due to the difficulties involved in obtaining reliable single-unit recordings at the primary site of cross-correlation. Previous studies have attempted to address the question of whether simple summation over frequency components alone is sufficient to explain the ITD response curves of neurons located at different levels of the pathway, but the results have been ambiguous (Yin, Chan

and Irvine 1986; Mazer 1998; Fujita and Konishi 1991; Mori 1997; Peña and Konishi 2000). These studies compared the response to various narrow-band stimuli to the response to broadband stimuli, which proves problematic due to issues involved in representing a broadband stimulus as a finite sum of narrow-band stimuli. The previously published data leave open the possibility that another, possibly nonlinear, mechanism beyond integration across frequencies is at play in determining the response to broadband stimuli, and hence to assist in the elimination of phase ambiguity. Using reverse correlation, we can avoid this complication. Therefore, our first goal is to determine whether the neurons of NL exhibit frequency-domain linearity, which will serve a two-fold purpose: validating the coincidence detection model, and establishing a baseline for the interaction of frequency bands.

Upon having addressed that question, we will then examine the neurons of ICcc to see if this relationship is maintained. Maintenance of such a relationship would indicate that the cross-correlation of the signals at the two ears in individual frequency bands is of use to the animal, lending further credence to the conclusions of chapter 3 that ICcc may play a role in sound recognition. Alternatively, we might see a disruption of this relationship with the aim of accelerating the suppression of side peaks in the lateral shell.

4.1 Methods

See chapter 2 for general methods, and section 3.1 for details on the use of reverse correlation to calculate the effective stimulus.

4.1.1 Data collection

Data for estimation of monaural filters were obtained using the same protocol as reverse correlation, except that the stimuli were binaurally uncorrelated broadband stimuli. Because of the constraints of the analysis, the target number of spikes with binaurally uncorrelated stimuli was roughly twice that of reverse correlation. In the case of binaurally uncorrelated stimuli, the ITD is by definition undefined, and there were no interleaved stimulus conditions as a result.

4.1.2 Analysis

Power spectral density (PSD) was estimated with the MATLAB implementation of Thomson's multi-taper method with the time-bandwidth product set to 5/2 (for signals sampled at 48,077 Hz) and 5/4 (for signals sampled at 30 μs). In neurons with significant covariance eigenvectors, the PSD was estimated by averaging the PSD of the STA and the significant covariance eigenvectors. We use three characterizations of the spectral properties of a signal. BW_{10} (BW_5) is the 10 dB (5 dB) bandwidth of the signal; that is, it is the width of the peak of the periodogram 10 dB (5 dB) below the peak. CF_{10} (CF_5) is the frequency on which the BW_{10} (BW_5) is centered, and PF is the peak frequency of the periodogram of the signal. Care was taken to ensure that multiple peaks did not confound these measures.

NL

The reverse correlation was done with a 20 ms window at a sampling rate of 48,077 Hz, while the ITD curve was sampled with 30 μs resolution over a range of no more than 4.8 ms. To remove the effects on power spectra estimation that arise due to this difference, the reverse correlation data was down-sampled by a spline to the sampling rate of the ITD curve, and then a time window matching that of the ITD curve was chosen about the maximum absolute value of the STA.

Model of NL response

Binaurally uncorrelated stimuli were presented repeatedly, and then reverse correlation was done twice, using the two families of stimuli (left ear and right ear). The resulting spike-triggered averages were used as estimates of the corresponding monaural filters. The same inhomegenous Poisson process as described in section 3.1 was used to generate spikes. However, the instantaneous firing rate parameter was derived by convolving the stimuli by the appropriate monaural filter, summing them, and then shifting (to represent a threshold) and scaling to obtain the desired mean firing rate. The nonlinear weighting function of Rust et al. (2004) was not used in this case. That method depends on having a set of filters which are orthogonal; this is a condition that is not guaranteed by

our method of filter extraction. The mean firing rate over the reverse correlation dataset was used as a target to determine the spiking threshold by gradient descent for the model STRF; the mean firing rate at the characteristic delay for the ITD curve was used similarly for the model ITD tunign curves. In all cases, only the spikes from 40 ms after stimulus onset to the end of the stimulus were considered. This was done to avoid effects of the onset transient, which the model does not attempt to replicate. In the case of the reverse correlation, the same stimuli were used to produce the model output as were used in the original experiment. In the case of the ITD curve, computing constraints prevented us from saving the original stimuli, and a different set of stimuli generated using the same parameters was used.

ICcc

To remove rectification in the ITD tuning curves of ICcc neurons, a Butterworth bandpass filter was used (MATLAB). The pass-band was centered on the center frequency of the first significant covariance eigenvector (see chapter 3), with a width equal to the center frequency; this is the maximum possible width, as rectification creates new peaks at integer muliples of the center frequency. The stop bands were set to be one-eighth of the center frequency beyond the width of the pass-band. Low- and high-pass attenuation was set to 10 dB, and the ripple parameter was set to 5 dB.

4.2 Results

The responses of NL neurons to ITD contained in periodic and broadband signals differ over an extreme range of ITDs, but are indistinguishable within a physiological range (fig. 4.1). This indicates an interaction between component frequencies within the NL frequency tuning bands, while confirming the existence of the general problem of phase ambiguity.

4.2.1 Nucleus laminaris

If NL neurons do behave as cross-correlators, then the Wiener-Khinchin theorem (Appendix A) predicts a precise relationship between the power spectrum of the ITD response curve and the

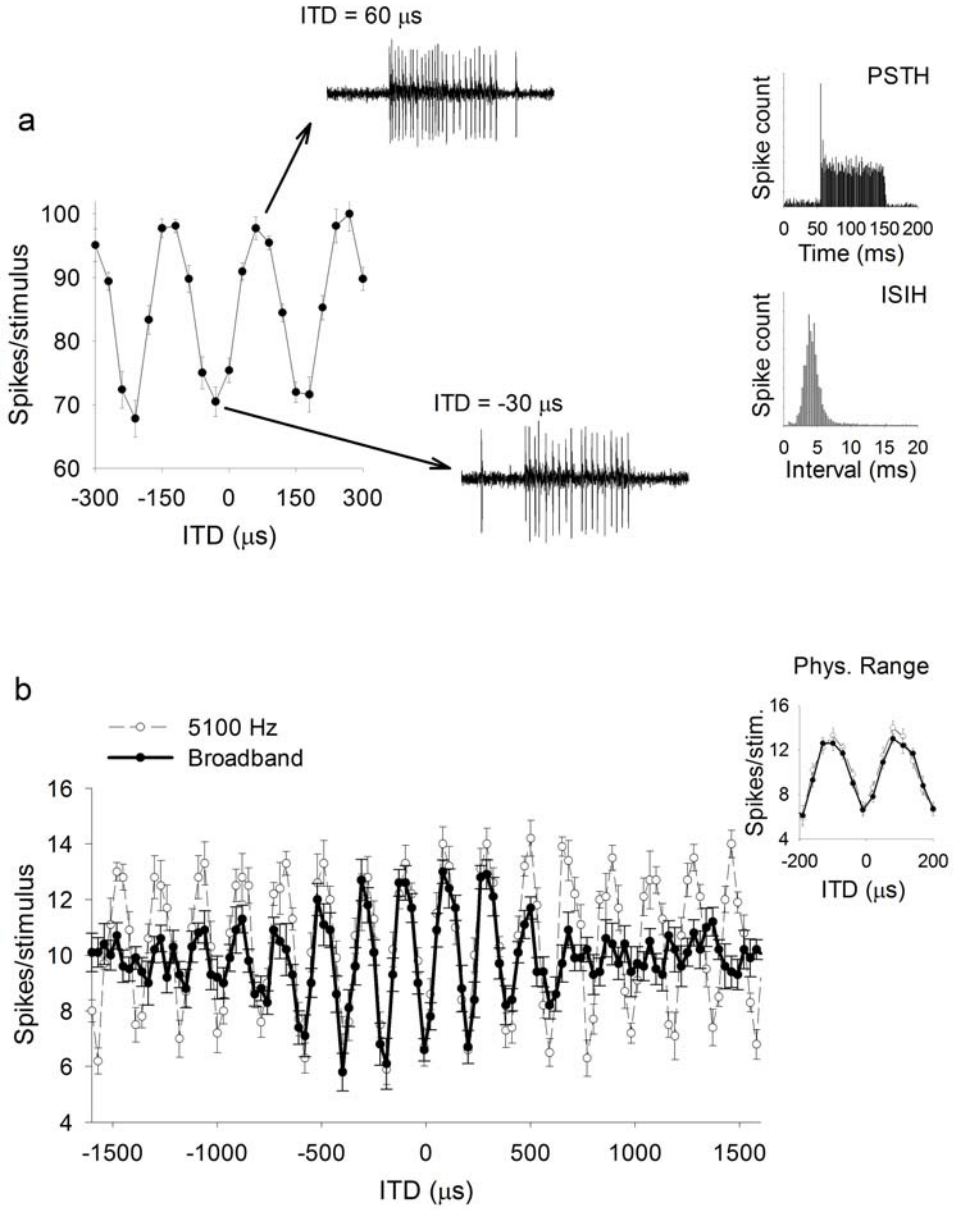


Figure 4.1. ***a***: Laminaris neuron recorded with "loose patch." When the firing rate is plotted as a function of ITD in response to broadband stimuli and the range of ITDs is limited to those that might be encountered under normal listening conditions, the resulting curve can show multiple peaks of similar amplitude, similar to a cosine. Two example traces show the response to favorable and unfavorable ITDs indicated by the arrows. The post-stimulus time histogram (PSTH) and interspike interval histogram (ISIH) are shown on the right. ***b***: When the range of ITDs presented is expanded, a more complex structure becomes evident. Whereas the response is periodic for tonal stimulation (*thin line*), the amplitude of the peaks decay when broadband sound is used as a result of the response to a narrow-band (*bold line*). For the physiological range (Phys. Range; inset) the response to tones and broadband signals is practically indistinguishable.

power spectrum of the square of the Fourier transform of the effective stimulus: the portion of the stimulus that passes the various cochlear and neural filters and succeeds in stimulating the neuron (for a justification of the applicability of the Wiener-Khinchin theorem under these conditions, see Appendix B). By using a loose-patch technique (Peña et al. 1996, 2001) we were able to obtain the quality and stability of the recordings (fig. 4.1) necessary to apply reverse correlation analysis (de Boer and de Jongh 1978; Klein et al. 2000) to estimate the effective stimulus. The power spectral densities of both the spike-triggered average and the significant eigenvectors of the spike-triggered covariance were averaged and compared to the power spectral density of the ITD response curve (fig. 4.2), using the 5 dB bandwidth and the 10 dB bandwidth respectively[1]. The correlation was good, with a linear regression near unity (regression $0.86x+363$ Hz, $r^2 = 0.82$, $n = 18$). A large offset between the two is explained by the fact that the ITD response curve was sampled at much lower resolution than the reverse correlation data; when the reverse correlation data was down-sampled to account for this, the offset dropped dramatically (fig. 4.3; regression: $1.03x + 60.4$ Hz, $r^2 = 0.87$, $n = 17$). It was necessary to remove one high frequency unit (CF_{10} of STA = 6,500 Hz) because the sampling rate compensation destroyed the structure of the STA.

We also compared the PSD of ITD curves of 77 neurons with the iso-intensity frequency-tuning curve, which is a plot of mean spike rates as a function of frequency for the same sound intensity. Correlation was observed as in the previous comparison, though considerably weaker, in the sense that the slope was less than unity (fig. 4.4). Squaring the iso-intensity frequency-tuning curve as per the Wiener-Khinchin theorem did not improve this. Since the iso-intensity frequency-tuning curve is known to be a poor estimate of the effective spectral response of the neuron (Spezio and Takahashi 2003), this lower degree of correlation is to be expected.

4.2.2 Model of nucleus laminaris response

Our findings support the theory that the neurons of NL can be treated as cross-correlators. However, working in MSO, Batra and Yin (2004) argued that those coincidence detector neurons did not

[1]Since dB are a logarithmic unit, comparing the bandwidths at 5 dB and 10 dB is equivalent to comparing the PSD of the ITD curve to the PSD of the square of the Fourier transform of the effective stimulus.

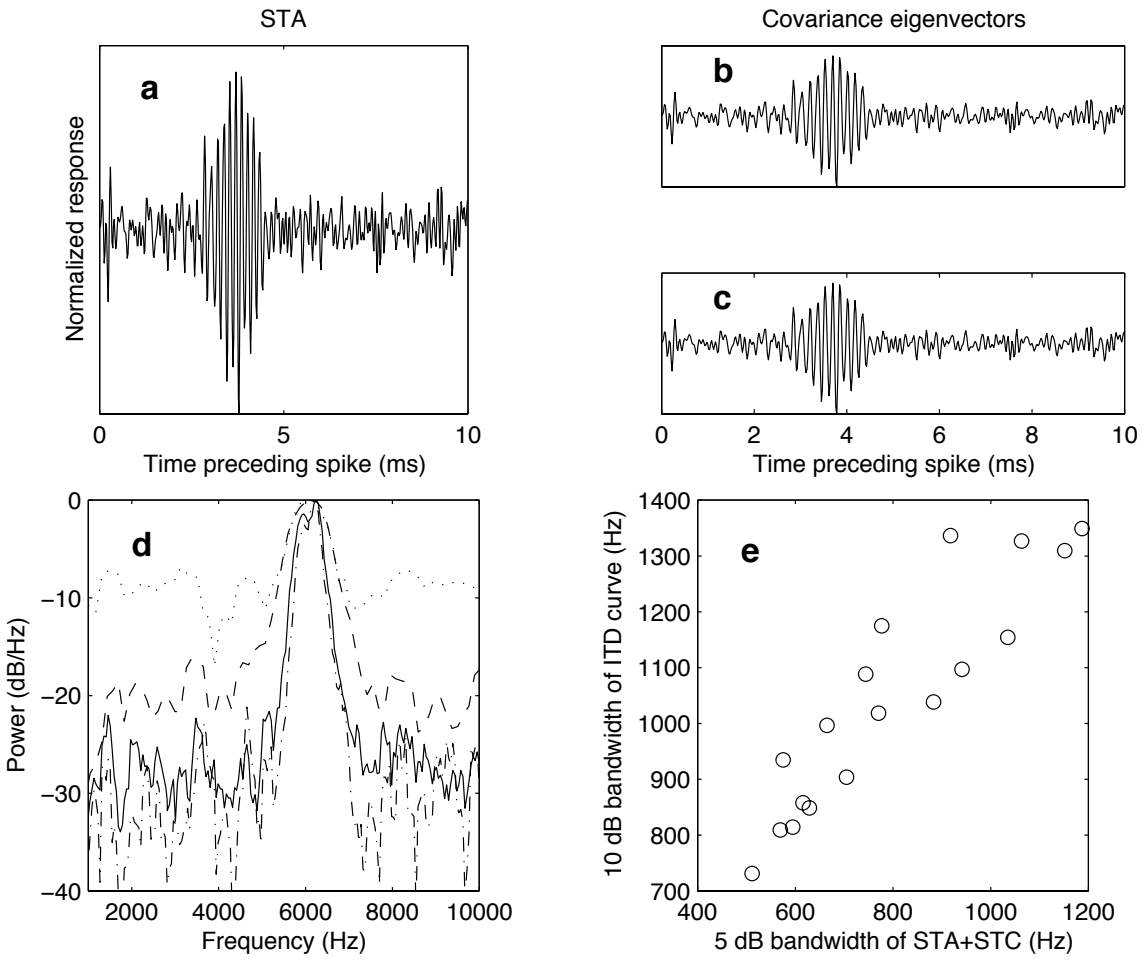


Figure 4.2. Some neurons had not only a spike-triggered average (STA; ***a***), but as many as two significant and coherent eigenvectors (Covariance eigenvector) of the covariance matrix (***b*** and ***c***). The power spectral density (PSD) of the STA alone (***d***; *dashed-and-dotted line*) is broadened by incorporating the covariance eigenvectors (***d***; *solid line*). By adjusting the sampling resolution, the PSD of the reverse correlation (***d***; *dashed line*) matches the halved PSD of the ITD curve (***d***; *dotted line*) through to the –10 dB mark. The BW_5 of the PSD of the STA plus covariance, without sampling rate compensation, is shown plotted against the BW_{10} bandwidth of the ITD curve in ***e***.

behave as ideal cross-correlators. Based on the relationship between the phase locking of the MSO neurons to monaural and to binaural stimuli, they observed that the phase locking in the binaural cases was less than would be predicted from the monaural responses. In an attempt to resolve these discrepancies, we collected additional data to test a model of ITD response based on the coincidence detection hypothesis.

The premise of the hypothesis is simple in concept: the effective time constant of the coincidence detectors is fast enough such that the spiking threshold is attained only when a certain number of

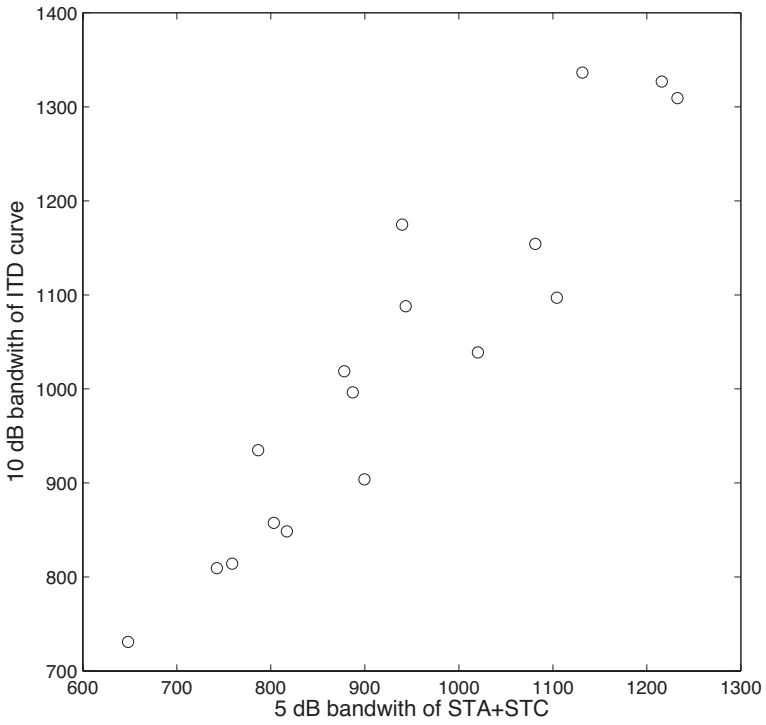


Figure 4.3. A plot of the BW_5 of the PSD of the down-sampled reverse correlation (STA and STC included) against the BW_{10} of the ITD curve. The correlation is improved over that seen in fig. 4.2.

excitatory post-synaptic potentials (EPSPs) occur within a few hundred microseeconds of each other. These EPSPs come from two populations of afferent neurons, one for each ear, and the instantaneous firing rate of each afferent neuron can be approximated as a linear filter of the relevant monaural stimulus (de Boer and de Jongh 1978). NL neurons tuned to frequencies above 2 kHz have been described as having very little dendritic arborization (Carr and Konishi 1990; Carr and Boudreau 1993). The lack of a dendritic tree suggests little opportunity for complex inter-synaptic interactions. If these assumptions are true, it should be possible to estimate the response of NL neurons to arbitrary stimuli using only two linear filters, one for the population of afferent neurons relaying information for each ear.

To estimate the monaural filters, in six neurons we presented binaurally uncorrelated stimuli. The coincidence detector model predicts that the resulting spike train contains spikes that can be grouped into one of three categories: those resulting from coincidences between spikes from the left-side monaural input population, those elicited by coincidences in the right-side monaural input

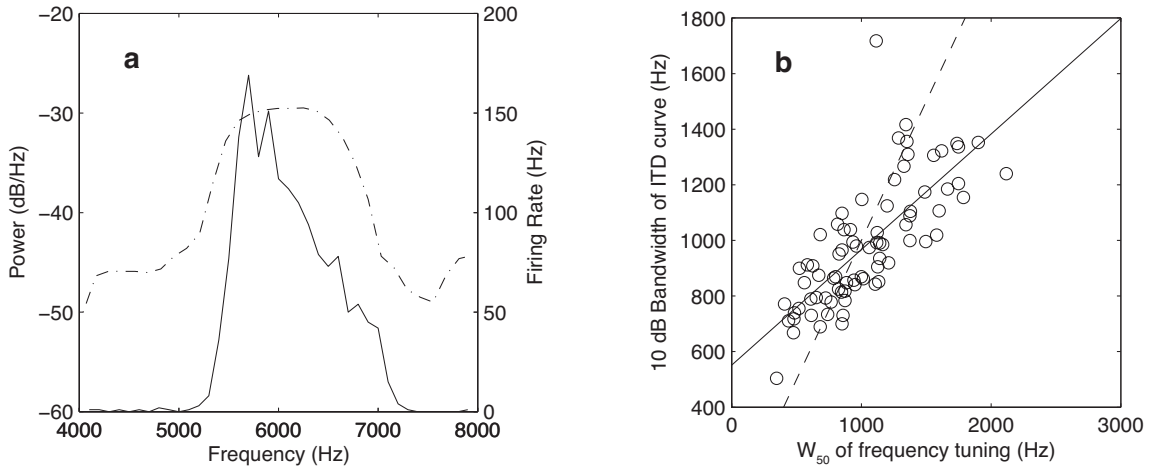


Figure 4.4. ***a***: While the PSD of the ITD curve (*dashed line*) and the iso-intensity frequency tuning curve (FTC; *solid line*) cover the same approximate frequency range, the FTC is asymmetric and more sharply peaked. ***b***: The correlation between the BW_{10} of the PSD of the ITD curve and the W_{50} of the FTC at half-height gives a slope considerably less than unity, with the regression line (*solid*) intersecting the unity line (*dashed*) within the data set.

population, and those arising from coincidences between the two channels. In a spike-triggered average between the spike train and the left-side stimulus, only the first category of spikes will contribute in a consistent manner; all other spikes are uncorrelated with the left-side stimulus by construction, and will be eliminated by the averaging process. By estimating the monaural filters in this manner rather than by using monaural stimulation, we avoid the possible confound of leaving the unstimulated side in an undefined state.

These filters were then used in the model to generate both the spike-triggered average and the ITD response curves to binaurally correlated sound. Both the spike-triggered averages and the ITD curves of the data and the model were good matches in all case (fig. 4.5a–f). In addition, the relationship between the PSDs of the modeled ITD curves and the PSDs of the modeled STAs (fig. 4.5g) are comparable to the relationship seen in the collected data, within the constraints of the small sample size (regression: $0.73x + 477$ Hz, $n = 6$). In three of the neurons (fig. 4.5c, f, e), a latency shift in the STA is apparent between model and data. The model's latency is always consistent with the latency as predicted by the individual monaural filters. Moreover, this latency shift does not affect the model's capability to predict the tuning to ITD. As the data used for the

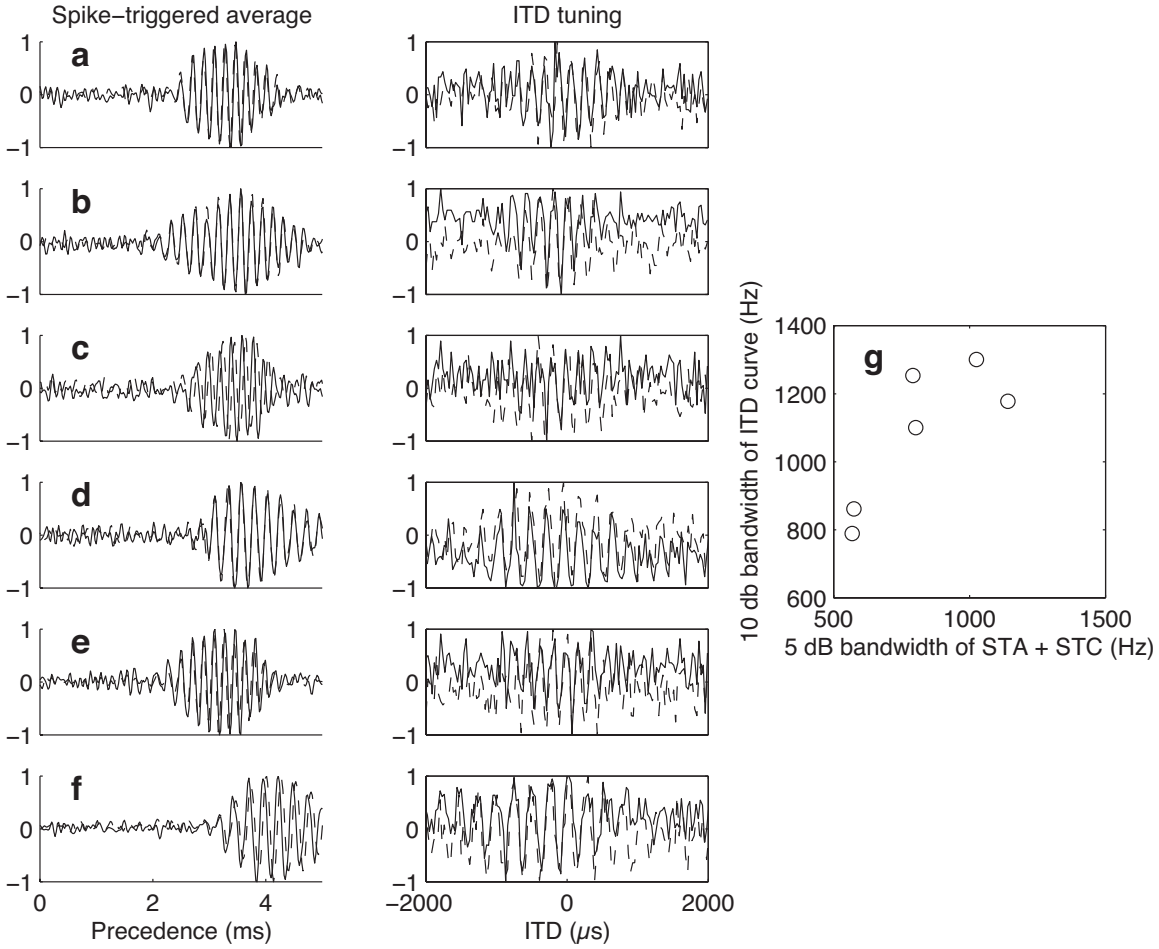


Figure 4.5. The spike-triggered averages of the data and the model show high correspondence (*left column*), beyond the latency shift in neurons $\boldsymbol{c}$, $\boldsymbol{e}$, and $\boldsymbol{f}$, as mentioned in the text. The normalized ITD curves of the six neurons (*dashed lines*) correspond well with the normalized ITD curves of the corresponding models (*solid*) for those ITDs in which coherent tuning is apparent (*right column*). The comparison of the PSDs of the STAs and the ITD curves of the model is shown in $\boldsymbol{g}$, as in fig. 4.2.

estimation of the STA and the binaurally uncorrelated data set were collected with as much as half an hour interval, changes in the amplitude of the spikes may explain this.

The shapes of the predicted and collected ITD curves were also reasonably similar, including the location of the CDs and the rate of decay at extreme ITDs. The peak-to-trough ratios were highly dissimilar between data and model. The model ITD curves fell to a firing rate of 0 spikes per second for unfavorable ITDs, while the data generally saw a low maximal peak-to-trough difference (median 66 spikes per second), with a minimal firing rate well above 0 spikes per second.

4.2.3 ICcc

In ICcc, a majority of the neurons studied show a clear effect of rectification when considered over a broad range of ITDs (18 of 28). This rectification has a significant effect on the PSD of the neurons (fig. 4.6). As a result, the Wiener-Khinchin theorem cannot properly be said to apply to ICcc neu-

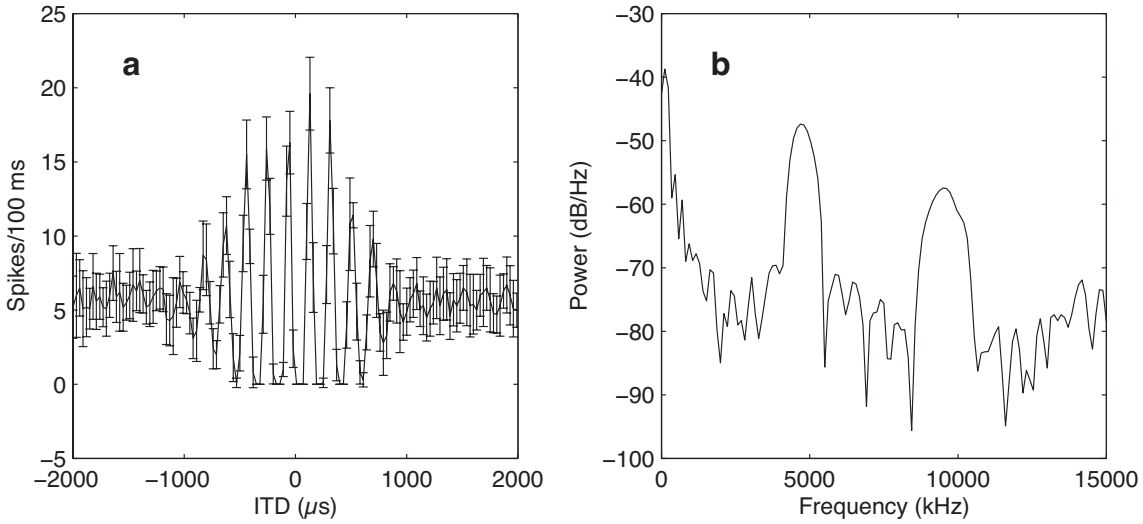


Figure 4.6. ***a***: An example ICcc tuning curve over an extended range of ITDs. The half-wave rectification of this example is clear. ***b***: The PSD of the ITD tuning curve in ***a***. The low-frequency component is due to the non-zero mean, an effect which rectification makes more difficult to remove. A more serious effect is the emergence of additional peaks in the PSD at integer multiples of the base frequency of the curve (in this case, approximately 5,200 Hz).

rons. Intuitively, there is no way to use knowledge of the inputs alone to derive the threshold of the neuron. Mathematically speaking, the Wiener-Khinchin theorem describes a particular input-output relationship; even if this relationship holds between afferent inputs and the sub-threshold membrane potential, the rectification that occurs in the spiking output represents an extra, unpredicted stage in the input-output relationship.

This is somewhat unsatisfying, and it is reasonable to ask: in the absence of rectification, does the Wiener-Khinchin theorem apply? It is possible to attempt a reconstruction of the unrectified curve by bandpassing the rectified ITD tuning curve (fig. 4.7a, b). When this is done, and the analysis done above repeated, a strong linear correlation is again observed (fig. 4.7c; regression: $1.28x - 14.9$ Hz,

$r^2 = 0.77$, $p < 10^{-6}$, $n = 24$).[2] This suggests that ignoring the effects of rectification, the Wiener-Khinchin result holds in ICcc. Presumably intracellular recordings that studied the subthreshold membrane potential would directly confirm this finding.

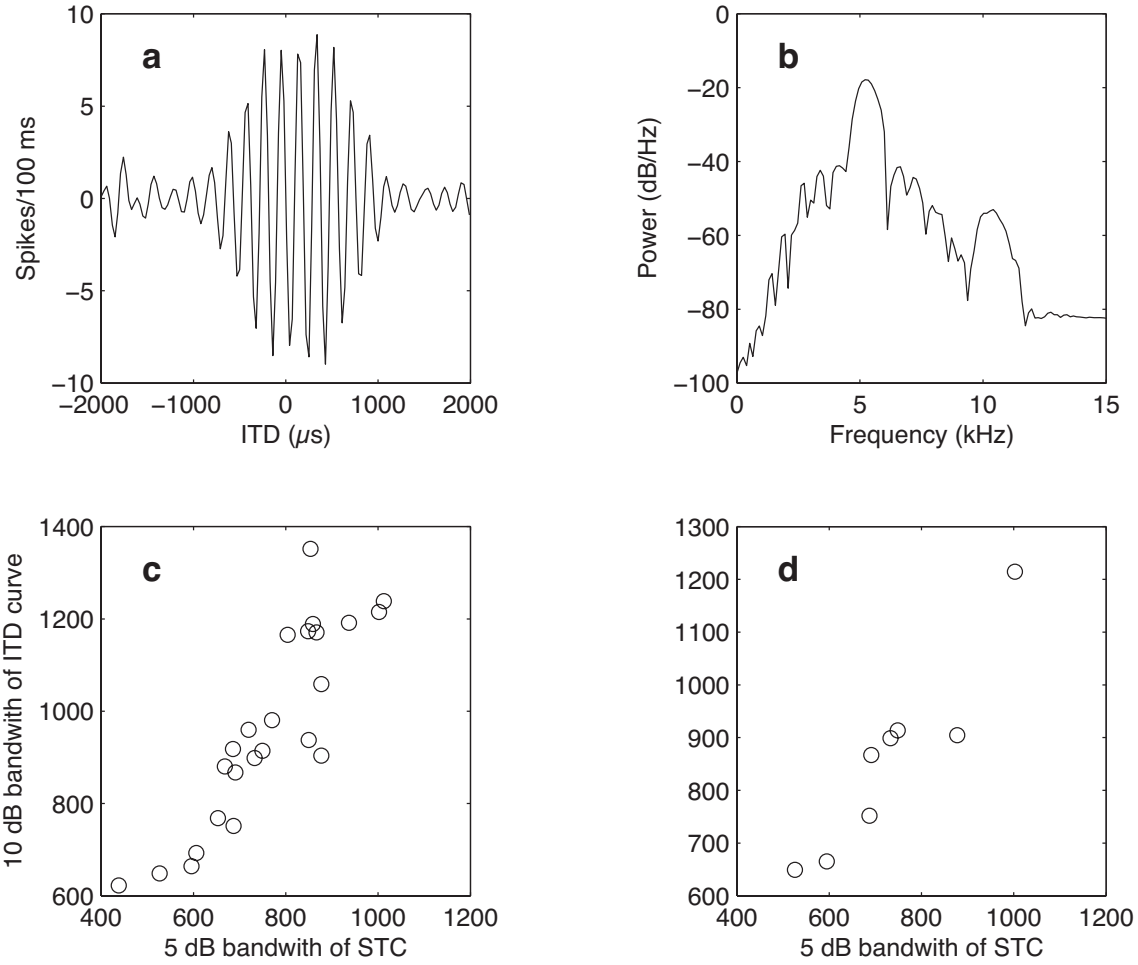


Figure 4.7. ***a***: The ITD tuning curve from fig. 4.6 has been bandpassed to remove the effects of the rectification; in order to accomplish this, "negative firing rates" are introduced. The PSD of the new curve (***b***) shows only a single strong peak at the CF of the tuning curve. ***c***: Once all of the ITD curves have been bandpassed, it is possible to plot the 5 dB bandwidth of the PSD of the reverse correlation against the 10 dB bandwidth of the PSD of the ITD curve. A strong linear correlation is apparent. ***d***: The ITD curves that do not show rectification demonstrate the same linear relationship without the need to bandpass. Both ***c*** and ***d*** include the sampling rate compensation.

To confirm that the observed result was not simply an artifact of the band-passing, the analysis was restricted to the eight neurons with significant covariance chosen by eye to show no effect of

[2]For one of the ICcc neurons, no reverse correlation data was available; only those neurons which had at least one significant covariance eigenvector were included (see chapter 3).

rectification, and the bandpassing omitted. No significant change in the correlation was observed (fig. 4.7d; regression: $1.12x + 37.9$ Hz, $r^2 = 0.88$, $p < 10^{-3}$, $n = 8$).

Neurons of ICcc were in general only weakly responsive to binaurally uncorrelated stimuli, making it impossible to collect enough spikes to attempt to estimate "monaural" filters. Additionally, the neurons of ICcc are not coincidence detectors, and the logic underlying the model used for NL neurons above does not properly apply. As such, no attempt was made to directly model the responses of ICcc.

4.2.4 Envelope coding

According to the cross-correlation hypothesis, the ITD tuning curve of a neuron generated using anti-correlated stimuli, e.g., stimulation where the signal to one ear has been inverted, should be the inversion of the ITD tuning curve generated using correlated signals (Yin, Chan and Carney 1987). In the mammalian IC, there exist neurons which do not demonstrate this effect (Joris 2003), indicative of locking to the envelope of the stimulus, rather than to the phase.

To test whether this might be the case in the owl as well, we collected ITD tuning curves using anti-correlated stimuli for 8 NL neurons and 17 ICcc neurons, and then computed the cross-correlation between these curves and the standard ITD tuning curves. A significant positive correlation would be a clear demonstration of envelope coding, and significant negative correlation is the prediction of the straight cross-correlation hypothesis. Correlations that are not significant might be examples of mixed coding (Joris 2003). In all the NL neurons, the correlated and anti-correlated ITD tuning curves were negatively correlated ($p < 10^{-10}$), and the same was true in ICcc ($p < 10^{-5}$). Thus there is no evidence for envelope coding in the owl's ITD pathway, though the possibility cannot be ruled out that such coding is used at the highest frequencies in the owl's hearing range, which this study did not sample (highest CF_{10}: 7,000 Hz). As it stands, these results further validate the cross-correlation hypothesis.

4.3 Discussion

These results suggest that at the locus of computation of the ITD in the barn owl, no additional mechanisms for the suppression of non-CD ITDs are at play and their behavior can be characterized as the cross-correlation between the stimuli presented to the two ears, up to a scale and a shift. Consistent with the theory presented here, Yin, Chan and Irvine (1986), proceeding from an assumption of classic cross-correlation behavior, showed a correspondence between the sync-rate curve (the rate-frequency response curve weighted by the synchronization coefficient) and the Fourier transform of the ITD curve. While direct comparison is difficult, it appears that their results are a close match for the iso-intensity frequency tuning curve data presented here. Given that their sync-rate curve shares many of the limitations in estimating the effective frequency tuning of iso-intensity frequency tuning curves, it is likely that the relationship between spectral and ITD tuning also holds for the neurons they examined. Yin, Chan and Irvine did their work in the inferior colliculus of cats, which does not feature coincidence detector neurons; instead, it receives input from the coincidence detectors of MSO, and is properly homologous to ICcc.

The results from ICcc are somewhat more ambiguous. The emergence of rectification results in the Wiener-Khinchin theorem being violated, and hence, strictly speaking, cross-correlation behavior is no longer maintained in ICcc. At the same time, there is no evidence that there are any mechanisms disturbing the spectral-ITD relationship at play other than rectification, which might have been introduced to accomplish, e.g., an accelerated suppression of side-peaks to deal with phase ambiguity. In fact, it is not immediately clear that the rectification should be considered of significance at all. Previous reports have not reported this rectification in the barn owl's ICcc (Wagner, Takashi and Konishi 1987, 2002). This is because those studies restricted themselves to the range of ITDs that the owl can actually experience under free-field conditions. Under those conditions, rectification is not apparent (fig. 4.8). Thus, over the short ranges of cross-correlation lag that the auditory system implements, the effects of this rectification are essentially negligible. The implications of rectification in ICcc in the context of coding will be discussed in more detail in chapter 5. Regardless, what our data demonstrates is that up to some degree of rectification due to thresholding, the response of

ICcc neurons to ITD is determined by the same relationship to their spectral tuning as we seen in NL.

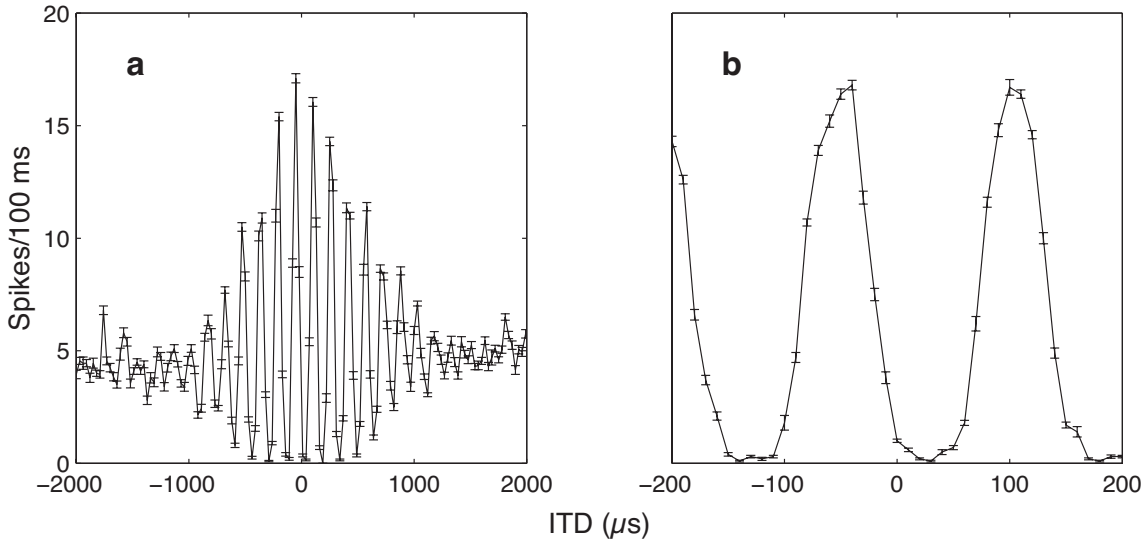


Figure 4.8. ***a***: Another example of a ICcc neuron whose ITD tuning curve shows rectification. ***b***: The ITD tuning curve of ***a*** restricted to ITDs of ± 200 μs. Without the reference point of the tails of the curve in ***a***, there is no clear sign of rectification; the relative broadness of the troughs is easily attributed to noise.

The model presented here is the first model of NL response that works with stimuli other than pure tones. While models of coincidence detection have been designed that work for broadband stimuli (for example, see Cai, Carney and Colburn 1998; Hancock and Delgutte 2004), they are are generally based on data from the mammalian inferior colliculus. A crucial difference in our model from other models of NL is our approach. Previous models have used cellular models, either integrate-and-fire (Gerstner et al. 1996) or compartmental (Grau-Serrat, Carr and Simon 2003) methods. Because there is as of yet no intracellular data on the NL neurons from the adult barn owl, the parameters of these models are largely based on data from the chick. We instead chose to use the simplest possible model based on the abstract description of the coincidence detection hypothesis. By comparing this model to the data, it is possible to determine in what areas it is necessary to introduce additional levels of detail to account for the data.

Our model of nucleus laminaris response consistently produced lower firing rates at trough ITDs than observed in the data. This is consistent with the observation of Batra and Yin (2004) that

the neurons of MSO were likely in the case of binaural stimulation to produce a relatively large number of spurious coincidences. These spurious coincidences would be expected to dominate at unfavorable ITDs. The model does not have a spiking representation of the input, which eliminates this stochastic effect, and hence the troughs show little to no activity. The fact that the shape of the ITD curve is the same in both model and data, however, indicates the contribution of these spurious coincidences is constant across ITD. Even though there is an elevated rate of spurious coincidences in the data, the ITD curve remains a shifted and scaled version of the ITD curve produced by the perfect coincidence detection model, and hence the relationship between ITD and firing rate can still be treated as cross-correlation, though with reduced signal power. This reduction could in theory be easily overcome through the use of population responses.

Approximating cross-correlation using an additive mechanism is known to be problematic in general. One issue is that addition is sensitive to the relative magnitudes of the inputs. However, NL is insensitive to differences in intensity in the inputs if either broadband noise or the CF of the neuron was used. This balancing of the inter-aural intensity difference (IID), thought to be accomplished via the superior olive (Viete, Peña and Konishi 1997), addresses that concern. The presence of this feedback loop through the superior olive also provides motivation to be careful when comparing monaural and binaural responses, as in Batra and Yin (2004). When presenting a monaural stimulus the IID is not well defined, and it is not clear what effect the IID compensation circuit might have.

A more pressing matter with an additive mechanism is that it does not properly deal with the sign of the signals: for example, synchronous large negative deflections will be reported as anti-correlated activity. However, a strong argument against the use of a multiplicative mechanism can be found in the fact that the spikes of NL neurons are locked to the phase of the auditory stimulus (which explains the coherent spike-triggered average in fig. 4.2; Carr and Konishi 1990). Multiplication of two sinusoid signals leads to the creation of a new sinusoid with different component frequencies, which would result in phase locking to frequencies not present in the input. Consistent with this, attempts to recreate the spike-triggered average using a model that multiplied together the two

filtered monaural waveforms failed. All together, our results indicate that a simple additive neural mechanism that does not distinguish between ipsi- and contra-lateral inputs is sufficient to implement a coincidence detector that can be treated as a cross-correlator.

Chapter 5

Tuning to Interaural Time Difference

To reiterate the key points of the Introduction, the hind- and midbrain nuclei of the barn owl responsible for the computation of sound location have been well-characterized anatomically (Takahashi and Konishi 1988b, 1988a). Two segregated pathways process the two primary binaural cues, ITD and the interaural level difference (ILD; Sullivan and Konishi 1984). The neurons of the nucleus laminaris (NL) act as coincidence detectors that perform the basic computation for the neural coding of ITD (Carr and Konishi 1990). Because cochlear hair cells respond to only a limited range of frequencies, and because the narrowness of this frequency tuning bandwidth is maintained through to NL, NL neurons respond equally well to ITDs that are phase-equivalent within their frequency tuning range (Carr and Konishi 1990; Peña and Konishi 2000). Thus, a given firing rate can correspond to multiple ITDs, and the true ITD cannot be determined from the response of a single NL neuron. Only in the nucleus ICx, following convergence across frequency bands, is ITD encoded unambiguously (Takahashi and Konishi 1986).

While NL is the initial locus of the computation of ITD, it is not the last nucleus in the exclusively ITD-responsive pathway. It projects both to ICcc, the terminus of the ITD pathway (Takahashi and Konishi 1988b), and to the dorsal lateral lemniscal nucleus, pars anterior (LLDa, previously referred to as VLVa; Takahashi and Konishi 1988a). LLDa in turn projects to ICcc (Moiseff and Konishi 1983) as well as the nucleus basalis (Wild, Kubke and Carr 2001). There are no reports that LLDa receives inputs other than from NL. Cohen, Miller and Knudsen (1998) reported a feedback projection from

the forebrain auditory arcopallium to the inferior colliculus, but they did not differentiate ICcc from the lateral shell, which also receives input from the ILD pathway and is physiologically distinct. Previous work has described the response to ITD seen in ICcc as similar to that observed in NL (Wagner, Takashi and Konishi 1987, 2002). In particular, there is no published evidence of frequency convergence at this stage of processing.

In chapter 3, we established that both NL and ICcc encode information about the spectrotemporal properties of the stimulus, and that they do so with similar resolution in both temporal and spectral domains. This spectrotemporal tuning is also the primary determining factor in the shape of the ITD tuning curve (chapter 4). Thus, our results so far have served to quantify the general observation that NL and ICcc have similar response properties. What remains to be shown is some domain in which their responses are dissimilar, and hence allows us to deduce a role for the post-laminaris ITD pathway.

To accomplish this, we here compare the ITD response of NL and ICcc neurons independent of considerations of spectral tuning. Single units were recorded in both nuclei, and ITD tuning curves using broadband noise were collected for each unit over a broad range of ITDs. We observed that with NL neurons it was necessary to average the responses of upwards of five broadband noise samples per ITD in order for visible structure in the ITD tuning curve to emerge. Conversely, ICcc neurons required only one or two noise samples per ITD, indicating an increase in response reliability. At the same time, ICcc neurons had a greater maximal peak-to-trough response difference than seen in NL. Since the variability of firing rate for each ITD was approximately the same across both the NL and ICcc populations, this results in an improved ability of a single unit to discriminate between nearby ITDs.

5.1 Methods

See chapter 2 for general methods.

5.1.1 Analysis

An ITD tuning curve is generated in the following manner. To calculate the mean firing rate at a particular ITD, a family of unique broadband stimuli (5–10) is presented, each stimulus having the same ITD, and the number of spikes generated for each stimulus is averaged. A single stimulus presentation is referred to as a *trial*, and a full set of trials, one per ITD, is referred to as a *block*. Trials are presented in random order within blocks; blocks are repeated until the desired number of trials is achieved.

To examining convergence to the mean, only those neurons for which we were able to collect 10 blocks were included. We first computed the ITD curve generated by averaging across ten trials per ITD condition (ITD_{10}), and then normalized that by a factor $F_{\mathrm{ITD}_{10}}$ to have a root-mean-square of 1. The error between ITD_{10} and ITD_n (where ITD_n is the ITD curve generated by averaging across a randomly selected subset of n trials per ITD condition, and then normalizing by $F_{\mathrm{ITD}_{10}}$) is given by $E_n = \sum_i (\mathrm{ITD}_{10}(i) - \mathrm{ITD}_n(i))^2$, where $\mathrm{ITD}_n(i)$ is the mean firing rate of ITD_n at the ITD i. For each n, E_n was recomputed fifty times using different subsets of trials in ITD_n to give an estimate of the mean and standard deviations. The resulting curve was then fit by an exponential of the form $c_1 e^{(n-1)/c_0}$, $n = 0 \ldots 9$. By design, c_1 is approximately E_1, and thus has no relevance to the question of rate of convergence to the mean.

We estimated the single-unit Fisher information as $I_F(i) = (\mathrm{ITD}'_{10}(i))^2/(T \cdot \mathrm{ITD}_{10}(i))$, where T is the duration of the stimulus (Dayan and Abbott 2001). ITD'_{10}, the derivative of ITD_{10} with respect to ITD, was estimated using the `gradient` function of MATLAB.

The rectification index (RI) of an ITD tuning curve is computed by first normalizing the ITD curve to have values between 0 and 1. Then, the response to the 30 ITDs with the largest absolute value (15 each from positive and negative ITDs) was averaged. A RI of 0.5 indicates no rectification; a RI of 0 is consistent with negative half-wave rectification, while a RI of 1 would be positive half-wave rectification.

5.1.2 Modeling

To model the response of an ICcc neuron, we used a leaky integrate-and-fire (IAF) model (Koch 1999). The membrane voltage of the model is given by:

$$V(t) = \frac{1}{R}I(t) - \frac{1}{\tau_m}\frac{dV(t)}{dt}$$

where $I(t)$ is the input current, R is the membrane resistance, and $\tau_m = RC$ is the membrane time constant (resistance times capacitance). Since we are not interested in the absolute values of $V(t)$, we can without loss of generality set $R = 1$. Thus, the only parameter of interest is the membrane time constant, τ_m. When $V(t)$ exceeds a threshold voltage V_{th} (determined by gradient descent to obtain the desired firing rate), a spike is initiated, the membrane voltage is reset to 0, and is held at 0 for a refractory period T_{ref}. In our model, $T_{\text{ref}} = 1$ ms, and the target firing rate was 100 spikes/second.

The input current $I(t)$ is given by

$$I(t) = \sum_{n=1}^{N}\sum_{i} g(t - t_{n_i})$$

where N is the number of afferents, t_{n_i} is the time of the ith spike of the nth afferent, and

$$g(t) = \begin{cases} \frac{1}{\tau_s}e^{-t/\tau_s} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

where τ_s is the synaptic time constant. The t_{n_i} are generated using the model of NL response that was presented in chapter 4, with a threshold set to produce a mean firing rate of 100 spikes/second. By presenting the same stimuli repeatedly to the NL model N times, we simulate a family of identically-tuned afferents. Sampling of $I(t)$ and $V(t)$ was done at 100 kHz.

Details of the computation of the shuffled auto-correlogram (SAC) can be found in section 3.1.

5.2 Results

We collected long-range (at least ±2000 μs) ITD curves for 31 NL and 28 ICcc neurons (fig. 5.1a, b). In both groups of curves, the amplitude of the peaks that occur within the owl's physiological range (±180 μs; Moiseff and Konishi 1981) are of similar amplitude, indicative of the narrow frequency tuning and the absence of frequency convergence that have been mentioned in previous work (Mazer 1995; Wagner, Mazer and von Campenhausen 2002). However, previous studies have not had access to comparable data from both nuclei, and hence did not quantify this observation. Plots of the W_{50} of the iso-intensity frequency tuning curve versus F_{50} for the 28 ICcc neurons and 80 NL neurons (including 49 previously recorded neurons which used 50 ms stimuli and are thus not included in any other analyses; fig. 5.1c) showed a similar distribution, consistent with a lack of frequency convergence. The slope of the linear regression for each data set falls within the 95% confidence interval of the linear regression of the other data set. In both nuclei, a significant dependence of bandwidth on the center frequency ($p < 10^{-4}$) was observed.

5.2.1 Evolution of tuning across stimulus repetitions

We compared the reliability of the ITD coding in both nuclei by examining the dependence of the mean response on the number of stimulus presentations. NL neurons' fluctuations in firing rate for a single stimulus presentation, whether due to noise or a dependence on instantaneous spectral properties, mask the fluctuations in firing rate due to the ITD of the signal. Only with averaging across multiple trials, an act assumed to be equivalent to averaging across a greater period of time or a number of similarly tuned neurons, does the dependence of firing rate on ITD become apparent. Conversely, in ICcc the dependence is clear with only a single stimulus presentation (fig. 5.2). This increase in reliability results in fewer ICcc neurons being required to encode ITD in a shorter expanse of time.

To quantify the results observed in figure 5.2, we compared the sum squared error (SSE) between the ITD tuning curve generated by using n trials per ITD ($n = 1 \ldots 9$) and the ITD tuning curve generated with the full set of 10 trials (fig. 5.3a, b). The absolute magnitudes of these numbers are

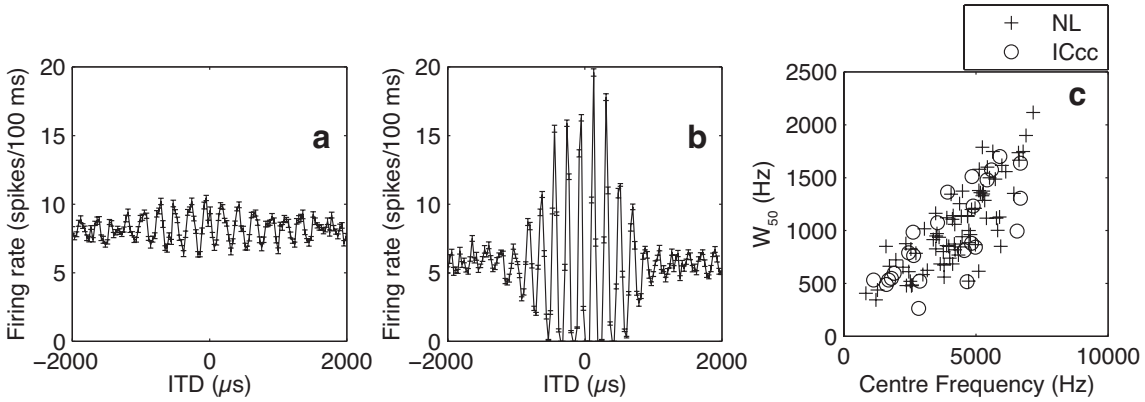


Figure 5.1. Examples of NL (***a***) and ICcc (***b***) ITD tuning curves. Note the difference in the dynamic range of response, as well as the similarity in the amplitude of the error bars (standard error of the mean). Consistent with a lack of frequency convergence, when only those portions of the curves within ± 200 μs are considered, the peaks of both the NL and ICcc tuning curves show no sign of relative suppression. ***c***: Width at half-height of the iso-intensity frequency tuning curve plotted against center frequency for NL (+) and ICcc ($\circ$). The distributions are not different.

not of interest, as they are primarily influenced by the range of values over which the ITD curve varies (a point we will address in the following section). The parameter of interest is the rate at which these curves fall to zero. A rapid convergence to saturation at zero indicates that averaging over a few trials is indistinguishable from averaging over the full set of ten. Conversely, a slow convergence indicates that even later trials contribute to the overall shape of the tuning curve. This rate of convergence was established by fitting each of the convergence curves with an exponential. The time constants for 24 NL neurons and 26 ICcc neurons[1] are summarized in fig. 5.3c. The median time constants of the two populations are significantly different ($p < 10^{-6}$, Kruskal-Wallis), indicating that ICcc neurons converge to their mean ITD curve with fewer trials per ITD than NL neurons.

In this analysis, we used an extreme range of ITDs that greatly exceeds the physiological range. This raises concerns that the difference in tuning might reflect only changes in the extreme periphery

[1]For some of the neurons in both nuclei, we were unable to obtain a full ten trials per ITD, hence the reduced numbers here.

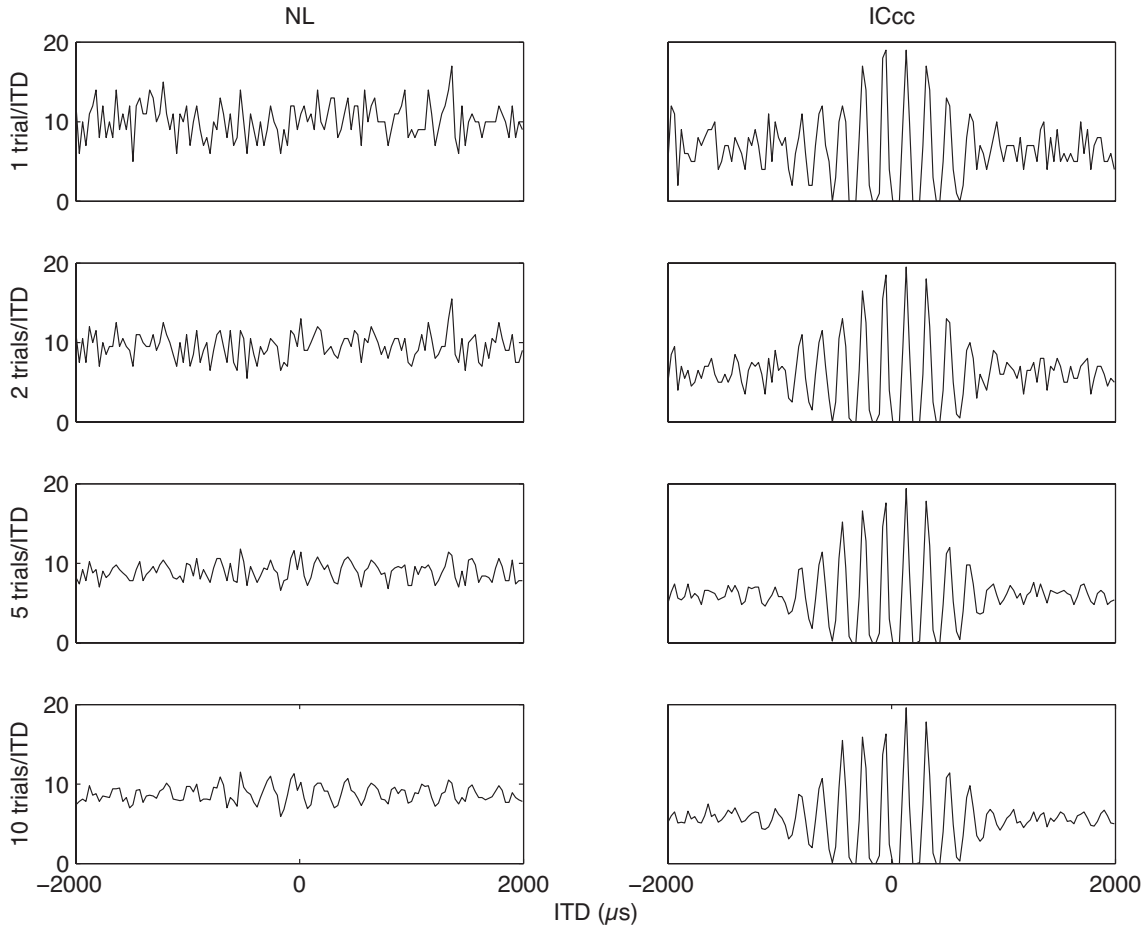


Figure 5.2. The same two ITD tuning curves as in figure 5.1, but here broken down by the number of separate broadband noise samples (trials) presented in order to generate the mean response per ITD. The y-axis has units of spikes/100 ms; error bars have been left off for visual clarity. Note that the tuning in the ICcc neurons is present immediately, but only begins to be apparent with five trials per ITD for the NL neuron.

of the response curves, and hence not be of any behavioral relevance. To address this point, we redid the analysis and considered only ITDs in the range of ± 200 μs. The time constants of convergence remained significantly different at the $p < 10^{-6}$ level, indicating that this result is not an artifact of the long-range ITD curves.

Having established that only one or two trials is in general necessary to establish ITD tuning with ICcc neurons, it is reasonable to wonder exactly how short a stimulus can be to develop this tuning. Figure 5.5 illustrates this point. In neurons that could develop tuning within a single trial, we often found that tuning developed quite rapidly, often within the first 25 ms of the stimulus. It

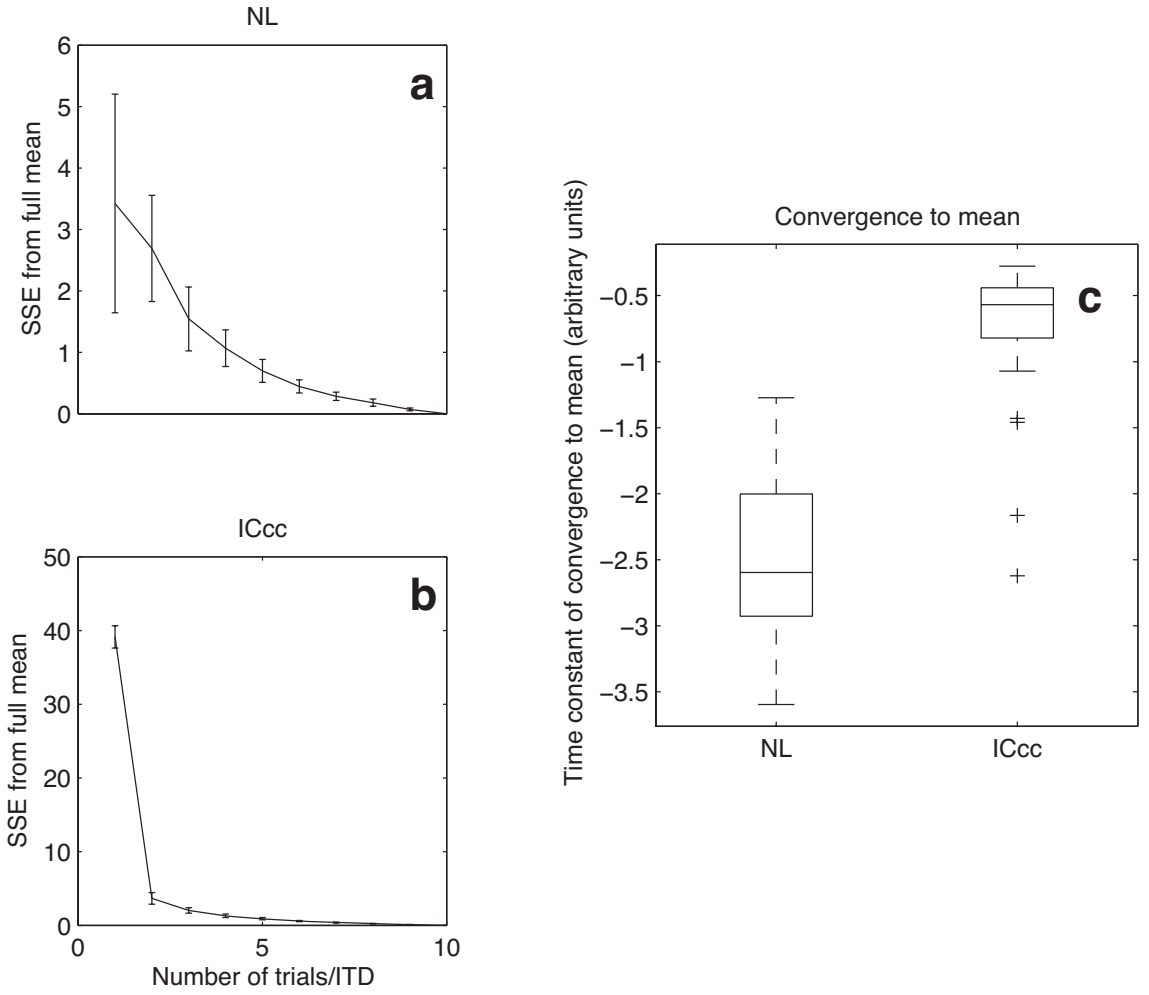


Figure 5.3. Examples of convergence curves for NL (***a***, *top left*) and ICcc (***b***, *bottom left*) using the same neurons as in figure 5.1. The majority of the difference between a single-trial ITD tuning curve and the full ITD tuning curve is eliminated with the addition of only a single trial to the averaging process for the ICcc neuron. Conversely, fall-off from the contribution of additional trials is much slower in the NL neuron. Error bars are standard deviations produced by Monte Carlo resampling. ***c***: Population data (see main text). Boxes extend from lower quartile to upper quartile of the sample, with the center line marking the median. Outliers (+) are data points greater than 1.5 times the interquartile range of the sample.

should be recalled that we use stimuli with a 5 ms rise (chapter 2), and that the latency in ICcc varies from 5 to 8 ms (figs. 3.5, 3.6), so that the first 25 ms can in fact reflect as little as the response to 10 ms of stimulus.

Quantifying this effect is problematic. Putting aside the issue of the range in latencies, to quantify this would require a precise definition of quality of tuning. We cannot use the same analysis as before,

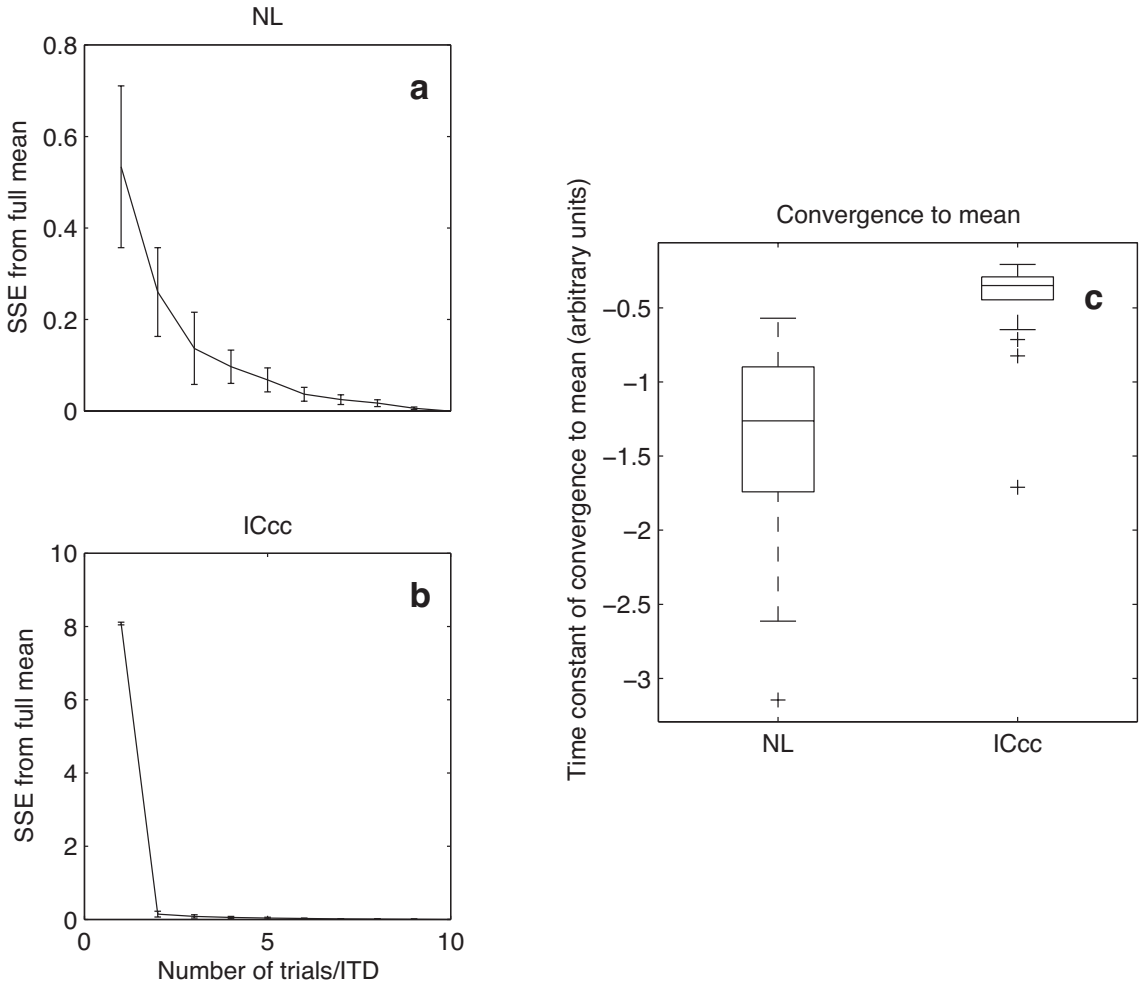


Figure 5.4. This is the same analysis as in figure 5.3, but limiting the range of ITDs considered to ± 200 μs. There is no change in the results.

as SSE requires a meaningful standard of comparison to be interpreted. We are further hampered by the fact that by necessity we have only a single data point per ITD condition. Since this result is not critical to the question this work hopes to address, we choose not to attempt to quantify it.

5.2.2 Dynamic range

Figure 5.1 suggests that there is also a difference in dynamic ranges between NL and ICcc neurons. This observation is borne out by figure 5.6a. The median dynamic range of the 28 ICcc neurons examined was 161 spikes/s, compared to 65.7 spikes/s, the median of 31 NL neurons (significant at $p < .001$, Kruskal-Wallis).

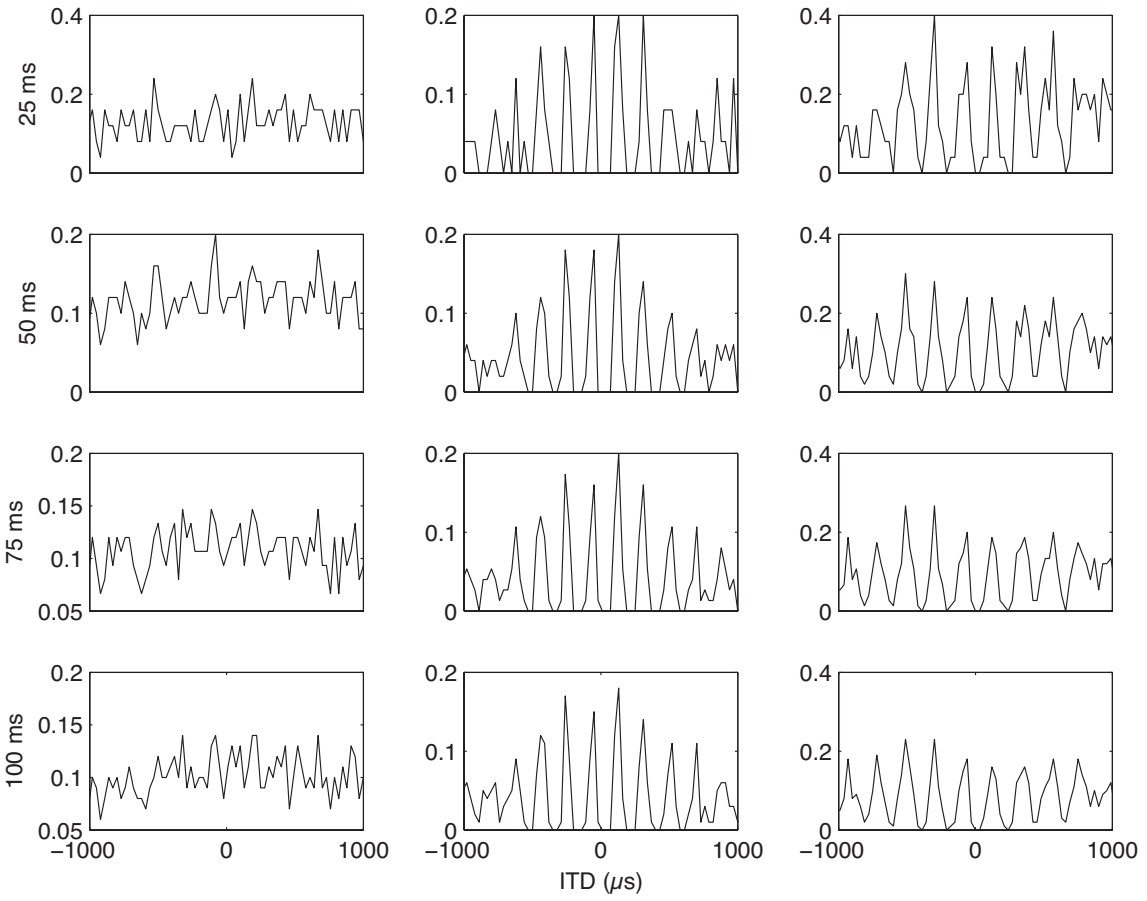


Figure 5.5. The ITD response of three example ICcc neurons (*columns*) in terms of normalized firing rate computed over the first 25 ms (*first row*), 50 ms (*second row*), 75 ms (*third row*), and the full trial (*bottom row*). In neurons that did not develop tuning using only a single stimulus presentation per ITD (*first column*), using a subset of the stimulus only served to exacerbate the situation. However, when a single trial was sufficient, the tuning developed rapidly, clearly visible even when only the first 25 ms of the stimulus was considered.

While the dynamic ranges were different, the error bars of the ITD tuning curves were not significantly different ($p > 0.1$, Kruskal-Wallis). An increase in dynamic range, and hence in the slope of the tuning curve, along with a constant amount of error intuitively suggests that the tuning curves of ICcc neurons will provide a finer discrimination of nearby ITDs, as nearby ITDs will "separate" (have non-overlapping error bars) sooner. Mathematically, this idea is expressed by the Fisher information (Dayan and Abbott 2001). Estimating the single-unit Fisher information as the slope of the tuning curve at an ITD x divided by the average firing rate at x, and taking the maximum across the entire tuning curve, we obtain the population data of figure 5.6b. The ICcc

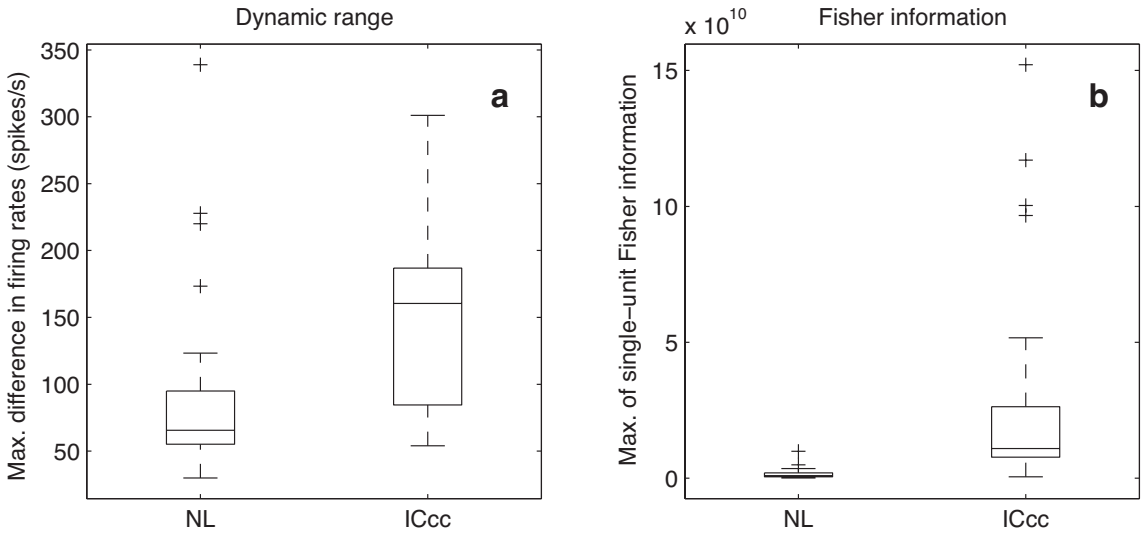


Figure 5.6. ***a***: Dynamic ranges. Dynamic range was defined as the maximum difference in firing rates through an ITD tuning curve. ***b***: Maximum single-unit Fisher information. Box plots as in figure 5.3.

neurons have a higher single-unit Fisher information, and hence a better ability to resolve nearby ITDs, than the NL neurons ($p < 10^{-6}$, Kruskal-Wallis).

5.2.3 Rectification of response

In a majority of the neurons of ICcc, we observed some effect of rectification (fig. 5.7). In other words, for most of the ICcc neurons, the response for peripheral ITDs was less than the midpoint between maximal and minimal responses. To quantify this, we introduce the *rectification index* (RI). If an ITD tuning is normalized to have values between 0 and 1, then the mean firing rate of the peripheral ITDs in an unrectified curve should be 0.5.

The results of the RI analysis are shown in figure 5.8. The NL neurons are tightly clustered about a RI of 0.5. In contrast, the ICcc population is skewed to lower RIs, indicative of a tendency towards rectification in the population (population medians different by Kruskal-Wallis, $p < 10^{-4}$).

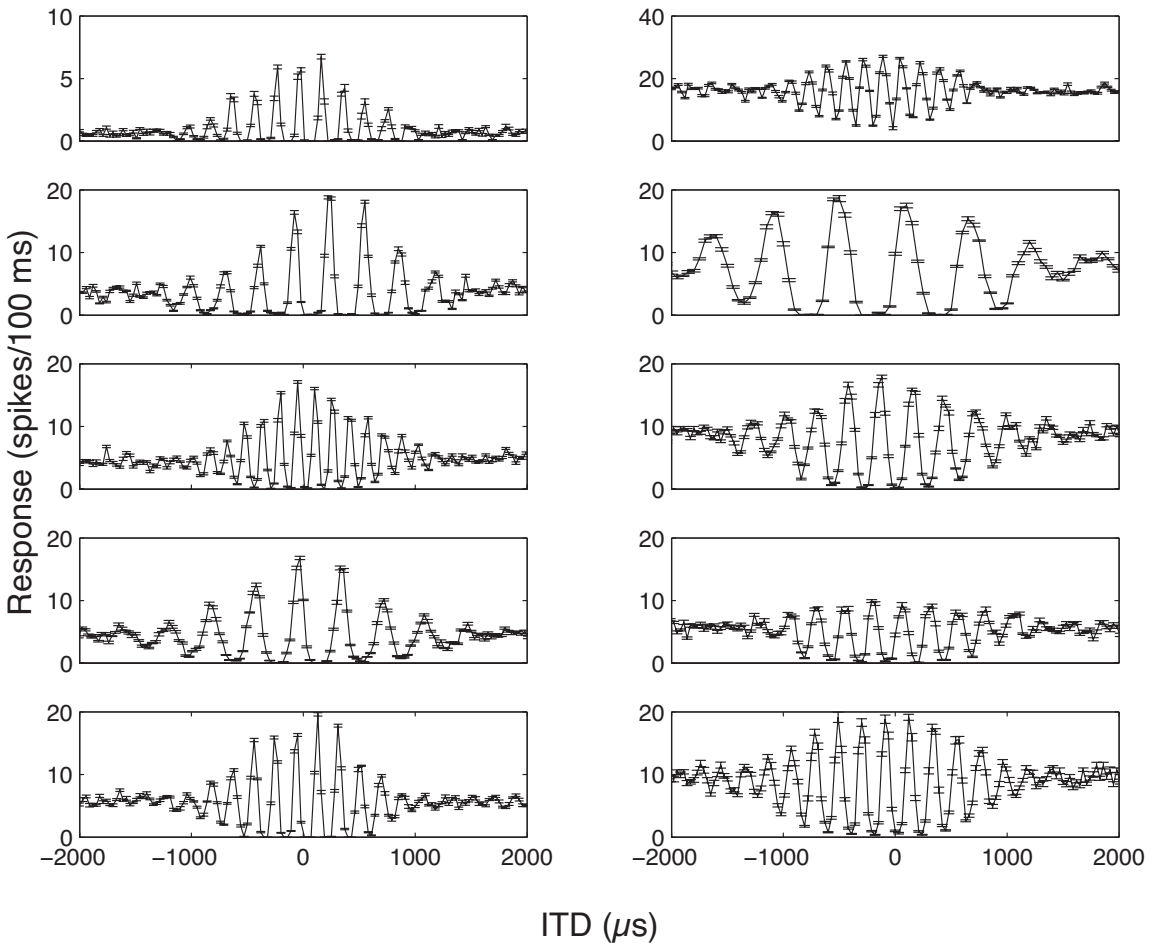


Figure 5.7. Eight ITD response curves from ICcc neurons illustrating the presence of rectification. Not all neurons (*right-hand column*) displayed rectification, but in a majority of neurons there was a pronounced effect.

5.2.4 Model of ICcc neurons

The above results suggest that ICcc response derives from a simple summation of the outputs of a number of similarly-tuned NL neurons. We would like to use the model of NL response we developed in chapter 4 to explore the mechanisms necessary to produce the observed responses in ICcc. However, the NL model was not perfect. In particular, because the ITD tuning curves of the NL model had dynamic ranges closer to what we observed in the ICcc data than in the NL data, an ICcc model built on the NL model cannot be used to meaningfully explore the evolution of ITD tuning. However, the reverse correlation of the NL model was a good match to the data. We can

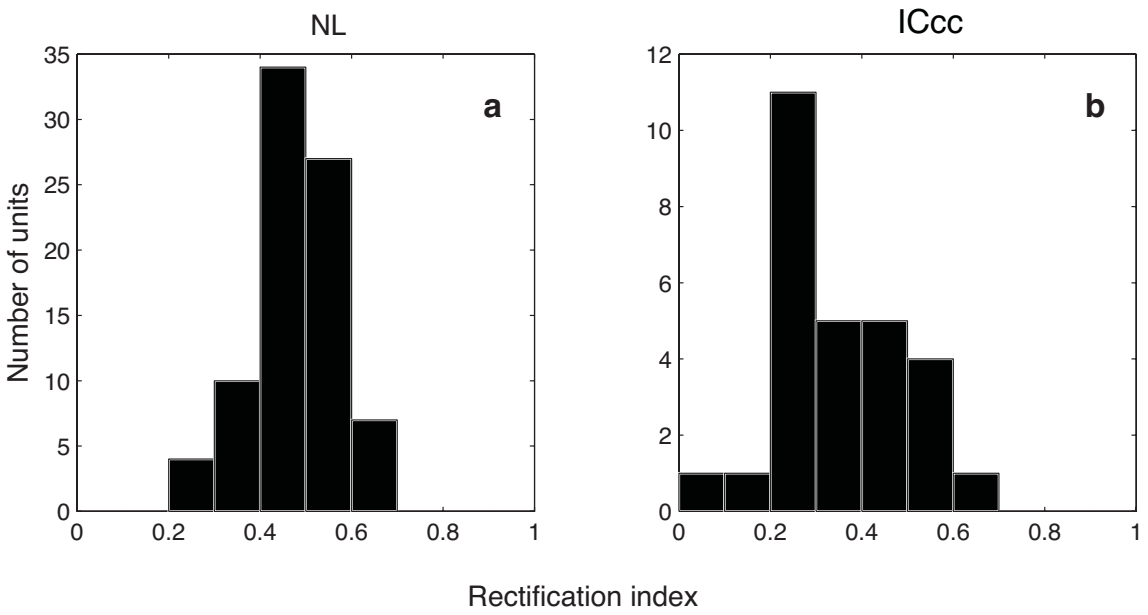


Figure 5.8. Histograms of the RIs for 82 NL neurons (***a***) and 28 ICcc neurons (***b***); previously collected NL neurons using stimuli of 50 ms for the ITD tuning curve were included in this analysis. Median of the NL neurons: 0.481; Median of the ICcc neurons: 0.323.

thus use it to address the question: what mechanisms are necessary to abolish phase-locking while maintaining a spike-triggered covariance?

Our model ICcc neuron is implemented as a simple leaky integrate-and-fire model with three parameters: τ_s, the synaptic time constant, τ_m, the membrane time constant, and N, the number of afferent inputs. The different afferent inputs were generated from repeated presentations of a single stimulus to a particular binaural model of NL from chapter 4 (in all the figures shown here, the model of fig. 4.5f was used). Because the NL model is stochastic, repeated presentations are not identical, allowing us to treat them as if they were the simultaneous output of identically-tuned neurons.

Initially, we reduce the model to two parameters by setting $\tau_s = \tau_m = \tau$. The reverse correlation of the model is illustrated in figure 5.9. For τ greater than 1 ms, the STA is clearly abolished. When $\tau = 1$ ms, then there is some apparent structure in the STA, which is clearest when $N = 2$. STC_1 only showed structure for $\tau = 1$ ms and $\tau = 5$ ms. There is little apparent effect of the number of inputs.

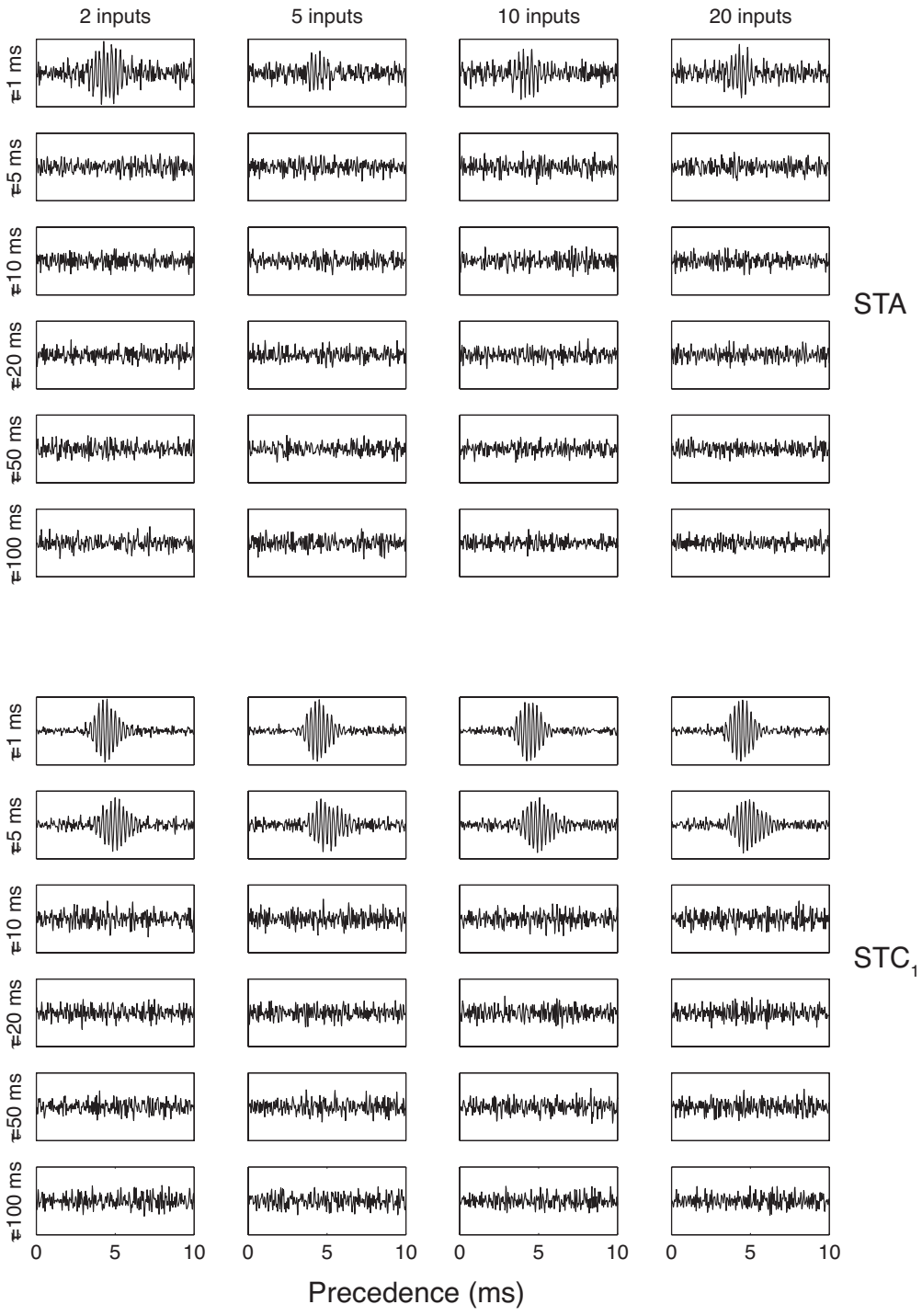


Figure 5.9. STA (*top*) and STC_1 (*bottom*) for a model ICcc neuron, for different values of the time constants of the integrate-and-fire model (*rows*; $\tau_s = \tau_m$) and number of model NL inputs (*columns*). The y-axis is $[-0.02, 0.02]$ for the STA and $[-0.2, 0.2]$ for STC_1; the tick-marks have been removed for legibility.

We can also look at the amount of variability in the response using the shuffled auto-correlogram method (SAC; Joris 2003; chapter 3). The SACs of the same conditions are shown in figure 5.10. There is a jaggedness present in all the correlograms; this is due to the fact that the temporal resolution of the spikes in the model is relatively large compared to the histogram bin-width of the correlogram. However, it is apparent that there is here a dependence on the number of inputs, with higher N leading to taller and narrower peaks. The outer half-height widths of these correlograms are given in table 5.1, and confirms this trend. In general, the NL model output itself tended to have more variability than was present in the data. Stipulating that, however, an N of 5 or more (for $\tau = 1$ ms) or 20 or more (for $\tau = 5$ ms) produced a SAC half-height width that is reasonably consistent with the observed data.

Table 5.1. Outer half-height widths of the ICcc model

	2 inputs	5 inputs	10 inputs	20 inputs
$\tau = 1$ ms	2.0122 ms	1.2440 ms	1.0029 ms	0.6934 ms
$\tau = 5$ ms	14.0255 ms	1.7646 ms	1.5448 ms	1.0980 ms
$\tau = 10$ ms	15.0428 ms	14.5322 ms	14.0043 ms	1.7043 ms
$\tau = 20$ ms	16.5271 ms	17.5133 ms	17.0365 ms	16.5049 ms
$\tau = 50$ ms	19.9021 ms	19.9070 ms	19.9072 ms	19.9001 ms
$\tau = 100$ ms	19.9007 ms	19.9102 ms	19.9139 ms	19.9013 ms

Based on this, we would conclude that ICcc should have a time constant of 1–5 ms and require as many as 20 convergent inputs. There is no published histology or cell count data that would allow us to determine whether the proposed number of afferents is plausible. The limitations we have put on τ, however, are fairly strict; measurements of the membrane time constant for most neurons range from 10 to 100 ms (Koch 1999). So far, we have been equating the membrane time constant τ_m (how quickly the membrane voltage can respond to changes in current) with the synaptic time constant τ_s (the half-life of the current injection due to a single input spike). In general, these can be different. When we fix τ_s to be 1 ms, and vary τ_m alone, we find that changes in τ_m have very little effect (fig. 5.11). When $\tau_m = 1$ ms, and also in some cases where $N = 2$, there is some evidence of a structured STA. For all other cases, there is no STA and a structured STC_1. Conversely, if we fix $\tau_m = 100$ ms, and vary τ_s, we find that there is no structured spike-triggered covariance for any

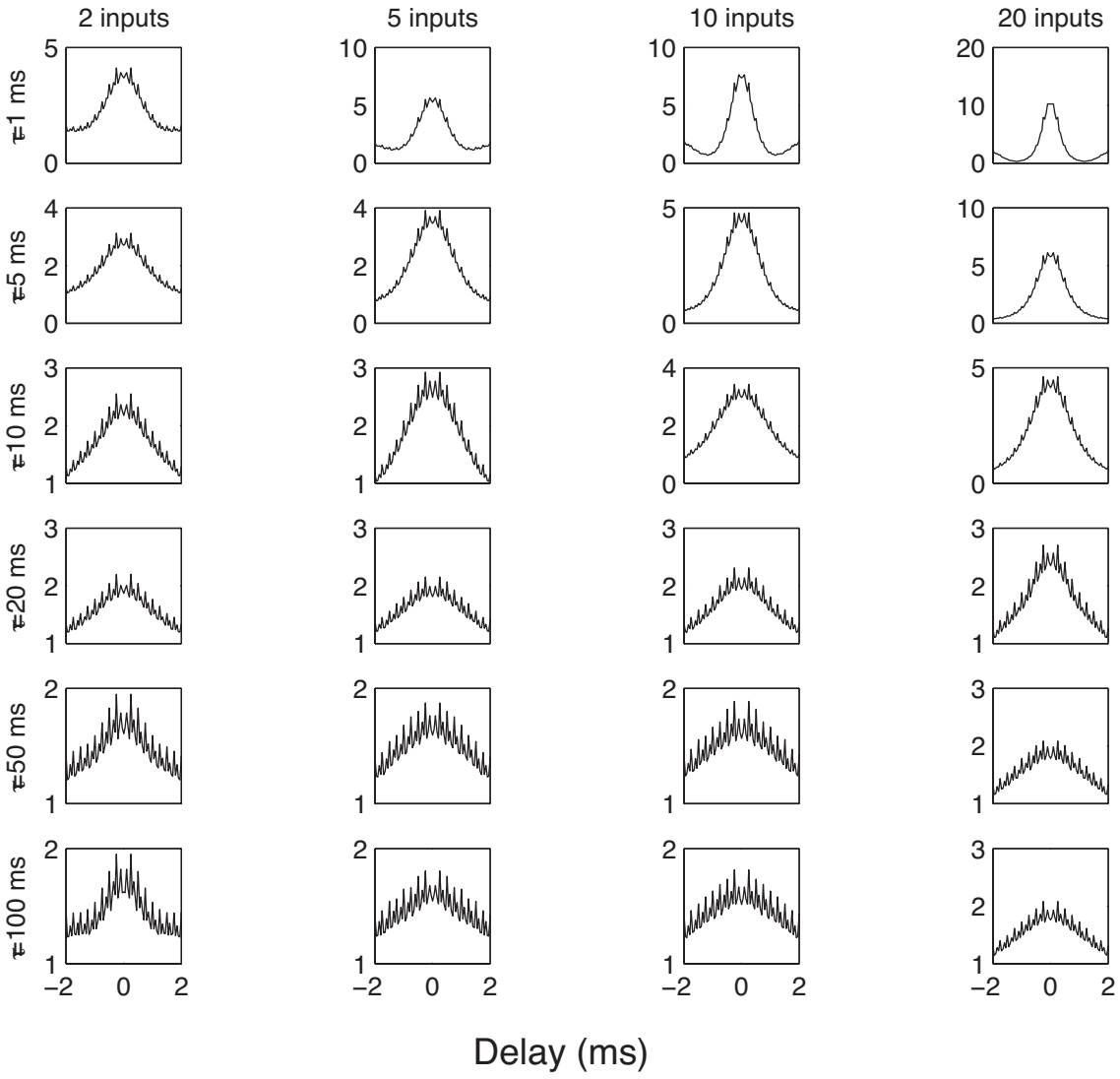


Figure 5.10. SAC for a model ICcc neuron, for different values of the time constants of the integrate-and-fire model (*rows*; $\tau_s = \tau_m$) and number of model NL inputs (*columns*).

τ_s greater than 1 ms (fig. 5.12).

5.3 Discussion

A basic tenet of neuroscience is the idea of averaging across a population of noisy inputs to achieve a more reliable measure of the encoded variable. This appears to be the computation that ICcc neurons are performing. The constant variability from NL to ICcc is indicative of a linear summation, and

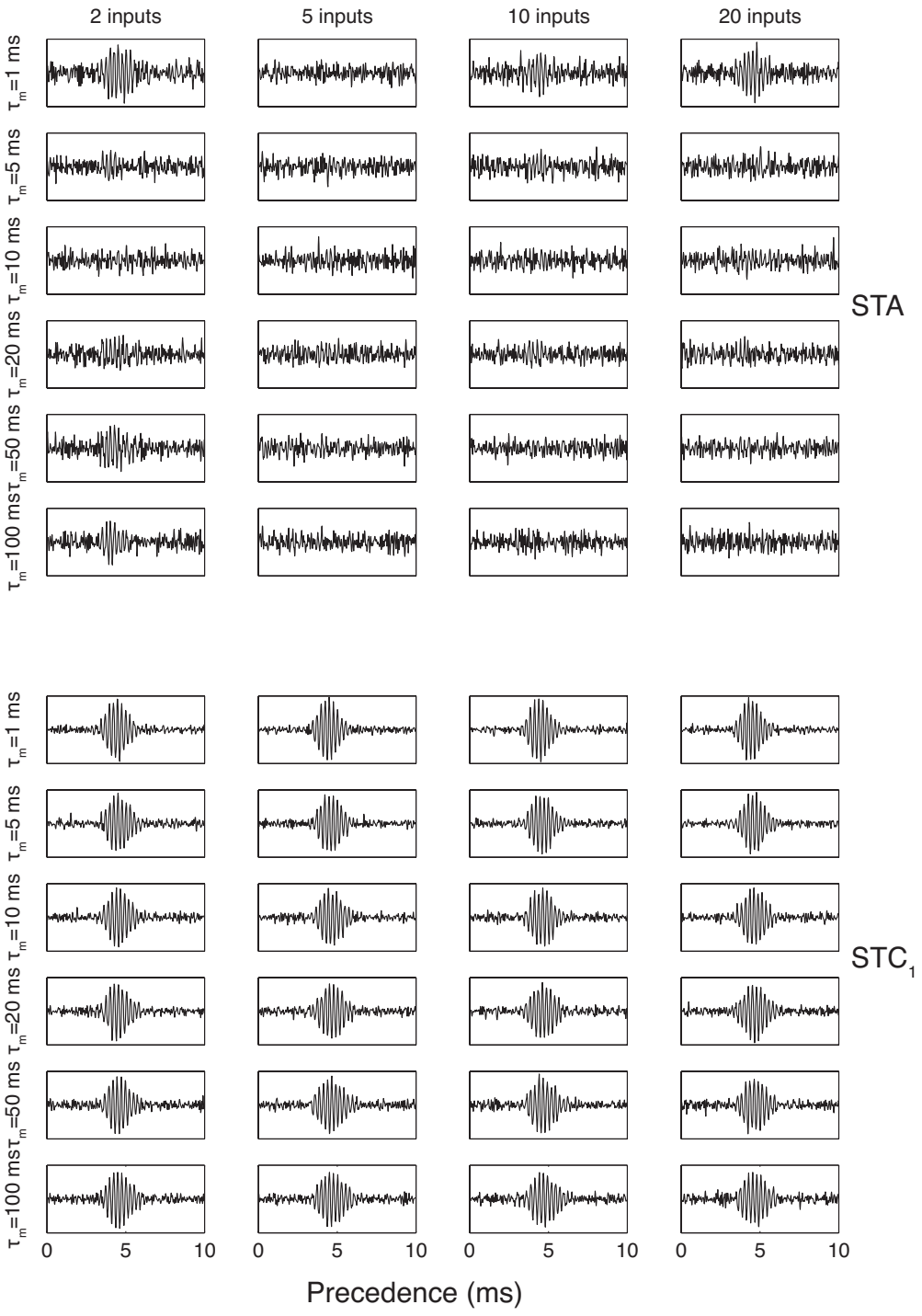


Figure 5.11. STA (*top*) and STC_1 (*bottom*) for a model ICcc neuron, for different values of τ_m and number of model NL inputs (*columns*), with fixed $\tau_s = 1$ ms. The y-axis is $[-0.02, 0.02]$ for the STA and $[-0.2, 0.2]$ for STC_1; the tick-marks have been removed for legibility.

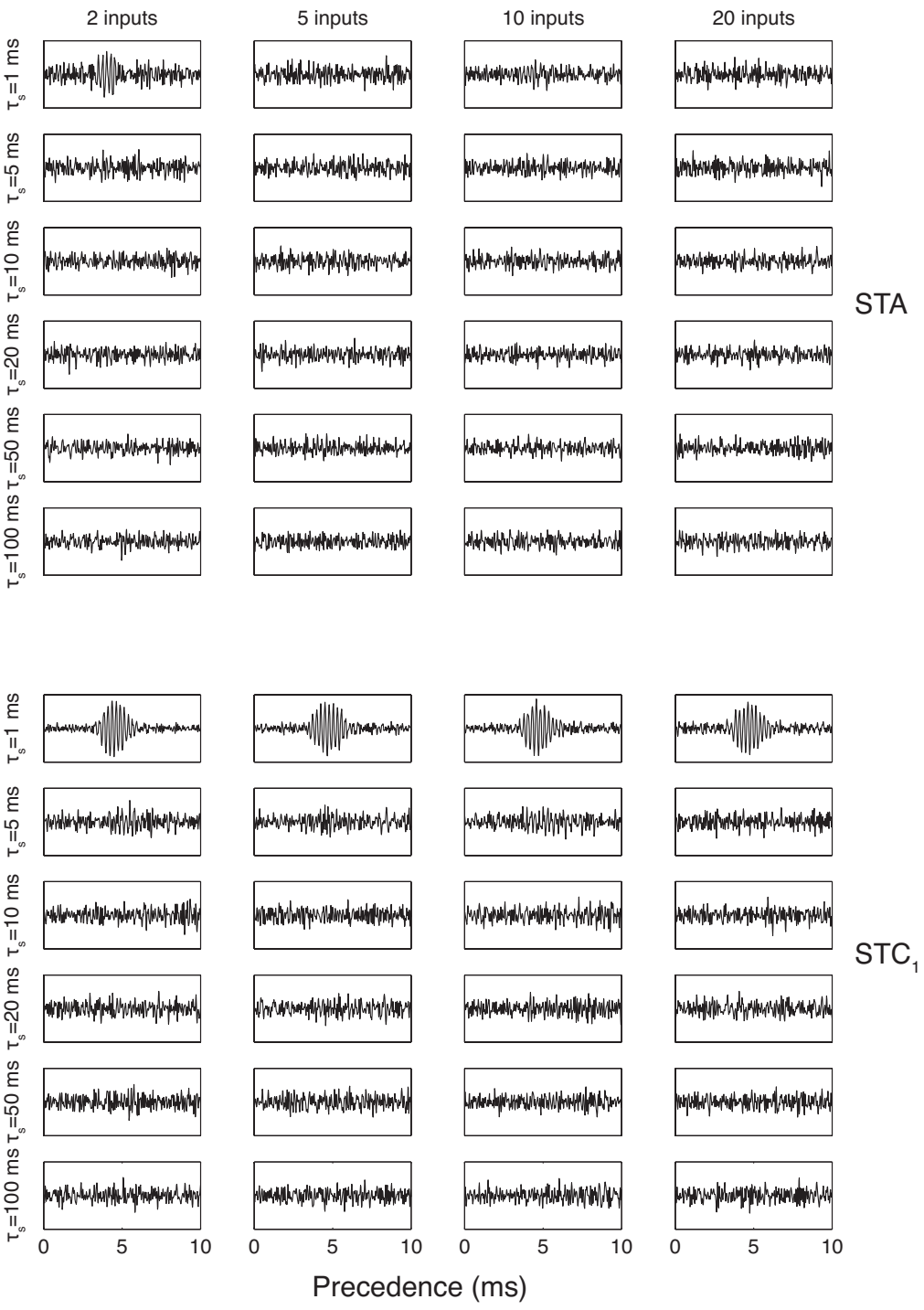


Figure 5.12. STA (*top*) and STC_1 (*bottom*) for a model ICcc neuron, for different values of τ_s and number of model NL inputs (*columns*), with fixed $\tau_m = 100$ ms. The y-axis is $[-0.02, 0.02]$ for the STA and $[-0.2, 0.2]$ for STC_1; the tick-marks have been removed for legibility.

this is consistent with the increase in the dynamic range. The difference in dynamic range as shown in figure 5.6a is likely an underestimation of the degree of summation present. As we demonstrate in figure 5.8, there is a degree of rectification in the ITD tuning curves of ICcc that was never present in the NL tuning curves. If we consider the theoretical dynamic range of the unrectified tuning curve (that is, the one predicted by the Wiener-Khinchin theorem), then it will exceed the actual dynamic range by a factor of nearly two. A consequence of this rectification is the possibility of obfuscation of ITDs in the troughs. As can be seen in a few of the examples of figure 5.7, there are times when a range of unfavorable ITDs are all encoded with zero firing rate. The conclusion is that ICcc neurons combine a large number of inputs from NL in order to ensure that changes in ITD within the owl's physiological range induce firing rate fluctuations that extend over the majority of the neurons' dynamic range, even at the possible expense of rectification due to thresholding.

It is tempting to conclude from this that the troughs of the ITD curve are somehow irrelevant, but this is not entirely correct. While this is true when coding is perceived from the single-unit perspective, it is generally believed that neural information is coded across a population of neurons (Panzeri et al. 1999; Dayan and Abbott 2001; Sahani and Dayan 2003; Johnson and Ray 2004; Latham and Nirenberg 2005). Given that horizontal position is obtained by converging ITD information across frequency channels, this is certainly true in the case of the owl. Interestingly, the optimal distribution of tuning curves as predicted by Fisher information comes from a population which has widely varying tuning curves, and for which many neurons respond for any given stimulus condition (Dayan and Abbott 2001). Normally, these conditions are conflicting: a sharply tuned response curve normally responds to comparatively few stimulus conditions. However, in both NL and ICcc, the periodic nature of the response guarantees that they will respond to the majority of possible ITDs; in accordance with this, Harper and McAlpine (2004) showed that the distribution of ITD curves in the barn owl matched the optimum distribution as predicted by Fisher information. Our results demonstrate that while NL and ICcc have the same distribution and shape of ITD curves (chapter 4), the increase of dynamic range leads to a greater satisfaction of the first condition of the optimal population code. Because of the information processing inequality, ICcc can have no more

information than NL. However, the improvement of single-unit tuning in ICcc is a sign that the information is made more accessible than in NL, and that acquiring position information requires examining fewer ICcc neurons than would be necessary in using NL.

While rectification may not have an effect on the ITD of correlated signals within the physiological range, its effects remain significant. At extreme ITDs, broadband signals are effectively uncorrelated: thus, the response to these ITDs is expected to be the same as the response to binaurally uncorrelated stimuli (fig. 5.13). The response to binaurally uncorrelated stimuli in turn is closely related to the response to monaural stimuli. Our results thus suggest that in ICcc a suppression of monaural responses is apparent. This is consistent with previous observations that monaural stimulation becomes less effective ascending through the auditory pathway (Moiseff and Konishi 1983). Combined with the observation from chapter 4 that the ITD curves of ICcc are determined by spectral tuning, the conclusion is that monaural responses can be eliminated through the successive process of thresholding.

The increase in dynamic range serves to reduce the noise in ITD tuning in another, slightly more subtle manner. As we described in chapter 3, both NL and ICcc neurons have firing patterns that encode the spectrotemporal properties of the stimulus. As a result of this, there will be variations in the firing rate that are due to the spectral properties of the stimulus, and not the ITD; this problem is only exacerbated with short stimuli, which provide less time over which to average and eliminate such effects. However, these fluctuations would be expected to be of the same magnitudes in both nuclei, a fact confirmed by our comparison of the standard deviations of the ITD tuning curves, and they would have less impact within ICcc, as they would be of relatively less significance compared to the changes in mean firing rate that arise out of changes in ITD. Therefore, another effect of the increase in dynamic range is to allow ICcc to encode two orthogonal stimulus dimensions (spectrotemporal properties and ITD) with minimum interference.

Our model results suggest that phase-locking can be abolished and significant covariance can be maintained with a convergence of five or more inputs on a neuron with a synaptic time constant of 1 ms and a membrane time constant larger than 1 ms, and otherwise unbounded. It should be

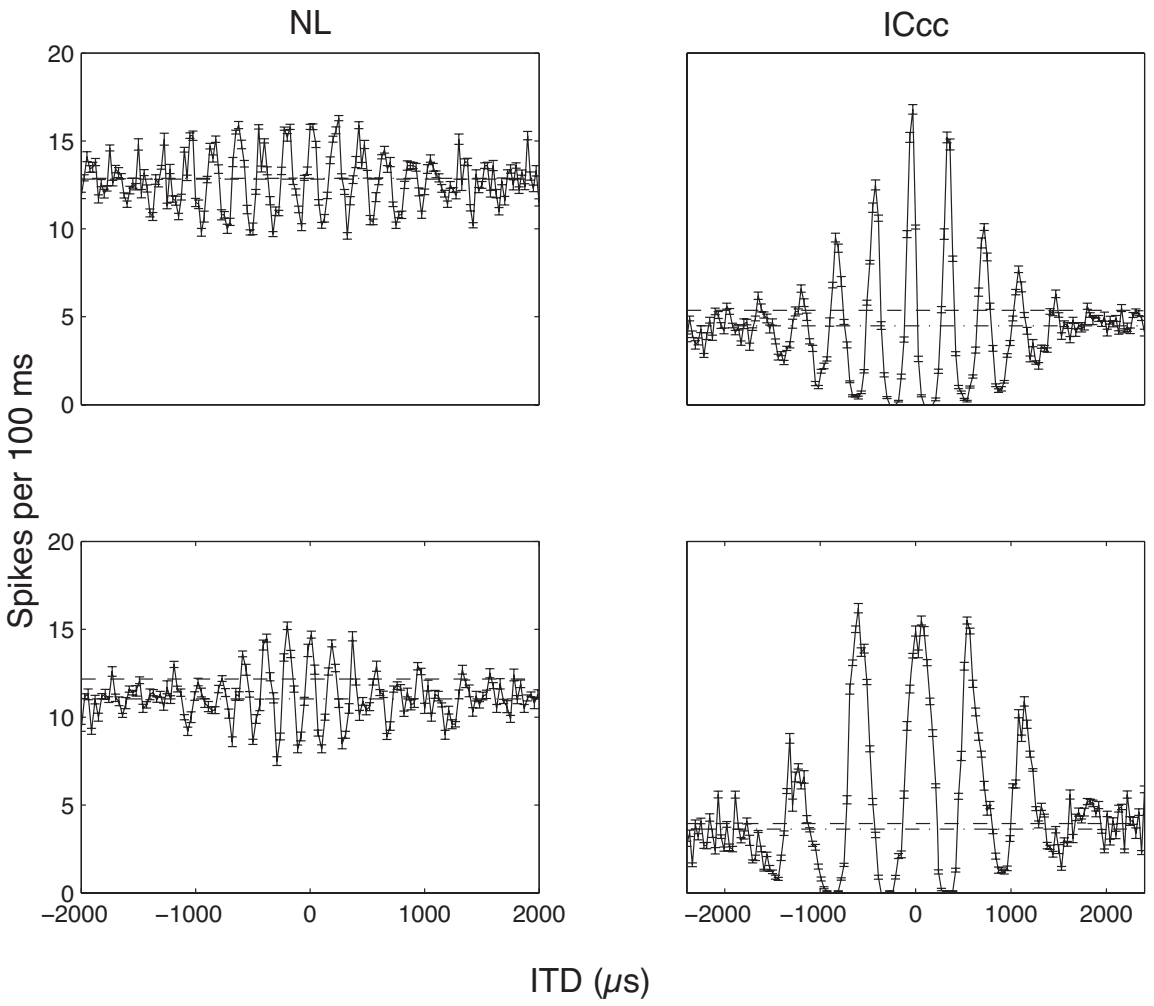


Figure 5.13. Example ITD curves for NL (*left column*) and ICcc (*right column*). The mean firing rate in response to binaurally uncorrelated stimuli is superimposed (*dashed line*), along with the mean firing rate of the peripheral ITDs (*dashed and dotted line*; estimated using the same range of ITDs as for the RI). Both firing rates are similar in all cases.

stated clearly that there could be and likely are other mechanisms at play. For example, the afferent axons are likely to be of different lengths, which could easily introduce enough jitter to abolish phase-locking. Hence, as a prediction of the actual anatomy and biophysics, our model should be approached with caution. The model does, however, establish a minimum mechanism to accomplish the conversion from STA to STC; the fact that the requirements of this mechanism are relatively undemanding suggest that the abolishment of phase-locking does not require any active mechanism. In other words, the loss of phase information is most likely not a deliberate move of benefit to the

auditory system, but rather a relaxation of the effort required to maintain phase information.

Though the results are clear, it is not evident what benefit the system gains by doing this averaging explicitly within ICcc, as opposed to performing it at the same time as the frequency convergence or the emergence of combination selectivity to ITD and ILD take place. The implication of these results is that it is important to have an accurate estimate of the ITD alone within a narrow frequency band before integrating across frequency channels. It is known that ICcc projects not only to the lateral shell of the central nucleus of the inferior colliculus (ICcl), but also directly to the thalamus (Cohen, Miller and Knudsen 1998; Proctor and Konishi 1997). Since the thalamus also receives projections from the lateral shell, there is no *a priori* reason based solely on considerations of sound localization to require a thalamic projection from ICcc. That such a projection does exist suggests a particular role for band-limited ITD information in the thalamic processing stream. Because interaural correlation, which is the basis of ITD detection, will be influenced not only by the location of the sound in space, but by features of the acoustic environment such as the presence of echoes, the existence of multiple sound sources, and distorting effects of the environment, it is plausible that it plays some role in non-localization perceptual tasks. Additionally, work in ICcl has suggested that it may be that frequency convergence is a gradual process, occurring in a cascade of neurons that terminates in the true space-specific neurons of ICx, rather than in a single step (Mazer 1995). There may be a biophysical constraint on the number of inputs that can be managed by a single lateral shell neuron that requires that the summation for the purposes of noise reduction in the ITD domain occurs before any process of frequency convergence begins.

The question also arises why this averaging to eliminate noise must be done after NL. It seems likely that the neurons of NL are already performing near the neural limits for coincidence detection. The temporal jitter of the inputs is significant compared to the frequencies involved (Köppl 1997b), and the time scales of the coincidences require specialized neurons with fast time constants (Han and Colburn 1993; Gerstner et al. 1996). Under these conditions, greater reliability may not be possible within the coincidence detectors themselves, requiring that an additional stage of processing perform the necessary averaging.

Another factor that may play a role is the need to segregate feedback circuits at different stages in the processing of ITD. There are two feedback projections involved in this portion of the ITD pathway. As described above, the arcopallium is reported to project back to the central nucleus of the inferior colliculus (Cohen, Miller and Knudsen 1998), and it is possible though currently unconfirmed that this projection may include ICcc. This feedback is hypothesized to play a role in attentional modulation in sound localization (Gaese and Wagner 1992); a projection to ICcc would suggest the need to manipulate the ITD-specific frequency channels independent of ILD. In addition, a looping projection from NL to the superior olivary nucleus (SO) back to NL again exists (Introduction; Lachica, Rübsamen and Rubel 1994), which is believed to mediate the elimination of intensity effects on ITD tuning in NL (Peña et al. 1996). SO's projections, however, are not limited to NL. If the forebrain feedback does indeed need to influence ITD information alone, then without ICcc, the auditory arcopallium feedback would then be required to project to NL, and from NL through SO and all of its targets.

Since we have taken a black-box approach to the function of the post-laminaris ITD pathway, it is impossible for us to distinguish how much of the difference between ICcc and NL responses arises due to the intermediate effects of LLDa, as opposed to the computation of ICcc neurons using the direct NL projection. However, none of the differences we describe seem to necessitate the added complexity of the LLDa projection to account for them. The published literature on LLDa is sparse (Moiseff and Konishi 1983; Albeck and Konishi 1995), and suggests tuning properties that are similar to NL and ICcc, which is expected from anatomical considerations. The fact that there is no need to require the action of LLDa to explain our results suggests that its computational role may be one not triggered by the stimulus conditions used in this study. It may have a function in localizing moving sound sources, or in the identification of multiple sources. Of course, even though the actions of LLDa are not seemingly required to explain the observed data, this does not eliminate the possibility of a contribution of LLDa in the observed averaging. However, any possible effect of a delayed input from LLDa on the output of ICcc neurons is concealed by the onset dynamics of the response. To distinguish the NL and LLDa inputs would require intracellular recordings that would

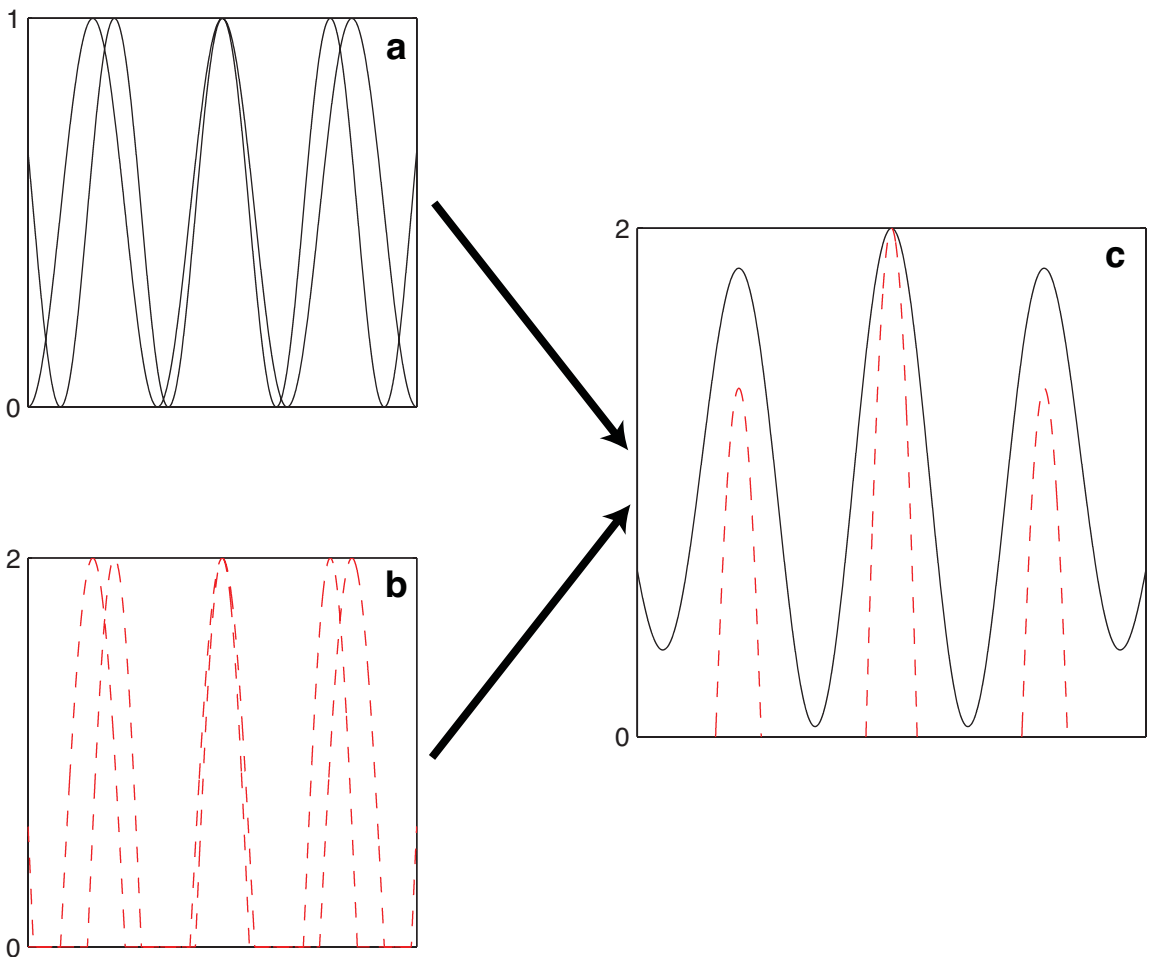


Figure 5.14. A schematic to illustrate the effects of dynamic range on frequency convergence. In ***a***, two cosines of different frequencies are plotted. In ***b***, the same cosines are plotted, but the dynamic range has been increased by a factor of two. In addition, the cosines have been half-waved rectified, which means the dynamic range is effectively greater by a factor of four. ***c***: The curves of ***a*** and ***b*** have been summed and plotted on the same axes, with the result of the sum of ***b*** (*dashed line*) shifted to have the same peak as the sum of ***a*** (*solid line*), and then clipped to match the same dynamic range. Even though the same frequency components were used for both sums, the sidepeaks of the sum of ***b*** are considerably more suppressed than the sidepeaks of **a**.

allow identification of subthreshold inputs with different latencies or data collected with inactivation of neural centers.

Strictly speaking, averaging suggests that the dynamic range of the averaging unit should be on the same order as the dynamic range of its inputs. However, the dynamic range is larger in ICcc than in NL. This increase may serve to accelerate the process of frequency convergence that occurs in the next stage of the sound localization pathway (Mazer 1998). The premise of frequency

convergence, confirmed in the nucleus ICx (Takahashi and Konishi 1986) includes summation across frequency channels in the ITD domain (Takahashi and Konishi 1986; Mori 1997; Mazer 1998) with thresholding to eliminate peaks that do not correspond to the true ITD (Peña and Konishi 2000, 2002). This process is influenced by the absolute magnitude of the component ITD tuning curves; the larger their initial amplitude, the larger the absolute difference between true and secondary peaks in the summed curve will be, simplifying the task to be accomplished by threshold (fig. 5.14).

It has been shown that the owl can localize sounds as short as 10 ms in duration (Konishi 1973). Our results indicate that there is a move towards reliable short-time scale ITD encoding on a single neuron level within the sound localization pathway. The spiking response of ICx neurons, which feature low firing rates with little or no sustained response (Wagner 1990; Peña and Konishi 2000, 2002), represents the culmination of this trend, and it has been reported that single ICx neurons can in fact match the behavioral performance (Bala, Spitzer and Takahashi 2003). Experiments in ICx have indicated that summation and thresholding of inputs is a crucial component of the neuronal computation of space specificity (Peña and Konishi 2000, 2002). The computations in ICcc provide a necessary basis for this, with the amplification of dynamic range and the reduction of noise working together to ensure that only the desired portions of the ITD response will exceed threshold.

Chapter 6

Summary of Results

We have demonstrated that in addition to encoding the interaural time difference (ITD) in the mean firing rate, the neurons of nucleus laminaris (NL) respond to complex stimuli with a patterned response. This pattern encodes details about the instantaneous spectral properties of the signal with a temporal resolution on the order of a millisecond.[1] The ITD response itself, consistent with the predictions of the cross-correlation model for the computation of ITD, is determined entirely by the neuron's spectral tuning. The ITD tuning of a single NL neuron is noisy, requiring the equivalent of a half-second or more of stimulus to obtain reliable tuning.

In the same manner, the core region of the central nucleus of the inferior colliculus (ICcc) also encodes both spectral properties and ITD. Its spectrotemporal tuning is essentially identical to that of NL, though phase information is now lost, and the same general relationship between spectral and ITD tuning is observed. However, ICcc neurons encode ITD over a greater dynamic range. Partially as a function of the increase in dynamic range minimizing the impact of instantaneous firing rate fluctuations due to spectral properties of the stimulus, and presumably in part because of averaging across a population of NL neurons with similar tuning properties, the ITD tuning is more clearly encoded and develops in coherency within the context of a much shorter effective stimulus.

Thus, ICcc is NL with improved single-unit tuning. This trend seems at first light to stand in

[1] Properly speaking, these statements are true only if the position of the neuron within the ITD and frequency maps throughout the nucleus are known. That is, the characteristic delay (CD) of a neuron is determined by its position within NL, and the mean firing rate provides a measure of the distance of the ITD of the stimulus from the CD of the neuron. Similarly, the pattern of response indicates the time of occurence of energy within a particular frequency band for a stimulus, with the identity of that frequency band being given by the neuron's location within the tonotopic map. Under experimental conditions, we have *a priori* knowledge of the stimulus, which allows us to interpret a neuron's output without reference to the maps. From the perspective of computations within the system, however, the organization provided by the maps is crucial.

contrast with the accepted view, informed in large part by work in the mammalian cortex, that the tuning of single units is of vastly lesser importance, and that it is only by considering a population of neurons that information can be extracted. However, the improvement in single-unit tuning also serves to improve the ease with which information may be extracted out of the population, as well. That an improvement in single-unit tuning is seen is in large part a function of the small number of explicit stimulus dimensions that are encoded at this stage in the auditory processing hierarchy. A single neuron can convey information only through three channels: its mean firing rate, the exact timing of its spikes, and the neuron's position within some map or maps. In ICcc, neurons of similar bandwidths and temporal profiles explicitly encode spectral and ITD information, and the tuning profiles for and interactions between these cues are simple. As a result, the necessary stimulus dimensions fit within the coding capabilities of a single neuron. In contrast, studies in mammalian auditory cortex demonstrate much more variability in tuning properties, and sensitivity to more properties of the stimulus (Linden and Schreiner 2003), and the receptive fields of these neurons can be complex and nonlinear (Depireux et al. 2001; Linden et al. 2003); similar variety in tuning properties (Cohen and Knudsen 1996), along with a high degree of plasticity (Miller and Knudsen 2001), occur in the owl's auditory arcopallium. The need to represent stimuli using a higher number of dimensions in high-level processing exceeds the coding capacity of a single neuron, making the response of multiple neurons the basic words of the code and thus makes single unit tuning more difficult to decipher (Dayan and Abbott 2001).

Before the auditory system can perform such complex computations, however, it requires a reliable and concise encoding of the raw stimulus parameters. Providing such information appears to be the task of the post-laminaris ITD pathway. The phase encoding used in the computation of ITD is converted into a spectrotemporal stimulus representation, which is provided in a fashion effectively independent of the ITD, providing a remarkable example of the nervous system to recycle and re-represent information in order to encode multiple stimulus dimensions using a limited coding language.

Appendix A

The Weiner-Khinchin Theorem

Given a function h which is a function of time t, we can find an equivalent representation H which is a function of frequency f. The *Fourier transform* relates these two functions:

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} h(t) e^{2\pi i f t} dt \\ h(t) &= \int_{-\infty}^{\infty} H(f) e^{-2\pi i f t} df. \end{aligned}$$

We say that $h(t)$ and $H(f)$ form a *transform pair*, and denote it as:

$$h(t) \iff H(f).$$

The cross-correlation of two possibly complex signals is given by:

$$C_{gh}(\tau) = \int_{-\infty}^{\infty} \overline{g(t-\tau)} h(t) dt$$

where τ is the lag, and $\overline{g(t)}$ is the complex conjugate of $g(t)$.[1] Therefore, we have:

$$\begin{aligned} C_{gh}(\tau) &= \int_{-\infty}^{\infty} \overline{g(t-\tau)} h(t) dt \\ &= \int_{\infty}^{\infty} \left[\overline{G(f)} e^{2\pi i f (t-\tau)} df \int_{-\infty}^{\infty} H(f') e^{-2\pi i f' t} df' \right] dt \end{aligned}$$

[1] Normally, we deal with the case where the signals are real, and hence $\overline{g(t)} = g(t)$. Thus, the complex conjugate is often discarded from the definition, but it is required here.

$$
\begin{aligned}
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{G(f)}H(f')e^{-2\pi i f\tau}e^{-2\pi i(f'-f)t}dt\,df'df \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{G(f)}H(f')e^{-2\pi i f\tau}\left[\int_{-\infty}^{\infty} e^{-2\pi i(f'-f)t}dt\right]df'df \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{G(f)}H(f')e^{-2\pi i f\tau}\delta(f'-f)df'df \\
&= \int_{-\infty}^{\infty} \overline{G(f)}H(f)e^{-2\pi i f\tau}df \\
\iff & \overline{G(f)}H(f).
\end{aligned}
$$

In the case where $g = h$ (which is the case in all relevant experiments done in this work), we have:

$$
\begin{aligned}
C_{hh}(\tau) &\iff \overline{H(f)}H(f) \\
&\iff |H(f)|^2. \qquad \text{(A.1)}
\end{aligned}
$$

This result is known as the *Wiener-Khinchin theorem* or the *cross-correlation theorem*.

The *two-sided power spectral density* (PSD) of a signal h at a frequency f is given by:

$$
\mathrm{PSD}_h(f) = \frac{|H(f)|^2 + |H(-f)|^2}{2}.
$$

In the case where h is real-valued, $H(f) = H(-f)$, and we have

$$
\mathrm{PSD}_h(f) = |H(f)|^2.
$$

Therefore, by Equation A.1, we have:

$$
\begin{aligned}
\mathrm{PSD}_{C_{hh}}(f) &= |(|H(f)|^2)|^2 \\
&= |\mathrm{PSD}_h(f)|^2 \\
&= (\mathrm{PSD}_h(f))^2.
\end{aligned}
$$

(since the PSD is always real and postive). Thus, the PSD of the autocorrelation of a signal is given by the square of the PSD of the signal itself. Since we are presenting signals at the two ears with perfect binaural correlation (and hence, those signals are indentical), this is the basic result we wished to establish. For the purposes of this dissertation, we will say that any two signals g and h are in a *Wiener-Khinchin relationship* if they satisfy:

$$\mathrm{PSD}_g(f) \quad = \quad (\mathrm{PSD}_h(f))^2$$

for all f.

Appendix B

Running Cross-Correlation

Let $x(u)$ and $y(u)$ be two integrable signals, 0 everywhere outside $u \in [0, T)$, with values in $[-1, 1]$. The classic cross-correlation of these two signals is:

$$C_{xy}(\tau) \quad = \quad \int_{-\infty}^{\infty} x(u-\tau)y(u)du$$

where τ is the lag. The conditions placed on x and y are not necessary for the definition of cross-correlation. However, they are necessary for later parts of the derivation, and match the nature of the signals used in the context of this research.

The Wiener-Khinchin theorem states that:

$$C_{xy} \quad \Longleftrightarrow \quad \overline{\mathcal{F}(x)}\mathcal{F}(y)$$

where $\mathcal{F}(x)$ is the Fourier transform pair of x and $\bar{\cdot}$ represents the complex conjugate (see Appendix A).

However, what Licklider (1959) demonstrated was that the instantaneous firing rate of a coincidence detector would be given by the *running cross-correlation* of the two stimuli. Choose a function $G(u)$ such that as $t \to \infty$, $G(u) \to 0$, and $G(u) = 0$, $\forall u < 0$. Then, the running cross-correlation is given by:

$$RC_{xy}(t, \tau) \quad = \quad \int_{-\infty}^{\infty} x(u-\tau)y(u)G(t-u)du.$$

The running cross-correlation is a function of both time and lag, and is therefore distinct from the cross-correlation. In particular, no result equating the running cross-correlation to any function of the Fourier transform exists. Therefore, to apply the Wiener-Khinchin theorem to neural data, it is necessary to first demonstrate the conditions under which running cross-correlation data can be treated as a classic cross-correlation.

In analyzing our data, we consider the mean number of spikes produced per stimulus. This is equivalent to averaging across time, and the *time-averaged running cross-correlation* is then:

$$\begin{aligned} TARC_{xy}(\tau) &= \frac{1}{T}\int_0^T RC_{xy}(t,\tau)dt \\ &= \frac{1}{T}\int_0^T \int_{-\infty}^{\infty} x(u-\tau)y(u)G(t-u)du\,dt \\ &= \frac{1}{T}\int_{-\infty}^{\infty} x(u-\tau)y(u)\left[\int_u^T G(t-u)dt\right]du. \end{aligned}$$

Let $G(u) = H'(u)$ for $u > 0$. Then,

$$\begin{aligned} TARC_{xy}(\tau) &= \frac{1}{T}\int_{-\infty}^{\infty} x(u-\tau)y(u)\left[H(T-u)-H(0)\right]du \\ &= \frac{1}{T}\left[-H(0)\int_{-\infty}^{\infty} x(u-\tau)y(u)du + \int_{-\infty}^{\infty} x(u-\tau)y(u)H(T-u)du\right] \\ &= \frac{1}{T}\left[-H(0)C_{xy}(\tau) + \int_{-\infty}^{\infty} x(u-\tau)y(u)H(T-u)du\right]. \end{aligned}$$

Note that since $y(u) = 0$ when $u > T$, we need not be concerned by the fact that $H(T-u)$ is not defined when $u > T$.

From the above, we conclude that the time-averaged running cross-correlation is simply a scaled version of the normal running cross-correlation, plus an error term. In the context of neural processing, $G(u)$ is generally assumed to be related to the membrane time constant, and of exponential form (Yin, Chan and Carney 1987). Thus, we examine this error term in the case where $G(u) = e^{-u/k}$. If that is the case, then $H(u) = -ke^{-u/k}$, and we have:

$$TARC_{xy}(\tau) = \frac{k}{T}\left[C_{xy}(\tau) - \int_0^T x(u-\tau)y(u)e^{\frac{u-T}{k}}du\right].$$

Since x and y have range $[-1, 1]$, we see:

$$\begin{aligned}
\int_0^T x(u-\tau)y(u)e^{\frac{u-T}{k}}du &\leq \int_0^T e^{\frac{u-T}{k}}du \\
&\leq ke^{\frac{u-T}{k}}\Big|_0^T \\
&\leq k\left[1-e^{\frac{-T}{k}}\right] \\
&< k.
\end{aligned}$$

Similarly for -1, and so we have that the error term is in the range $(-k, k)$. However, it is not true that for all values of τ the absolute value of the error term will be less than $|C_{xy}|$. To see this, consider the case where $x(u-\tau)y(u)$ is the function $S(u)$:

$$S(u) = \begin{cases} -1 & u < T/2 \\ 1 & u \geq T/2 \end{cases}$$

In this case, C_{xy} will be zero, and the error term will be $2k(e^{-\tau/k} - e^{(T/2-\tau)/k})$, and hence nonzero.

In the case of the stimuli used in this study, a 50 ms long segment has C_{xy} with mean absolute values in the range of 3.31 ± 0.50 for the raw stimuli. If the stimuli are first convolved with the reverse-correlation estimated impulse response, C_{xy} is in the range of 14.6 ± 3.9. Since k is expected to be related to the membrane time constant, and hence have a magnitude no greater than on the order of 10^{-2}, it is reasonable to expect that the error term is insignificant. Thus, the time-averaged response of a coincidence detector should behave as classic cross-correlation, and the Wiener-Khinchin theorem applies.

Bibliography

Albeck YA, Konishi M (1995). Reponses of neurons in the auditory pathway of the barn owl to partially correlated binaural signals. *J Neurophysiol* **74**: 1689–1700.

Arthur BJ (2005). Distribution within the barn owl's inferior colliculus of neurons projecting to the optic tectum and thalamus. *J Comp Neurol* **492**: 110–121.

Bala ADS, Spitzer MW, Takahashi TT (2003). Prediction of auditory spatial acuity from neural images on the owl's auditory space map. *Nature* **424**: 771–774.

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

Batra R, Yin TCT (2004). Cross correlation by neurons of the medial superior olive: A reexamination. *JARO* **5**: 238–252.

Beckius GE, Batra R, Oliver DL (1999). Axon from anteroventral cochlear nucleus that terminates in medial superior olive of the cat: Observations related to delay lines. *J Neurosci* **19**: 3146–3161.

Blauert J (1997). *Spatial hearing: The psychophysics of human sound localization.* Cambridge, Massachusetts: MIT Press.

Blauert J, Lindemann W (1986). Spatial-mapping of intracranial auditory events for various degrees of interaural coherence. *J Acoust Soc Am* **79**: 806–813.

Brenowitz EA (1983). The contribution of temporal song cues to species recognition in the red-winged blackbird. *Anim Behav* **31**: 1116–1127.

Brown CH, Beecher MD, Moody DB, Stebbins WC (1980). Localization of noise bands by old world monkeys. *J Acoust Soc Am* **68**: 127–132.

Brown CH, Schessler T, Moody D, Stebbins W (1982). Vertical and horizontal sound localization in primates. *J Acoust Soc Am* **72**: 1804–1811.

Cai H, Carney LH, Colburn HS (1998). A model for binaural response properties of inferior colliculus neurons. I. A model with interaural time difference-sensitive excitatory and inhibitory inputs. *J Acoust Soc Am* **103**: 475–493.

Carney LH, Yin TC (1989). Response of low-frequency cells in the inferior colliculus to interaural time differences of clicks: Excitatory and inhibitory components. *J Neurophysiol* **62**: 144–161.

Carr CE (1993). Delay line models of sound localization in the barn owl. *Amer Zool* **33**: 79–85.

Carr CE, Boudreau RE (1991). Central projections of auditory nerve fibers in the barn owl. *J Comp Neurol* **314**: 306–318.

Carr CE, Boudreau RE (1993). Organization of the nucleus magnocellularis and the nucleus laminaris in the barn owl: Encoding and measuring interaural time differences. *J Comp Neurol* **334**: 337–355.

Carr CE, Konishi M (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci USA* **85**: 8311–8315.

Carr CE, Konishi M (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci* **10**: 3277–3246.

Cohen YE, Knudsen EI (1995). Binaural tuning of auditory units in the forebrain archistriatal gaze fields of the barn owl: Local organization but no space map. *J Neurosci* **15**: 5152–5168.

Cohen YE, Knudsen EI (1996). Representation of frequency in the primary auditory field of the barn owl forebrain. *J Neurophysiol* **76**: 3682–3692.

Cohen YE, Miller GL, Knudsen EI (1998). Forebrain pathway for auditory space processing in the barn owl. *J Neurophysiol* **79**: 891–902.

Curtis WE (1952). *Quantitative studies of echolocation in bats (Myotis l. lucifugus); studies of vision of bats (Myotis l. lucifugus and Eptessicus f. fuscus) and quantitative studies of vision of owls (Tyto alba pratincola)*. Ph. D. thesis, Cornell University.

Dasika VK, White JA, Carney LH, Colburn HS (2005). Effects of inhibitory feedback in a network model of avian brain stem. *J Neurophysiol* **94**: 400–414.

Dayan P, Abbott LF (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, Massachusetts: MIT Press.

de Boer E, de Jongh HR (1978). On cochlear encoding: Potentialities and limitations of the reverse-correlation technique. *J Acoust Soc Am* **63**: 115–135.

Depireux DA, Simon JZ, Klein DJ, Shamma SA (2001). Spectro-temporal response field characterizations with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* **85**: 1220–1234.

Dice LR (1945). Minimum intensities of illumination under which owls can find dead prey by sight. *Am Nat* **79**: 384–416.

Drullman R, Festen HM, Plomp R (1994). Effect of temporal envelope smearing on speech recognition. *J Acoust Soc Am* **95**: 1053–1064.

du Lac S, Knudsen EI (1990). Neural maps of head movement vector and speed in the optic tectum of the barn owl. *J Neurophysiol* **63**: 695–702.

Eggermont J (1993). Wiener and Volterra analyses applied to the auditory system. *Hearing Res* **10**: 191–202.

Egnor R (2000). *Cellular mechanisms for resolving phase ambiguity in the owl's inferior colliculus*. Ph. D. thesis, California Institute of Technology.

Feldman DE, Knudsen EI (1994). NMDA and non-NMDA glutamate receptors in auditory transmission in the barn owl inferior colliculus. *J Neurosci* **14**: 5939–5958.

Fujita I, Konishi M (1991). The role of gabaergic inhibition in processing of interaural time difference in the owl's auditory system. *J Neurosci* **11**: 722–739.

Gaese BH, Wagner H (1992). Precognitive and cognitive elements in sound localization. *Zoology* **105**: 329–339.

Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature* **383**: 76–78.

Goldberg JM, Brown PB (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms for sound localization. *J Neurophysiol* **32**: 613–636.

Grau-Serrat V, Carr CE, Simon JZ (2003). Modeling coincidence detection in nucleus laminaris. *Biol Cybern* **89**: 388–396.

Greenfield MD, Baker M (2003). Bat avoidance in non-aerial insects: The silence response of signaling males in an acoustic moth. *Ethology* **109**: 427–442.

Guinan, Jr. JJ, Norris BE, Guinan SS (1972). Single auditory units in the superior olivary complex. II. Locations of units categories and tonotopic organization. *Int J Neurosci* **4**: 147–166.

Han Y, Colburn HS (1993). Point-neuron model for binaural interaction in MSO. *Hearing Res* **68**: 115–130.

Hancock KE, Delgutte B (2004). A physiologically based model of interaural time difference discrimination. *J Neurosci* **24**: 7110–7117.

Harper NS, McAlpine D (2004). Optimal neural population coding of an auditory spatial cue. *Nature* **430**: 682–686.

Hecht S, Pirenne MH (1940). The sensibility of the nocturnal long-eared owl in the spectrum. *J Gen Physiol* **23**: 709–717.

Heffner RS, Heffner EH (1992). Evolution of sound localization in mammals. In A. N. Popper, R. R. Fay, and D. Webster (eds.), *The evolutionary biology of hearing*, pp. 691–715. New York, New York: Springer.

Hermes DJ, Aertsen AM, Johannesma PIM, Eggermont JJ (1981). Spectrotemporal characteristics of single units in the auditory midbrain of the lightly anaesthetised grass frog (*Rana temporaria*

L) investigated with noise stimuli. *Hearing Res* **5**: 147–178.

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

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

Johnson DH, Ray W (2004). Optimal stimulus coding by neural populations using rate codes. *J Comput Neurosci* **16**: 129–138.

Joris PX (2003). Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* **23**: 6345–6350.

Keller CH, Takahashi TT (2000). Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *J Neurophysiol* **84**: 2638–2650.

Kim PJ, Young ED (1994). Comparaive analysis of spectro-temporal receptive fields, reverse correlation functions, and frequency tuning curves of auditory-nerve fibers. *J Acoust Soc Am* **95**: 410–422.

King AJ, Hutchings ME (1987). Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: A map of auditory space. *J Neurophysiol* **57**: 596–624.

Klein DJ, Depireux DA, Simon JZ, Shamma SA (2000). Robust spectrotemporal reverse correlation for the auditory system: Optimizing stimulus design. *J Comput Neurosci* **9**: 85–111.

Knudsen EI (1983). Early auditory experience aligns the auditory map of space in the optic tectum of the barn owl. *Science* **218**: 174–186.

Knudsen EI, Blasdel GG, Konishi M (1979). Sound localization by the barn owl (*Tyto alba*) measured with the search coil technique. *J Comp Physiol* **133**: 1–11.

Knudsen EI, Brainard MS (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. *Science* **253**: 85–87.

Knudsen EI, Cohen YE, Masino T (1995). Characterization of a forebrain gaze field in the archistriatum of the barn owl: Microstimulation and anatomical connections. *J Neurosci* **15**: 5139–5151.

Knudsen EI, Knudsen PF (1983). Space-mapped auditory projections from the inferior colliculus to the optic tectum in the barn owl (*Tyto alba*). *J Comp Neurol* **218**: 187–196.

Knudsen EI, Knudsen PF (1985). Vision guides the adjustment of auditory localization in young barn owls. *Science* **230**: 545–548.

Knudsen EI, Konishi M (1978a). A neural map of auditory space in the owl. *Science* **200**: 795–797.

Knudsen EI, Konishi M (1978b). Space and frequency are represented seperately in auditory midbrain of the owl. *J Neurophys* **41**: 870–884.

Knudsen EI, Konishi M (1979). Mechanisms of sound localization in the barn owl (*Tyto alba*). *J Comp Physiol* **133**: 13–21.

Koch C (1999). *Biophysics of computation: Information processing in the single neuron*. New York, New York: Oxford University Press.

Konishi M (1973). How the owl tracks its prey. *Am Scient* **61**: 414–424.

Konishi M, Kenuk AS (1975). Discrimination of noise spectra by memory in the barn owl. *J Comp Physiol* **97**: 55–58.

Köppl C (1997a). Frequency tuning and spontaneous activity in the auditory nerve and cochlear nucleus magnocellularis in the barn owl *Tyto alba*. *J Neurophysiol* **77**: 364–377.

Köppl C (1997b). Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*. *J Neurosci* **17**: 3312–3321.

Lachica EA, Rübsamen R, Rubel EW (1994). GABAergic terminals in nucleus magnocellularis and laminaris originate from the superior olivary nucleus. *J Comp Neurol* **348**: 403–418.

Latham PE, Nirenberg S (2005). Synergy, redundancy, and independence in population codes, revisited. *J Neurosci* **25**: 5195–5206.

Lewis ER, Henry KR, Yamada WM (2002). Tuning and timing of excitation and inhibition in primary auditory nerve fibers. *Hearing Res* **171**: 13–31.

Licklider JCR (1959). Three auditory theories. In S. Koch (ed.), *Psychology: A study of a science*,

pp. 41–144. New York, New York: McGraw-Hill.

Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM (2003). Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol* **90**: 2660–2675.

Linden JF, Schreiner CE (2003). Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. *Cerebral Cortex* **13**: 83–89.

Louage DHG, van der Heijden M, Joris PX (2004). Temporal properties of responses to broadband noise in the auditory nerve. *J Neurophysiol* **91**: 2051–2065.

Manley GA, Köppl C, Konishi M (1988). A neural map of interaural intensity differences in the brain stem of the barn owl. *J Neurosci* **8**: 2665–2676.

Masino T, Knudsen EI (1990). Horizontal and vertical components of head movement are controlled by distinct neural circuits in the barn owl. *Nature* **345**: 434–437.

Matthews LH, Matthews BHC (1939). Owls and infra-red radiation. *Nature* **143**: 983.

Mazer JA (1995). *Integration of parallel processing streams in the inferior colliculus of the Barn Owl*. Ph. D. thesis, California Institute of Technology.

Mazer JA (1998). How the owl resolves auditory coding ambiguity. *Proc Natl Acad Sci USA* **95**: 10932–10937.

McAlpine D, Jiang D, Palmer AR (2001). A neural code for low-frequency sound localization in mammals. *Nat Neurosci* **4**: 396–401.

Middlebrooks JC, Green DM (1991). Sound localization by human listeners. *Annu Rev Psychol* **42**: 135–159.

Middlebrooks JC, Knudsen EI (1984). A neural code of auditory space in the cat's superior colliculus. *J Neurosci* **4**: 2621–2634.

Miller GL, Knudsen EI (2001). Early auditory experience induces frequency-specific, adaptive plasticity in the forebrain gaze fields of the barn owl. *J Neurophysiol* **85**: 2184–2194.

Moiseff A (1989). Bi-coordinate sound localization by the barn owl. *J Comp Physiol A* **164**:

637–644.

Moiseff A, Konishi M (1981). The owl's interaural pathway is not involved in sound localization. *J Comp Physiol* **144**: 299–304.

Moiseff A, Konishi M (1983). Binaural characteristics of units in the owl's brainstem auditory pathway: Precursors of restricted spatial receptive fields. *J Neurosci* **3**: 2553–2562.

Mori K (1997). Across-frequency nonlinear inhibition by GABA in processing of interaural time difference. *Hearing Res* **111**: 22–30.

Müller R, Schnitzler HU (1999). Acoustic flow perception in cf-bats: Properties of the available cues. *J Acoust Soc Am* **105**: 2958–2966.

Norberg NA (1977). Occurence and independent evolution of bilateral ear asymmetry in owls and implications on owl taxonomy. *Phil Trans Roy Soc Lond B* **280**: 375–408.

Panzeri S, Treves A, Schultz S, Rolls ET (1999). On decoding the responses of a population of neurons from short time windows. *Neural Comput* **11**: 1553–1577.

Payne RS, Drury W (1958). Marksman of the darkness. *Nat Hist* **67**: 316–323.

Peña JL, Konishi M (2000). Cellular mechanisms for resolving phase ambiguity in the owl's inferior colliculus. *Proc Natl Acad Sci USA* **97**: 11787–11792.

Peña JL, Konishi M (2002). From postsynaptic potentials to spikes in the genesis of auditory spatial receptive fields. *J Neurosci* **22**: 5652–5658.

Peña JL, Viete S, Albeck Y, Konishi M (1996). Tolerance to sound intensity of binaural coincidence detection in the nucleus laminaris of the owl. *J Neurosci* **16**: 7046–7054.

Peña JL, Viete S, Funabiki K, Saberi K, Konishi M (2001). Cochlear and neural delays for coincidence detection in owls. *J Neurosci* **21**: 9455–9459.

Pettigrew JD, Konishi M (1976). Neurons selective for orientation and binocular disparity in visual wulst of barn owl *(Tyto alba). Science* **193**: 675–678.

Pollack I, Rose M (1967). Effect of head movement on the localization of sounds in the equatorial

plane. *Percept Psychophys* **2**: 591–596.

Popper AN, Fay RR, editors (1995). *Hearing by bats*. New York, New York: Springer-Verlag.

Populin LC, Yin TCT (1998). Pinna movements of the cat during sound localization. *J Neurosci* **18**: 4233–4243.

Proctor L, Konishi M (1997). Representation of sound localization cues in the auditory thalamus of the barn owl. *Proc Natl Acad Sci USA* **94**: 10421–10425.

Quine DB, Konishi M (1974). Absolute frequency discrimination in the barn owl. *J Comp Physiol* **93**: 347–360.

Rayleigh (1876). Our perception of the direction of a source of sound. *Nature* **14**: 32–33.

Rayleigh (1907). On our perception of sound direction. *Philos Mag* **13**: 214–232.

Ross S (1994). *A first course in probability* (4th ed.). New York, New York: Macmillan College Publishing Company.

Ruggero MA, Rich NC, Recico A, Narayan SS, Robles L (1997). Basical-membrane responses to tones at the base of the chinchilla cochlea. *J Acoust Soc Am* **101**: 2151–2163.

Rust NC, Schwartz O, Movshon JA, Simoncelli E (2004). Spike-triggered characterizations of excitatory and suppressive stimulus dimensions in monkey V1. *Neurocomputing* **58–60**: 793–799.

Saberi K, Takahashi Y, Farahbod H, Konishi M (1999). Neural basis of an auditory illusion and its elimination in owls. *Nat Neurosci* **2**: 656–659.

Saberi K, Takahashi Y, Konishi M, Albeck Y, Arthur BJ, Farahbod H (1998). Effects of interaural decorrelation on neural and behavioral detection of spatial cues. *Neuron* **21**: 789–798.

Sahani M, Dayan P (2003). Doubly distributional population codes: Simultaneous representation of uncertainty and multiplicity. *Neural Comput* **15**: 2255–2279.

Schnitzler HU, Kalko EKV (2001). Echolocation by insect-eating bats. *BioScience* **51**: 557–569.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995). Speech recognition with primarily temporal cues. *Science* **270**: 303–304.

Shaw EAG (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J Acoust Soc Am* **56**: 1848–1861.

Smith PH, Joris PX, Yin TCT (1993). Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: Evidence for delay lines in the medial superior olive. *J Comp Neurol* **331**: 245–260.

Speakman JR, Anderson MF, Racey PA (1989). The energy-cost of echolocation in pipistrelle bats *(Pipistrellus-pipistrellus)*. *J Comp Physiol* **165**: 679–685.

Spezio ML, Takahashi TT (2003). Frequency-specific interaural level difference tuning predicts spatial response patterns of space-specific neurons in the barn owl inferior colliculus. *J Neurosci* **23**: 4677–4688.

Sterbing SJ, Hartung K, Hoffmann KP (2002). Representation of sound source direction in the superior colliculus of the guinea pig in a virtual auditory environment. *Exp Brain Res* **142**: 570–577.

Sterbing SJ, Hartung K, Hoffmann KP (2003). Spatial tuning to virtual sounds in the inferior colliculus of the guinea pig. *J Neurophysiol* **90**: 2648–2659.

Sullivan WE, Konishi M (1984). Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl. *J Neurosci* **4**: 1787–1799.

Sullivan WE, Konishi M (1986). Neural map of interaural phase difference in the owl's brainstem. *Proc Natl Acad Sci USA* **83**: 8400–8404.

Takahashi T, Konishi M (1986). Selectivity for interaural time difference in the owl's midbrain. *J Neurosci* **6**: 3413–3422.

Takahashi T, Konishi M (1988a). Projections of nucleus angularis and nucleus laminaris to the lateral lemniscal nuclear complex of the barn owl. *J Comp Neurol* **274**: 212–238.

Takahashi T, Konishi M (1988b). Projections of the cochlear nuclei and nucleus laminaris to the inferior colliculus of the barn owl. *J Comp Neurol* **274**: 190–211.

Takahashi TT, Keller CH (1992). Commissural connections mediate inhibition for the computation of interaural level difference in the barn owl. *J Comp Physiol* **170**: 161–169.

Takahashi TT, Wagner H, Konishi M (1989). Role of commissural projections in the representation of bilateral audtiory space in the barn owl's inferior colliculus. *J Comp Neurol* **281**: 545–554.

Thurlow WR, Runge PS (1967). Effect of induced head movements on localization of direction of sounds. *J Acoust Soc Am* **42**: 480–488.

van de Willigen RF, Frost BJ, Wagner H (1998). Stereoscopic depth perception in the owl. *Neuroreport* **9**: 1233–1237.

Vanderplank FL (1994). The effect of infra-red waves on tawny owls *(Strix aluco)*. *Proc Zool Soc Lond*: 505–507.

Venturi JB (1796). Considerations sur la connaisance de l'entendue que nos donne le sens do l'ouie. *Magasin encyclopedie, ou journal des sciences, des lettres et des arts* **3**: 29–37.

Viete S, Peña JL, Konishi M (1997). Effects of interaural intensity difference on the processing of interaural time difference in the owl's nucleus laminaris. *J Neurosci* **17**: 1815–1824.

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

Wagner H (1990). Rceptive fields of neurons in the owl's auditory brainstem change dynamically. *Eur J Neurosci* **2**: 949–959.

Wagner H, Mazer JA, von Campenhausen M (2002). Response properties of neurons in the core of the central nucleus of the inferior colliculus of the barn owl. *Eur J Neurosci* **15**: 1343–1352.

Wagner H, Takashi T, Konishi M (1987). Representations of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *J Neurosci* **7**: 3105–3116.

Waters DA (2003). Bats and moths: What is there left to learn? *Physiol Ent* **28**: 237–250.

Wild JM, Kubke MF, Carr CE (2001). Tonotopic and somatotopic representation in the nucleus

basalis of the barn owl, *Tyto alba*. *Brain Behav Evol* **57**: 39–62.

Wright BA, Lombardino LJ, King WM, Puranik CS, Leonard CM, Merzenich MM (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature* **387**: 176–178.

Yamada WM, Lewis ER (1999). Predicting the temporal responses of non-phase-locking bullfrog auditory units to complex auditory waveforms. *Hearing Res* **130**: 155–170.

Yin TCT, Chan JCK (1990). Interaural time sensitivity in the medial superior olive of the cat. *J Neurophysiol* **64**: 465–488.

Yin TCT, Chan JCK, Carney LH (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J Neurosci* **58**: 562–583.

Yin TCT, Chan JCK, Irvine DRF (1986). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise. *J Neurophysiol* **55**: 280–300.

Zhang XD, Heinz MG, Bruce IC, Carney LH (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J Acoust Soc Am* **109**: 648–670.