

Chapter 8

The Neural Processing of Frequency Modulations in the Auditory System of Bats

George D. Pollak

Keywords Communication calls • Echolocation • Frequency modulation • Inferior colliculus • Inhibition

8.1 Introduction

Bats rely to an inordinate degree on sound for survival. Bats, of course, are best known for their ability to echolocate, a form of biosonar used for both orientation and detecting, identifying, and capturing prey. Not only do they rely on hearing for orientation and hunting through echolocation, but hearing is also critically important for social communication. Many bats live in large colonies where they engage in a myriad of social interactions that are accomplished almost entirely with sound since they live in dark environments where visual displays are of no use (Griffin 1958). Their communication signals can be quite elaborate (Bohn et al. 2009), and some species are capable of vocal learning (Boughman 1998; Knornschild et al. 2006). Indeed, the repertoire of signals bats use for vocal communication is rich and sophisticated (Kanwal 1999; Bohn et al. 2008).

This chapter is concerned with the neural processing of signals that bats employ in both their echolocation and communication calls. Particular attention is given to the processing of communication calls and the roles of inhibition in creating response selectivity that enables the auditory system to distinguish among the various signals bats receive. Vocal communication was presumably used by their ancestors before bats took to the night sky to exploit a food supply for which there was little competition. It is noteworthy in this regard that the primary acoustic feature of virtually all echolocation calls is a brief, frequency-modulated (FM) sweep (Simmons et al. 1975; Neuweiler 1990); strikingly similar FM sweeps are also acoustic features of many communication calls emitted by bats (Schwartz et al. 2007; Bohn et al. 2008). Thus, the view presented in this

G.D. Pollak (✉)

Department of Neuroscience and Center for Perceptual Systems, The University of Texas at Austin, Austin, TX 78712, USA

e-mail: gpollak@austin.utexas.edu

chapter is that one of the primary challenges that drove the evolution of their auditory systems was to process communication signals, and adaptations required for echolocation were subsequently added or co-opted to enable the various species of bats to compete successfully for food resources in a wide range of different habitats.

8.1.1 Some General Comments on Echolocation

Echolocation, while exotic, is not as unique a perceptual ability as many believe. It evolved several times and is present in two species of birds, in cetaceans, in one species of megachiropteran bat, the Egyptian Rousette bat, *Rossetus aegyptiacus*, as well as in all microchropteran bats (Griffin 1958). The proposition that echolocation is not a unique perception that required new and special modes of processing is supported by the fact that humans can learn to echolocate with high precision (Thaler et al. 2011).

There have always been anecdotal stories about one or another blind person who displayed echolocation abilities that were so good that they appeared not to be blind at all (Griffin 1944, 1958). Recently, one person, Daniel Kish, has received particular attention. Kish has been sightless since he was a year old, yet he can mountain bike, navigate the wilderness alone, and recognize a building hundreds of feet away, all with echolocation. You can see his abilities to recognize buildings and other objects at <http://www.youtube.com/watch?v=PGMpswJtCdI&feature=fvwrel>. He echolocates by emitting clicks with his tongue, the same way birds and tomb bats do, and he can form remarkably precise images from the echoes he receives. The progress Kish has made and how he is teaching other blind people to use echolocation can be seen at http://www.youtube.com/watch?v=CRA-asTuP_Y and at <http://www.youtube.com/watch?v=xATiyq3uZM4&feature=related>.

Studies on echolocation in the blind suggest that the echoes they perceive evoke some activity in the visual portions of the cortex (Thaler et al. 2011). The importance of visual involvement is intriguing, but unclear, since these people did not report that they experienced “visual” images while echolocating, even though they had sight early in life. Nevertheless, they could form images of objects in the environment through sound, which shows that whatever circuits and mechanisms that enable bats to form images of objects in their environment by listening to echoes are also present in humans. Consistent with this notion is the evidence that the auditory systems of bats are similar to the auditory systems of all other mammals, with the same structures, wiring, and mechanisms for processing information that are possessed by other mammals (Pollak and Casseday 1986; Pollak et al. 1995; Winer et al. 1995). The distinguishing features of the auditory system of bats are not novel mechanisms. Rather, the distinguishing feature is that some mechanisms, which are common to all mammals, are far more pronounced in their auditory systems than in other mammals. The difference is quantity, not quality.

8.1.2 Themes of the Chapter

In the first section of the chapter, some of the communication signals bats emit are presented with the behaviors during which the signals are emitted to illustrate the richness and variety of their vocal repertoires. The subsequent sections then deal with the neural processing of communication calls, focusing on the processing of FM sweeps. Additionally, focus is on processing of FM sweeps in the inferior colliculus (IC), the midbrain auditory nucleus that receives projections from almost all lower auditory nuclei and then synthesizes that convergent information (Pollak and Casseday 1986; Casseday et al. 2002). The net result of those syntheses is that a variety of new response properties are either formed de novo in the IC or other response properties that have been formed in lower nuclei are sharpened or further modified in the IC.

The following sections on neural processing have three themes. The first theme is that the various computations employed by the IC endow its neurons with selectivities for features of FM sweeps, and those selectivities, in turn, are a principal feature for creating the response selectivities for echolocation as well as the various conspecific communication calls these animals hear in their daily lives. The second theme is that inhibition is the major sculptor in the auditory system, where inhibition acts to shape selective response properties out of excitatory inputs that are far less selective or even unselective. The third theme is that the IC population is heterogeneous, in that some cells form their selective response properties in a straightforward, linear manner, whereas other cells form similar response properties through more complex, non-linear processing.

8.2 The Vocal Repertoire of Bats

Before turning to the neural mechanisms by which acoustic signals are processed, the variety of communication calls bats use are considered first, as illustrated by the vocal repertoire of Brazilian free-tailed bats, *Tadarida brasiliensis*. These bats are common in the Southwestern United States where they live in caves with populations that often number in the millions. Here males use vocal signals to establish dominance hierarchies, maintain territories, garner females into harems, and defend their harems against intruding males. Females use vocal signals for recognition of and bonding with their pups, among other behaviors (Schwartz et al. 2007; Bohn et al. 2008).

To give a flavor of the variety of calls emitted by these animals, a sample of calls is shown in Figure 8.1, together with a notation about the behaviors the bats displayed during the emission of each call type. Each call is composed of one or more repetitions of a syllable or note. Each syllable is composed of multiple harmonics with spectral components that change in amplitude, and often in frequency, throughout its duration. The syllables range not only in duration (from 2

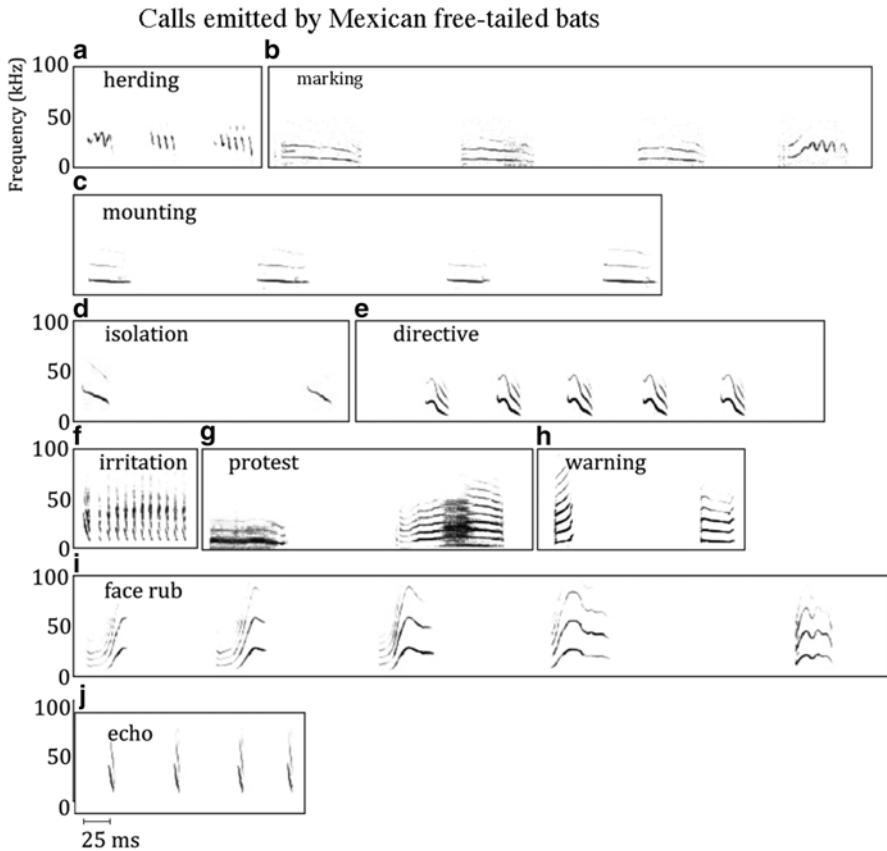


Fig. 8.1 Spectrograms that show the various communication calls emitted by Mexican free-tailed bats. **(a)** *Herding Calls* were emitted while males forcefully pushed one or more females with muzzle or wing into his territory. **(b)** *Marking Calls* were emitted by dominant males while rubbing their faces and gular glands on the surfaces of their territories. **(c)** *Mounting Calls* were emitted by males to convey dominance, when males would mount females and forcefully push their muzzles repeatedly between their shoulders. **(d)** *Isolation Calls* were emitted by pups immediately after birth and throughout development; pups called when they were isolated or hungry. **(e)** *Directive Calls* were emitted by females while giving birth and throughout pup development when females approached pups or in response to their pups' isolation calls. **(f)** *Irritation Calls* were emitted when bats were jostled by other bats. **(g)** *Protest Calls* were emitted in response to aggressive behaviors by other bats. **(h)** *Warning Calls* were emitted prior to aggressive encounters. **(i)** *Face-Rubbing Calls* were used for social bonding; they were emitted in roost sites while approaching another bat and rubbing their muzzles across the body of the bat. **(j)** Echolocation calls were emitted for orientation while the bats were flying (Reproduced from Bohn et al. 2008 with permission; ©Acoustical Society of America)

to 3 ms to over 100 ms) but also in their spectral structures. For example, some syllables are simply brief downward sweeping frequency modulations (FMs) (e.g., irritation call in Figure 8.1f), whereas others are more complex with both upward and downward FMs (e.g., directive, Figure 8.1e, and herding calls, Figure 8.1a), and yet others have only harmonic stacks of constant frequencies

(marking, Figure 8.1b, and mounting calls, Figure 8.1c). The temporal sequence in which the syllables are emitted is also an important feature that varies with behavioral context (Bohn et al. 2008). The syllables produced in several different calls associated with completely different contexts are indistinguishable except for differences in temporal intervals or the repetition rate at which the syllables are emitted.

The Brazilian free-tailed bats not only emit the simpler types of calls illustrated in Figure 8.1, they also sing elaborate “songs.” During the breeding season, free-tailed bats, as well as many other animals, emit simple repetitions of one or a few syllables that are generally referred to as mating or advertisement calls. In a few exceptional cases, such as songbirds (Catchpole and Slater 1995), whales (Payne and McVay 1971), and some bat species (Behr and von Helversen 2004; Bohn et al. 2009), these advertisement signals can be more complex vocalizations termed songs. The major difference between mating “calls” and “songs” is that songs are longer and more complex and contain multiple types of elements (e.g., syllables or notes) that are combined in a stereotypical manner (Catchpole and Slater 1995; Marler 2004). Therefore, songs have an added dimension of complexity in the form of syntax, the patterns by which elements are ordered and combined. Indeed, in most songs, element ordering is not random but instead is highly structured, with individual, regional, and/or species-specific patterns (Balaban 1988).

A remarkable feature of Brazilian free-tailed bats is that the ways in which phrases are combined to form songs follow broad syntactical rules, yet males dynamically vary phrase order from one rendition to the next. During the breeding season dominant males sing their courtship songs (Figure 8.2A) when females approach their territories (Schwartz et al. 2007; Bohn et al. 2009). Their courtship songs are composed of three types of phrases. A phrase is composed of one or more syllable types that form a distinct and reproducible unit, and the phrases are combined to form songs. The three phrases are chirps, trills, and buzzes (Figure 8.2A). Chirps are phrases composed of two types of syllables: “A” and “B” syllables. The A syllables are short (~5 ms) downward sweeping FMs. The B syllables are longer (~15 ms) and more complex than A syllables. The B syllables often begin with an upward FM followed by a longer downward FM, and some signals end with a second upward FM. Thus, their spectral contours often have multiple inflection points. Several A syllables always precede each B syllable, and the sequence of several A syllables followed by a B syllable is then repeated to form the chirp phrase.

The second type of phrase is the trill. Trills are composed of short (3–4 ms), downward FM syllables that are sometimes connected, resulting in sinusoidal patterns (Figure 8.2Ac). Trill syllables, whether discrete or connected, are produced as a distinct phrase or burst with durations of approximately 25 ms. Sequential trill phrases are often emitted in songs, but they are highly distinctive since each phrase is separated from the next by a silent interval that is much greater than the interval between syllables within each trill phrase.

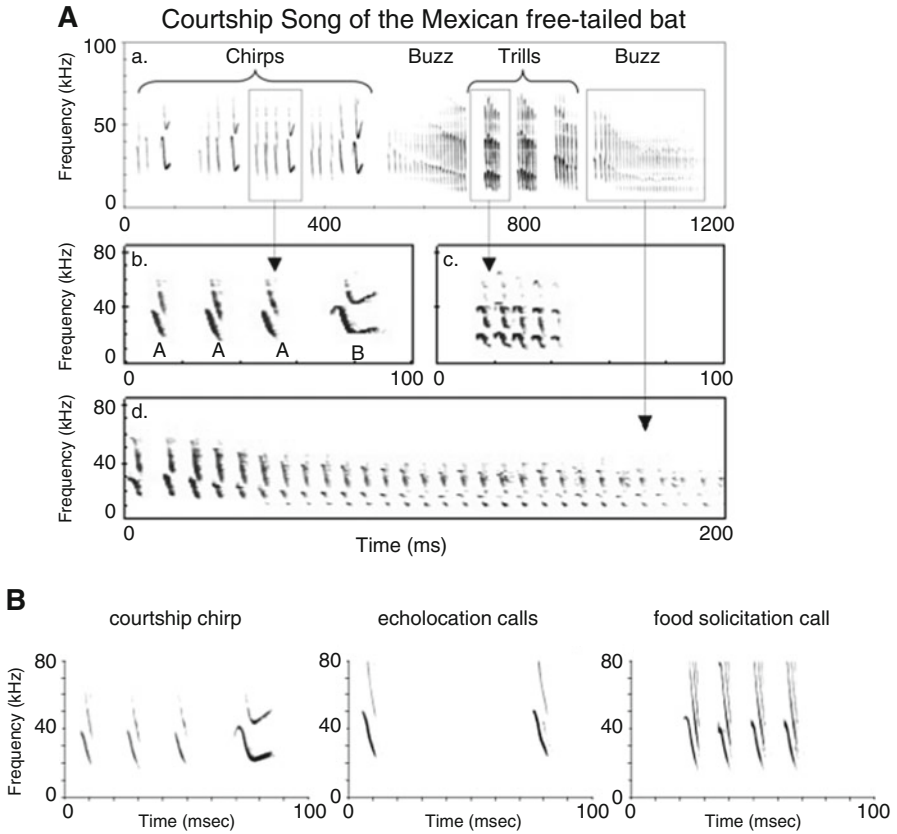


Fig. 8.2 Communication and echolocation calls emitted by Brazilian free-tailed bats. **(A)** The courtship song of a Brazilian free-tailed bat: **(a)** one complete song showing the three types of phrases: a chirp, buzzes, and trills; **(b)** Expanded section of a chirp phrase showing the A and B syllables; **(c)** Expanded section of a trill; **(d)** Expanded section of a buzz (Adapted from Bohn et al. (2009)). **(B)** The FM sweeps emitted in courtship chirps, echolocation calls, and food solicitation calls are strikingly similar

The third phrase in this song is the buzz (Figure 8.2Ad). Buzzes are also composed of short (3-ms) downward FM syllables that are always separated by a few milliseconds. Although the acoustical structures of trill and buzz syllables are similar, the phrases are distinguished by the number of syllables they contain: Trills have only 3–4 syllables, whereas buzzes have on average 35 syllables. They are also distinguished by the spectral structure of the syllables. The initial FM syllables in each buzz have relatively high beginning and end frequencies and

are followed by 5–10 syllables with progressively lower beginning and end frequencies (Figure 8.2Ad).

A significant feature of some of the FMs in the communication calls is that they are strikingly similar to the FMs these bats emit in their echolocation calls (Figure 8.2B) (Schwartz et al. 2007; Bohn et al. 2009). The similarity is in both the spectral and temporal structure of the FMs emitted in the two different contextual situations. An example is the similar structures of the FMs of the A syllables of the courtship songs, the FM sweeps of the food solicitation call, and the echolocation FMs emitted during certain phases of the echolocation cycle (Figure 8.2B). In all cases, the FMs are only about 1–3 ms in duration with fundamental frequencies that sweep downward from about 30–15 kHz. It follows then that the auditory system must process the two types of signals in the same way. The similarity in the structural features also underscores the hypothesis presented earlier, that the processing of communication signals is one of the primary tasks of the bats' auditory system, and adaptations required for echolocation were subsequently added or co-opted from the basic processing of the elements in communication calls. In the sections below, how the auditory system in bats processes and represents the various communication calls and songs they emit are discussed first and then attention is turned to the processing of brief FM sweeps.

8.3 Responses in the Colliculus Are Selective

When a series of echolocation and conspecific communication calls is presented to a group of isofrequency IC neurons (i.e., neurons tuned to the same frequency), most neurons respond to only a subset of the calls and not to others, and thus most neurons express response selectivity (Andoni et al. 2007; Andoni and Pollak 2011). Selectivity of this sort is seen in the IC of all mammals that have been studied (Holmstrom et al. 2007; Portfors et al. 2009). Selectivity is illustrated in Figure 8.3, which shows a suite of ten species-specific communication and echolocation calls from a Brazilian free-tailed bat and the responses that were evoked from four IC cells that were all tuned to about the same frequency. Each call had a different and unique spectrotemporal structure, was broadband with multiple harmonics, and was presented at an intensity that was at least 20 dB above the neuron's threshold at the frequency to which the neuron was most sensitive (its best frequency, BF). Thus, each call had suprathreshold energy that encroached upon each neuron's excitatory tuning curve. The differential responses to each of the calls showed that IC cells are not only selective but also that their selectivities are diverse in that the particular subset of calls that evoke discharges varied from neuron to neuron, even though the neurons were all tuned to the same frequency.

The selectivity for calls is shaped in the IC by the interaction of the excitatory and the inhibitory innervation that plays upon each IC cell (Klug et al. 2002;

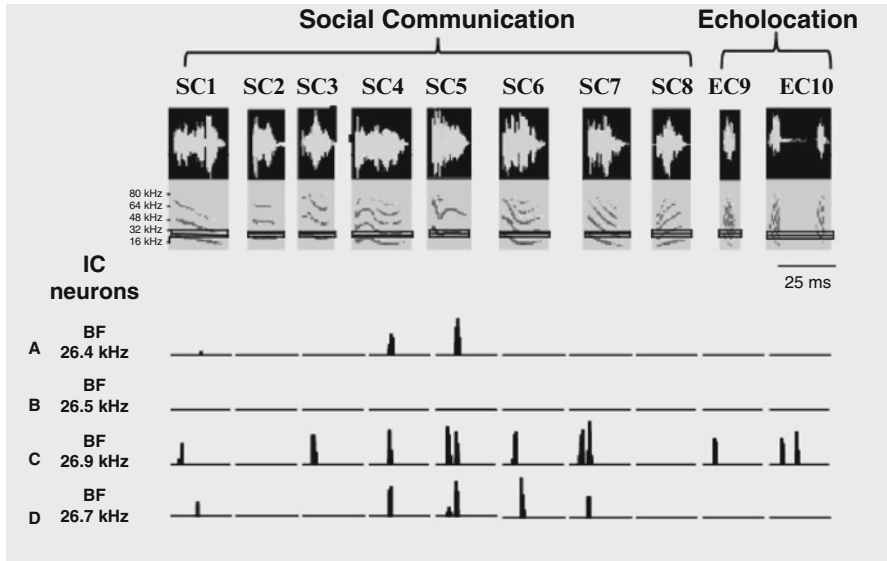


Fig. 8.3 Responses of four IC neurons to ten species-specific calls. Eight of the calls (SC1–SC8) are social communication calls and two others (EC9–EC10) are echolocation calls. The four IC cells are isofrequency and all tuned to about 26 kHz. The IC cells were selective in that each fired to only a subset of the ten calls although each of the calls had suprathreshold energy that swept through each neuron’s excitatory tuning curve. The selectivity was also heterogeneous in that each cell fired to a particular subset of calls that was different from the subset to which the other cells fired. One cell failed to fire to any of the calls (Adapted from Klug et al. 2002)

Xie et al. 2005). Although each IC neuron receives both excitatory and inhibitory innervation, the excitatory inputs are either non-selective or only slightly selective. Indeed, inhibition is the major sculptor in the IC, where inhibition shapes a selective response property out of a far less selective excitatory input.

The profound impact of inhibition on the selective responses evoked by communication calls is illustrated in Figure 8.4, which shows the responses of nine IC neurons to two different calls, social communication call 4 (SC4) and social communication call 6 (SC6). The responses evoked by the two calls were recorded before and while inhibition was blocked by the iontophoretic application of bicuculline, a drug that selectively blocks GABA_A receptors, and/or strychnine, which selectively blocks glycine receptors. Each of the nine neurons had a different BF and the cells are arranged from low to high, which corresponds to the tonotopic organization of the IC. Note that the calls had similar spectrotemporal features but evoked different responses among the population. Before inhibition was blocked, the nine neurons expressed different selectivities, since only three of the nine neurons responded to call SC4 and four different neurons responded to call SC6.

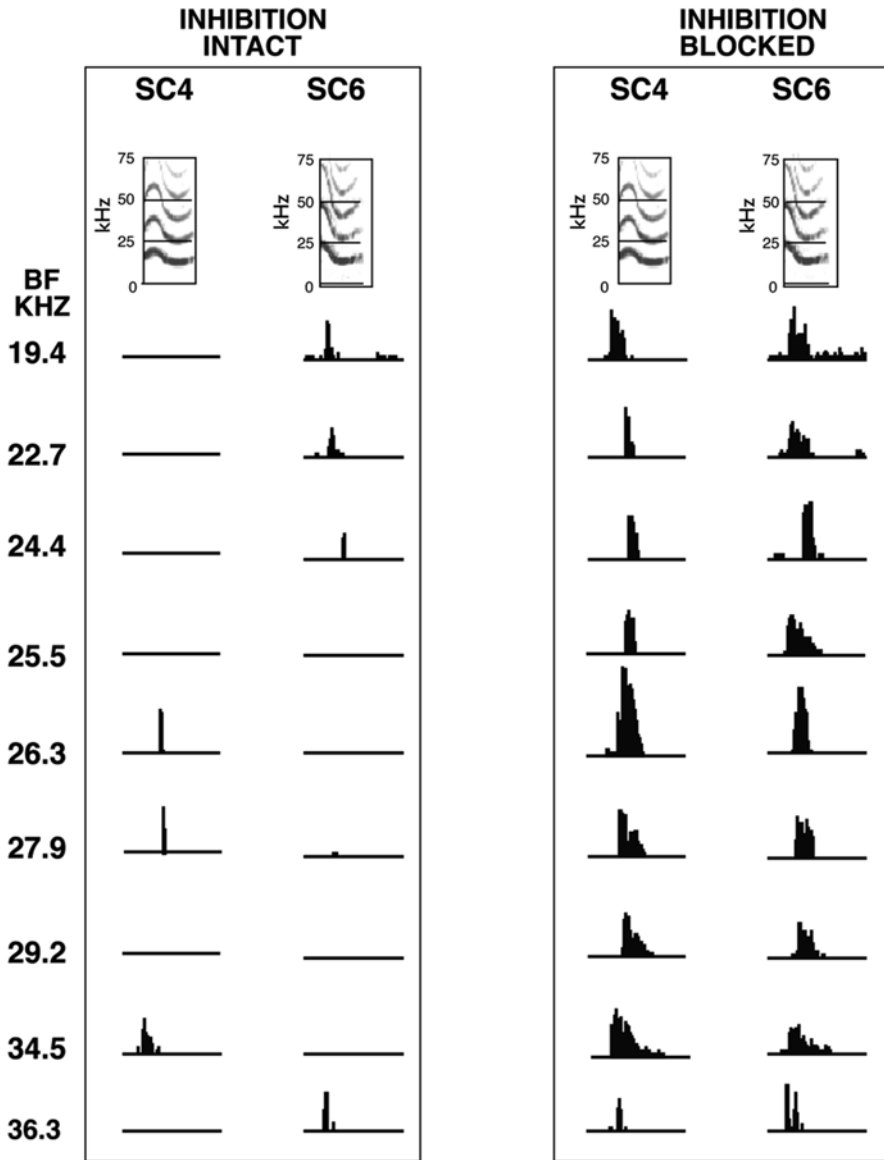


Fig. 8.4 Responses of nine IC neurons to two conspecific communication calls, SC4 and SC6, before and while inhibition was blocked. The BFs of the neurons are arranged from low to high, which corresponds to the tonotopic organization of the IC. Note that the calls had similar spectrotemporal features but evoked different responses among the population. Before inhibition was blocked, the selectivities of the two neurons were different, the three neurons that responded to call SC2 did not respond to SC4, and the four neurons that responded to SC6 did not respond to SC4. Blocking inhibition eliminated selectivities and all neurons responded to both signals (Adapted from Klug et al. 2002)

Blocking inhibition virtually eliminated selectivity and allowed all nine neurons to respond to both calls.

8.4 Spectrotemporal Receptive Fields Reveal the Importance of Sideband Inhibition

But what are the features of inhibition that act to shape response selectivity? The most important feature is the temporal and spectral structure of each neuron's sideband inhibition. Sideband, or surround inhibition as it is sometimes called, is composed of the frequencies that flank the excitatory frequency region of a neuron's tuning curve and evoke inhibition. Consistent with the results in Figure 8.4, other studies have also shown that when sideband inhibition is eliminated by the iontophoretic application of bicuculline and/or strychnine, IC neurons responded to many more calls than they did before inhibition was blocked or even responded to all of the calls presented (Andoni et al. 2007; Holmstrom et al. 2007; Mayko et al. 2012). Specifically, it must be the timing and magnitude of inhibition relative to excitation that underlies selectivity, but exactly how those features are expressed in each IC cell and how they differ among IC cells to create the diverse selectivities among isofrequency cells could not be determined from blocking inhibition alone.

To obtain a more detailed picture of both the excitatory and inhibitory fields in IC cells, a large number of complex signals called "moving ripples" were presented. Moving ripples are complex signals that contain a broad range of both spectral and temporal modulations that have been used by numerous investigators to generate spectrotemporal receptive fields (STRFs) (Kowalski et al. 1996). These signals were used to generate STRFs by a process analogous to spike-triggered averaging of the signals that preceded each spike (Andoni et al. 2007). The STRF derived from one IC cell is shown in Figure 8.5. The idea is that each ripple stimulus is a signal with a broad spectrum but unique spectrotemporal structure. When the different rippled stimuli are presented, frequencies in each ripple stimulus that are always present prior to a discharge are summed and thereby form the red region in the STRF. Frequencies that are rarely or never present prior to a discharge form the blue regions in the STRF. The frequencies represented in the red colors are presumed to be excitatory, whereas the frequencies in blue colors are presumed to be inhibitory. Moreover, whenever the neuron fires, some frequencies will be present in a random fashion, thereby generating the green background color in the STRF in Figure 8.5. Given these assumptions, the STRF provides a picture of relative magnitudes and temporal relationships of the excitation and inhibition that plays upon the cell.

If the STRF provides an accurate representation of the spectrotemporal arrangement of the excitatory and inhibitory innervation, then convolving the STRF with a suite of communication calls should yield predicted responses for that cell, and the predicted responses should be in close agreement with the responses that are actually evoked by each call. Convolution is a calculation in which two matrices are slid

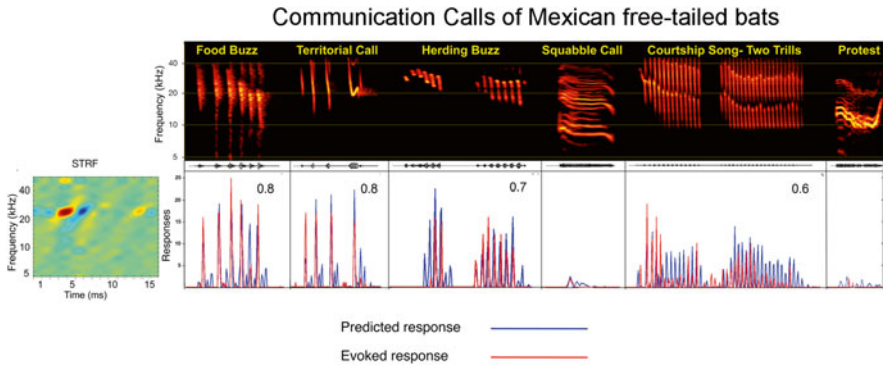


Fig. 8.5 The spectrotemporal receptive field (STRF) provides accurate predictions of responses to species-specific calls. The STRF derived from one cell in the inferior colliculus is shown on the far left. Spectrograms of each species-specific vocalization are shown in the top, with the evoked responses (*red*) and the responses predicted from the STRF (*blue*) displayed below each call. Convolution of the STRF with the spectrogram of each call generated predicted responses. The correlations between the predicted and actual responses are shown in *top right* of each panel. Convolutions predicted the call selectivity of the neuron because they predicted high response magnitudes for those calls that evoked strong responses, but they also predicted very low response magnitudes for the calls that evoked little or virtually no responses (Adapted from Andoni et al. 2007)

past each other: one constructed from the cell’s STRF and the other from the spectrogram of each call. At each point in time, the values of the two matrices are multiplied. If, at a point in time, a portion of the excitatory portion of the STRF overlaps with a part of the call spectrogram, a positive number is obtained. The positive number results from the multiplication of the energy in the call frequency at that point in time and the value of the corresponding excitation in the STRF. The value resulting from the multiplication indicates how strongly the neuron should have responded at that point in time. However, if the response of the STRF at that time point is inhibitory, the resultant value is negative and is recorded as zero, since there can be no negative number of spikes. Thus, the convolution takes into account not only whether the sound contains frequencies that enter the cell’s excitatory or inhibitory response regions but also the magnitudes of the excitation and inhibition at times when the spectral components of the sound and the excitatory and inhibitory regions of the STRF overlap. Assuming the neuron’s responses to complex calls are determined largely by the linear sum of responses evoked by the component frequencies in the call, the convolution of the STRF with each call provides a prediction of whether the neuron should respond to the call and, if so, how the neuron should respond in terms of relative response magnitude, latency, and temporal discharge pattern, i.e., it should predict the neuron’s discharge profile.

The STRF is, in essence, a linear filter that represents the optimal signal to which the neuron is tuned (Klein et al. 2000). Thus, the prediction is that the strongest responses should be evoked by stimuli that are most similar to the filter, the spectro-

temporal features of the neuron's STRF, and the more that the spectrotemporal features of the signal differ from the STRF, the weaker the predicted response. The responses predicted by the convolutions can then be compared or correlated with the discharge profile that is actually evoked by the same call, thereby providing a numerical value of correspondence.

In about 25 % of the IC cells, the responses evoked by the calls were accurately predicted by the convolutions (Andoni et al. 2007). An example is shown in Figure 8.5. The convolutions not only accurately predicted the calls to which the neurons responded, they also predicted the temporal discharge pattern evoked by each call. Equally important, they also predicted the calls to which the neurons did not respond. In short, the STRF in these cells captured the essential features of the cell and provided a picture of the relative magnitude and timing of excitation and inhibition, which in turn predicted how the cell would respond to any of the communication calls or to any other stimulus.

8.4.1 Spectrotemporal Receptive Fields Explain FM Directional and Velocity Selectivities

A basic response feature tested by FM sweeps is whether the neuron responds equally well to both upward and downward frequency sweeps or whether it responds only to, or most strongly to, one FM direction and thus expresses a directional preference or selectivity. Directional selectivity for FM sweeps is strongly influenced by inhibition, since blocking inhibition greatly reduces directional preferences in IC neurons (Casseday et al. 1997; Fuzessery et al. 2011). However, it is not inhibition per se that shapes directional selectivity. The important features are the relative timing of the excitation and inhibition. The timing is reflected in the tilting of the inhibitory fields along the spectrotemporal axis of the STRF (Andoni et al. 2007). Tilted inhibitory fields enhance directional preferences, or even create them, because signals sweeping in the non-preferred direction simultaneously evoke both excitation and inhibition, thereby suppressing responses to that FM direction (Figure 8.6, top panel). In contrast, signals sweeping in the preferred direction activate excitation and inhibition at different times, excitation first followed by inhibition, thereby allowing the cell to respond to the preferred direction (Figure 8.6, lower panel). This interpretation is supported by results obtained when inhibition was blocked by the iontophoretic application of bicuculline and/or strychnine (Andoni et al. 2007). Blocking inhibition not only reduced or even eliminated the inhibitory fields in their STRFs, but it also reduced FM directional selectivities in the IC.

The degree of tilt in the receptive field shapes both the neuron's directional selectivity and the FM velocity that evokes the strongest response (Andoni et al. 2007; Andoni and Pollak 2011). The response strength is determined by the correspondence between the tilt in the excitatory field and the rate of frequency sweep or FM velocity. Thus neurons with strong tilts are most sensitive to high FM velocities, whereas neurons with lesser tilts are most sensitive to lower FM velocities. Based

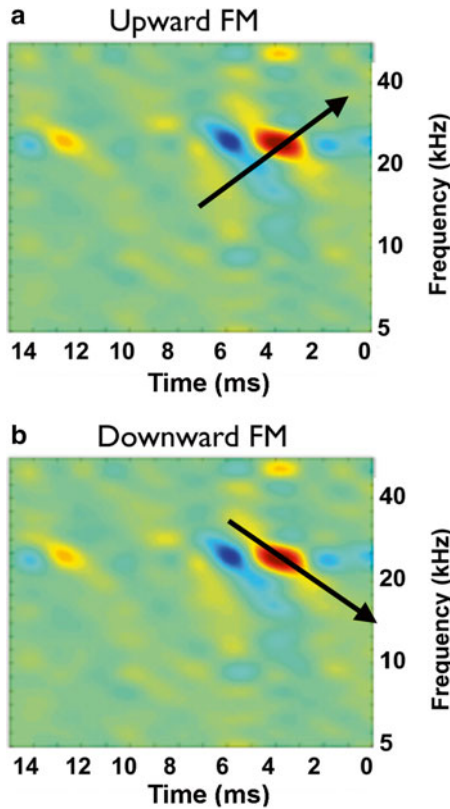


Fig. 8.6 Tilted spectrotemporal receptive fields impart directional selectivity for FM sweeps. Cell with a tilted (*inseparable*) receptive field. *Arrows* indicate how an *upward* (a) and *downward* (b) FM sweep would traverse the STRF at one point in time. The key feature is that at some point in time the downward FM will only sweep through the excitatory portion of the STRF without encroaching upon the inhibitory portion and thereby excite and drive the cell. In contrast, the upward FM will never encroach only upon the excitatory part of the STRF but rather will sweep through both its excitatory and inhibitory portions, which will suppress excitation thereby preventing the neuron from firing. This is the same STRF shown in Figure 8.5 but flipped in time (STRF is adapted from Andoni et al. 2007)

on these features, it was estimated that most IC neurons had best velocities between 5 and 100 octaves/s, with a mean of ~60 octaves/s (Figure 8.7c).

Of particular importance is the close agreement between the FM features in their conspecific communication sounds and the tuning for those FM features among the IC population (Andoni and Pollak 2011; Pollak et al. 2011a, b). As can be seen in the spectrograms of the various calls in Figures 8.1, 8.2, and 8.3, all echolocation and most communication signals emitted by Mexican free-tailed bats contain FMs. At least a portion of the FMs in almost all calls sweeps downward at velocities ranging from 0 to 250 octaves/s. Consistent with these signal features, the IC of all bats

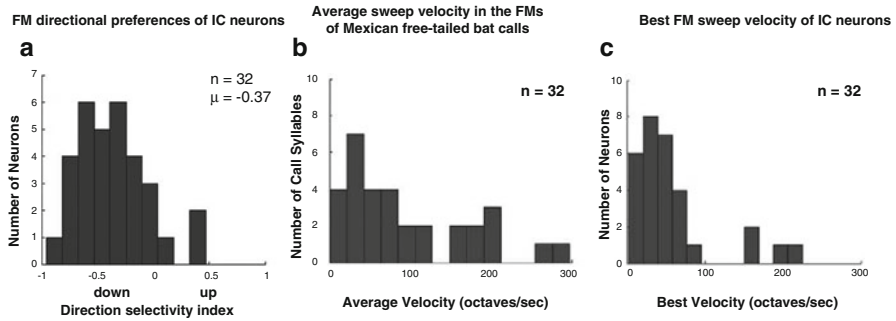


Fig. 8.7 Directional selectivities of IC cells, call velocities, and best velocities of IC cells. **(a)** Distribution of directional selectivity indices in IC cells. Downward preferring cells have negative indices, upward preferring cells have positive indices, and non-directional cells have an index of 0. A small number of cells preferred upward sweeps or were non-directional. The vast majority, however, preferred downward FMs. **(b)** Distribution of the FM velocities found in 21 calls. **(c)** Distribution of best velocities to which IC neurons are tuned. The distributions in B and C are well correlated ($r=0.7$), showing the close correspondence between the FM velocities in their communication calls and the FM velocities to which IC neurons are tuned. The STRFs of all cells shown yielded good predictions for responses to communication calls (Adapted from Andoni et al. (2007))

has cells selective for both upward and downward FMs, but the majority of cells are selective for the downward direction (Figure 8.7a) (Razak and Fuzessery 2006; Andoni et al. 2007). Moreover, the range of preferences for sweep velocities corresponds closely to the sweep velocities in the signals these animals emit (Figure 8.7b, c) (Andoni et al. 2007). Thus the structure of their excitatory and inhibitory fields biases many IC neurons for downward direction selectivity and shapes their responsiveness to the FM velocities and other features present in their vocalizations.

8.4.2 Predictive Spectrotemporal Receptive Fields Found in Minority of IC Neurons

The STRFs of cells in which the convolutions accurately predicted responses and explained response selectivities present a comprehensive view of the quantitative features of excitation and inhibition in both frequency and time. The cells that yielded predictive STRFs must have linearly added the response of inhibitory and excitatory frequencies evoked by the rippled stimuli. Since STRFs reflect the average signal generated by such linear additions, the average representation of the excitatory and inhibitory fields generated by ripple stimuli was appropriate for predicting responses to other complex stimuli, such as the communication calls.

The neuronal population in the IC, however, is heterogeneous (Li et al. 2010; Pollak 2012) and most IC cells did not behave in the relatively simple way that the IC cells described above did. Specifically, predictive STRFs were only found in

only 25 % of IC cells; the STRFs in most cells (~75 %) provided poor predictions or were non-interpretable (Andoni et al. 2007). Those cells apparently had either static or dynamic non-linear response properties that were stronger than the linear response properties extracted by the STRFs generated by ripples. Stated differently, there was no linear relationship between the magnitudes of the excitation and inhibition in time and frequency that would apply to every complex signal. Therefore, the STRFs computed for those cells could not predict the response to a new complex signal such as the conspecific calls. The reason for the lack of predictability is that the nonlinear interactions of excitation and inhibition would be different than the average STRF derived from the linearly summed ripples.

8.4.3 Most Neurons Had More Than One Spectrotemporal Filter

Neurons in which the STRF (generated by spike-triggered averaging) yielded poor predictions for calls had multiple spectrotemporal features of the stimulus that defined the neuron's overall receptive field (Andoni and Pollak 2011). In these neurons, the nonlinear combination of multiple spectrotemporal features, or filters, predicted the neuron's spiking responses. To evaluate how multiple filters influenced the responses to communication calls, a computation was used that was a spike-triggered covariance procedure somewhat similar to principal component analysis (Rust et al. 2005). This method yielded two or three relevant filters in most of the IC cells, where the first spectrotemporal filter captured the most information of the stimulus-response relationship of each neuron.

In a study by Andoni and Pollak (2011), the set of relevant spectrotemporal filters was not computed from rippled stimuli but rather was computed from the responses evoked by a large number of conspecific communication calls. Natural calls were used because previous studies showed that STRFs derived from natural stimuli in both the IC of songbirds (Woolley et al. 2005; Gill et al. 2006) and in the cortex of ferrets (David et al. 2009) are significantly different than the STRFs derived with synthetic stimuli, such as ripples or noise. Most importantly, the receptive fields derived with natural stimuli provided far better predictions of responses to natural calls than did the receptive fields derived with synthetic stimuli (Woolley et al. 2005).

Predicted responses for both electronically generated FMs and conspecific calls were then calculated using either the first most informative spectrotemporal filter alone or the two most informative spectrotemporal filters (Figure 8.8). The most significant finding was that the predicted responses were poor when only the first filter was used but improved significantly when two filters were used (Andoni and Pollak 2011). The correlation coefficient between the predicted and the evoked responses for calls had a mean of 0.46 with only one filter but increased to a mean of 0.61 when two filters were used. This showed that these neurons did indeed have two or more spectrotemporal filters that determined the responses to calls.

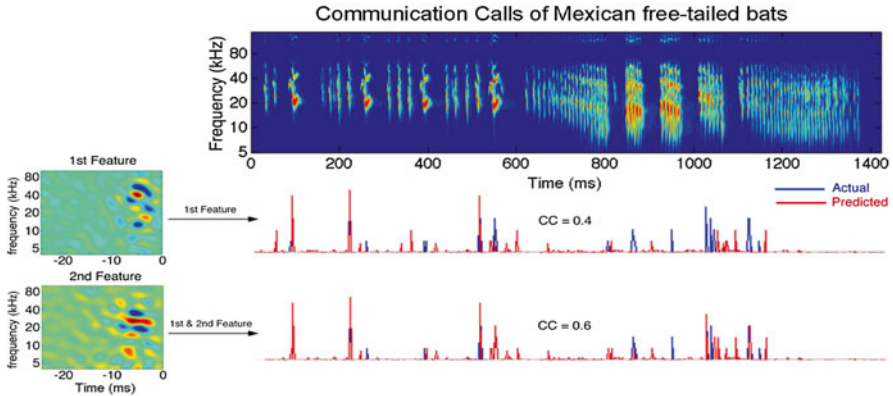


Fig. 8.8 Convolution with only first filter or feature and with both first and second features. Responses predicted from spectrotemporal features improve when multiple stimulus features are considered. The two most informative features of an IC neuron are shown. Convoluting the calls shown with only the first feature yielded poor predictions, with an average correlation coefficient between the predicted responses and those evoked by the calls of only 0.4. When both the first and second features were used to calculate the predicted responses, the correlation coefficient increased to 0.6. This shows that this IC neuron is tuned for multiple spectrotemporal features of natural calls (Adapted from Andoni and Pollak 2011)

The relevance of the two filters was further supported by the near-perfect agreement between the responses evoked by electronically generated FMs and the responses predicted with the non-linear combination of the two most relevant filters.

8.5 The Importance of Frequency Modulations for Call Selectivity

As was shown above, an important acoustic feature in the calls of bats is the structure of the FM components (Andoni et al. 2007; Andoni and Pollak 2011). FMs are not only prominent in both the echolocation and social communication calls emitted by bats (Bohn et al. 2008, 2009), FMs are also important components of communication signals in most animals, including humans (Doupe and Kuhl 1999). The structure of the FM component is important because IC cells are tuned for FM features and the velocity of the sweep and its direction, whether it sweeps upward or downward (Andoni et al. 2007; Fuzessery et al. 2011; Gittelman et al. 2012). Indeed, the specific tuning of IC cells for FM features is one of the principal features that determines the response selectivity for calls in the IC of bats (Andoni and Pollak 2011).

8.6 Directional Preferences for FMs Measured with In-Vivo Whole Cell Recordings

The previous sections emphasized the roles of inhibition in shaping directional preferences for FM sweeps. However, inhibition cannot be measured directly with extracellular recordings because extracellular electrodes only record spikes, which are evoked by excitation. With extracellular recordings, inhibition has to be inferred from the suppressive effects of some stimulus manipulation on the excitation evoked by another signal. Furthermore, with extracellular recordings there is an uncertainty about whether the observed spike suppression was due to inhibition at the IC or whether suppression was inherited from the inhibition that occurred in a lower nucleus that projects to the IC. To obtain a more direct and detailed view of sound-evoked inhibition, and to evaluate how the temporal features of inhibition interact with excitation to shape responses to complex signals, intracellular recordings were obtained with patch electrodes from the IC in awake bats in response to FM sweeps.

With intracellular recordings using patch electrodes, as with extracellular recordings, the discharges evoked in most IC cells exhibited a preference for downward sweeping FMs (Gittelman et al. 2009; Gittelman and Pollak 2011). With patch recordings, however, both the inputs to the cells (expressed in the amplitudes of post-synaptic potentials, PSPs) and their outputs, their discharges, are obtained. The selectivity differences of the inputs can be quantified by computing a PSP directional index (PSP amplitude evoked by the downward FM minus PSP amplitude evoked by the upward FM divided by the sum of the two amplitudes). Similarly, the selectivity differences of the outputs (discharges) are quantified by computing a discharge directional index (DSI) based on spike counts rather than PSP amplitudes. Thus the directional preferences of the inputs can be quantitatively compared to the directional preferences of the outputs.

In most IC cells, the differences in the discharge vigor evoked by the two FMs are substantially greater than the differences in the magnitudes of the excitatory postsynaptic potentials (EPSPs) evoked by the same signals (Gittelman et al. 2009). The discharge output of cell 2 in Figure 8.9, for example, was selective for the preferred (downward) FM; it fired to every presentation of the preferred FM but only fired 50 % of the time to the null (upward) FM and had a spike directional index of 0.33. In marked contrast, the EPSP amplitudes evoked by the two signals were very similar in amplitude. The PSP directional index was only 0.06. The disparity in the higher spike selectivity compared to the low EPSP selectivity is due to the nonlinear influence of spike threshold, where the larger EPSP evoked by the preferred FM in this cell was above threshold and evoked a discharge on every presentation, whereas the EPSP evoked by the null FM was only slightly smaller and just at threshold level. This cell illustrates the general finding that the inputs (PSPs) were less selective than the outputs (spikes). On average, the spike DSI was more than twice as large as the PSP-DSI among the IC population (Gittelman et al. 2009).

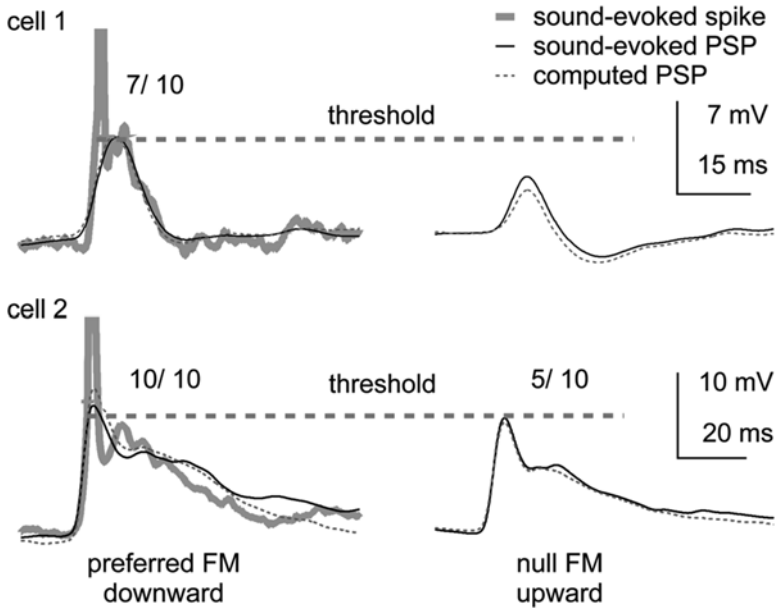


Fig. 8.9 Two directionally selective cells. Black traces are the measured postsynaptic potentials (PSPs) (mean of ten trials; spikes were removed by filtering), dashed traces are PSPs computed from derived conductance waveforms, and gray traces illustrate spiking with a single sweep response. Dashed line is spike threshold (Adapted from Gittelman and Pollak 2011)

8.7 The Role of Spike Timing for Creating Directional Selectivity

While the intracellular recordings with patch electrodes, like the recordings with extracellular electrodes, showed that most IC cells express directional preferences for FM sweeps, the comparison of EPSPs and spikes did not show by itself how the interactions of excitation and inhibition shaped the directional preferences of the cells. Previously, it was shown that for cells whose responses to calls could be predicted by their STRFs, the directional preferences were formed by the relative timing of the excitatory compared to the inhibitory inputs evoked by an FM sweep. The wide acceptance of this explanation is based on two principal observations. The first is that neurons selective for downward (or upward) FM sweeps have inhibitory fields that are lower (or higher) in frequency than the frequencies that activate their excitatory fields. These features were confirmed by the excitatory and inhibitory response fields in the linear STRFs, as illustrated by the cell in Figure 8.5, and these features were shown in a large number of previous studies in a variety of mammals (Razak and Fuzessery 2006; Kuo and Wu 2012). The second observation is that blocking inhibition reduces or eliminates directional preferences (Koch and Grothe 1998; Razak and Fuzessery 2009).

The timing hypothesis, which follows from the results of those experiments, posits that downward FM signals first sweep through the excitatory field, thereby evoking

an initial excitation, and slightly later in time, the signal sweeps through the inhibitory field (Covey and Casseday 1999; Zhang et al. 2003). With upward sweeping FMs, on the other hand, inhibition is activated first and the initial inhibition quenches the subsequent excitation. This is exactly the result obtained from the STRF of the IC neuron shown in Figures 8.5 and 8.6. The same arguments apply for upward preferring cells, but the frequencies of the excitatory and inhibitory fields are reversed.

But there is an additional implicit assumption in this explanation. Specifically, the explanation assumes that the inputs behave in a linear manner, where the excitatory and inhibitory inputs are evoked in synchrony with the spectrotemporal features of the signals. Thus, the same excitation and inhibition are evoked by downward and upward sweeping FMs, but the timing of excitation and inhibition is reversed because the temporal features of the signals are reversed. Moreover, it assumes that the cells are sensitive to even small differences in the arrival times of excitation and inhibition. These assumptions were used to explain the directional selectivity based on the STRF shown in Figure 8.5 and for all the other IC neurons that had linear STRFs. The strengths and relative timings of excitation and inhibition should simply be reversed as the FM direction is changed from upward to downward. In short, there is strong evidence from a variety of different studies in a variety of mammals to support the hypothesis that spectral arrangement of the excitatory and inhibitory fields generates FM directionality and that hypothesis also explains why blocking inhibition eliminates directionality.

8.7.1 FM Directional Selectivity Formed by Timing Disparities of Excitation and Inhibition Does Not Apply to All IC Cells

The IC is heterogeneous, as was pointed out previously, and a particular response property is formed in different ways among its neuronal population (Li et al. 2010; Pollak et al. 2011a, b). With regard to the formation of FM directional preferences, sensitivity for small differences in the timing of excitation and inhibition requires that such cells have features suitable for fast temporal processing, such as low input resistances and fast time constants. Such features would generate brief sound-evoked responses, where even small changes in the arrival of excitation and inhibition would change the amplitude of the sound-evoked response and hence the evoked firing rate. In cells with high input resistances and long time constants, in contrast, small changes in timing would be ineffective. The sound-evoked membrane potentials in these cells would change far more slowly with the arrival of each input, and these cells would be suited to integrate inputs over longer time periods. A recent study of IC cells in bats showed that about half of the cells in the IC do indeed have low input resistances that range from 40 to 100 megohms and fast time constants (Xie et al. 2008). Presumably these are the cells whose FM preferences are formed by the relative timing of excitation and inhibition. The other side of the finding is that about half of the IC population has high input resistances and long time constants, features that are inappropriate for sensitivity to small changes in the

timing of excitation and inhibition. It may well be that the non-linear cells with multiple filters had high input resistances and slow time constants, although there is no direct proof of this correspondence.

8.7.2 The Timing of Excitation and Inhibition Explored with Whole Cell Recordings

To evaluate the role of the timing of excitation and inhibition in IC cells with high input resistances, the excitatory and inhibitory conductances that generated the responses evoked by an upward and by a downward FM sweep were computed in a subset of IC neurons (Gittelman et al. 2009; Gittelman and Pollak 2011). Conductances reflect the number of ligand-gated receptors that are opened during excitation and inhibition. Thus, the excitatory conductance indicates the number of excitatory receptors opened in the IC cell by the release of glutamate from the pre-synaptic axons, whereas the inhibitory conductance reflects the number of inhibitory receptors opened in the IC cell by the release of GABA and/or glycine. Additionally, the latency and shape of the excitatory conductance waveform indicates the latency of excitatory innervation, the length of time the receptors are open, and the strength or magnitude of the excitation. The same applies to the inhibitory conductance waveform. The response of the membrane potential, the postsynaptic potential (PSP), is generated by the flow of currents through the conductive receptors. Those experiments showed several important features of the conductances evoked by the preferred and null FMs, as well as several other features of the EPSPs that would occur when the timing of inhibition is advanced or delayed (Gittelman and Pollak 2011).

The first important feature is that the excitatory conductances evoked by both the preferred and null FMs by themselves would have evoked a suprathreshold response. This finding is consistent with the general finding from extracellular studies that blocking inhibition, which simply eliminated inhibition and left only excitation, reduces or eliminates the directional preferences in almost all IC cells, allowing the cells to fire to both FMs. The second finding is that the amplitudes of the excitatory and inhibitory conductances evoked by the preferred FM and null FMs are almost always different. In other words, even though the spectral composition of the preferred and null FMs are identical but reversed in time, each signal does not evoke the same-but-time-reversed excitatory and inhibitory conductance waveforms. Rather, the excitatory and inhibitory conductance waveforms evoked by the preferred FM differ in either waveform shape or amplitude, or both shape and amplitude, from the conductances evoked by the null FM. The third finding is that there was no consistent relationship between the timing of the excitatory and inhibitory conductances evoked by the preferred FM compared to the null FM.

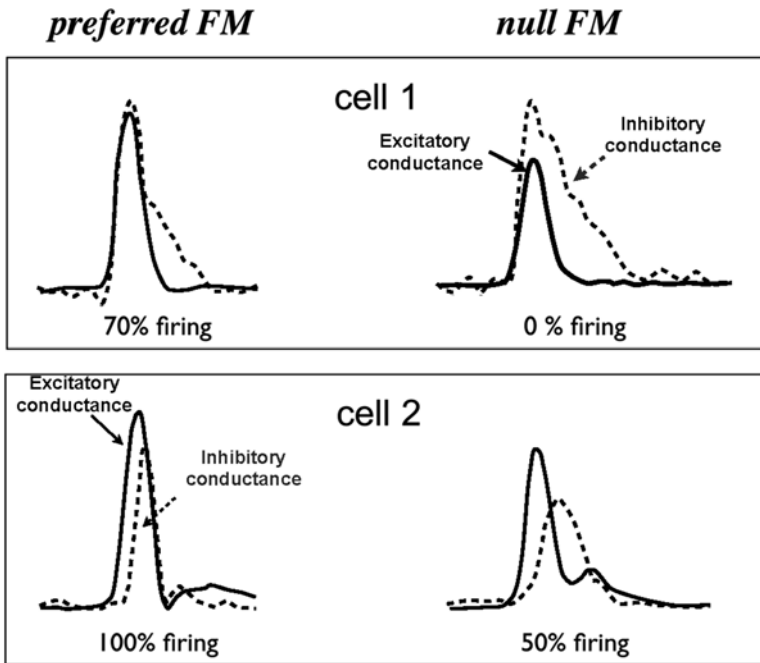


Fig. 8.10 Timing of excitation and inhibition provides no information about directional preferences. The calculated excitatory (*solid lines*) and inhibitory (*dashed lines*) conductances for a downward (*preferred*) and upward (*null*) FM sweep in two IC neurons. These are the same cells whose responses to the FMs are shown in Figure 8.9. In cell 1 (*top panel*), the timing of the excitatory and inhibitory conductances was virtually simultaneous for both the preferred and null FMs. In cell 2 (*bottom panel*), the excitatory conductance led the inhibitory conductance for both the preferred and the null FMs. However, the lead time of excitation was even greater for the null than the preferred. In both cells, the differences in the responses to the two FMs are due largely to the differences in the shapes and amplitudes of the excitatory compared to the inhibitory conductances rather than to their relative timing (Adapted from Gittelmann and Pollak 2011)

The two cells in Figures 8.9 and 8.10 illustrate two of the three features. Although not shown in Figure 8.9, the EPSPs of the preferred and null FMs, computed only from the excitatory conductances shown in Figure 8.10, were above threshold for both cells. In addition, the waveforms of the excitatory and inhibitory conductances of the preferred and null FMs in Figure 8.10 differed in shape and in peak amplitude (the exception is the peak amplitudes of the inhibitory conductances for cell 1, which were about the same). Finally, excitation and inhibition in cell 1 were virtually coincident for both the preferred and null FMs. In cell 2, in contrast, excitation led inhibition in the response to the preferred FM, but excitation led by an even greater amount of time in the response

to the null FM. For these cells at least, the notion that the preferred FM evokes an excitation that leads inhibition, whereas the null FM always evokes an inhibition that either leads or is coincident with the excitation, is not supported by the relative timing of the excitatory and inhibitory conductances. Since the relative timings of the excitatory and inhibitory conductances evoked by the preferred and null FMs differed from cell to cell, the relative timing of excitation and inhibition by itself provides little or no information about the preferences of these cells for the direction of an FM sweep. While the timing of inhibition and excitation contributed to the FM directional preference, in that inhibition and excitation had to overlap in time, the principal features that shaped directional selectivity in these cells were the magnitudes of inhibition compared to excitation coupled with the shapes of the conductance waveforms (Gittelman and Pollak 2011).

8.8 Combination Sensitivity

The results of the studies on FM directional selectivity illustrate that selective response properties, FM sweep directionality in this case, are formed in a variety of ways in the IC. Moreover, since FMs occur prominently in both echolocation and communication calls, the various ways in which FM directionality is formed applies to both echolocation and communication signals.

The idea of shared mechanisms for processing acoustic features of echolocation and communication calls is further illustrated by combination sensitive neurons in the auditory system of mustache bats (*Pteronotus parnellii*). Mustache bats emit biosonar signals composed of an initial long constant-frequency (CF) component and a terminal brief, downward sweeping FM component (Figure 8.11) (Pollak and Casseday 1986; Suga et al. 1987; Wenstrup and Grose 1995). The duration of the CF component can be up to 30 ms in duration, whereas the terminal FM is only 2–4 ms. Each call is emitted with a fundamental frequency and three harmonics, but the second harmonic always contains the most energy. Thus, the first CF harmonic, or fundamental, is emitted at about 30 kHz, and the terminal FM component sweeps downward by about 7 kHz, from 30 kHz to about 23 kHz. The CF of the second harmonic, the dominant harmonic, is emitted at about 60 kHz, and the second harmonic FM sweeps downward from 60 kHz to about 45 kHz. The CF of the third harmonic is emitted at about 90 kHz and the CF of the fourth at about 120 kHz, each with initial and terminal FM components as shown in Figure 8.11.

Suga was the first to discover combination sensitive neurons in the mustache bat's cortex (O'Neill and Suga 1979; Suga and O'Neill 1979). Combination sensitive neurons are distinguished by their far more vigorous responsiveness to two tones that have a specific-frequency relationship and temporal relationship, as compared to their responsiveness to a single tone at their BF. Hence the name "combination sensitive neurons."

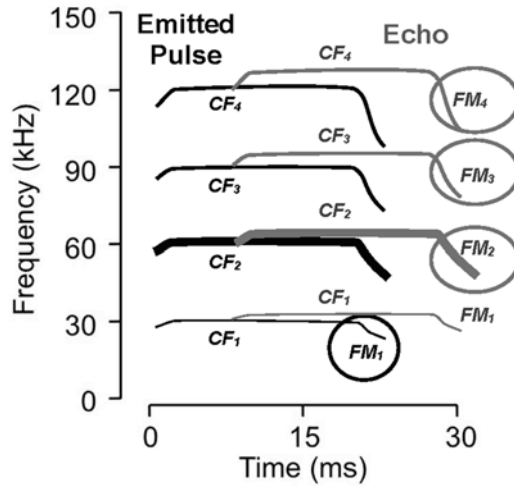


Fig. 8.11 Schematic sonogram of the sonar signals of mustache bats with the emitted pulse (*black lines*) and returning echo (*gray lines*). Line thickness indicates the relative strength of each signal harmonic. FM-FM neurons express response facilitation when presented with a frequency in the FM of the fundamental (FM1, *black circle*) followed in time by a frequency in the FM of either the second, third, or fourth harmonics (FM2, 3, 4, *gray ovals*) (Adapted from Portfors and Wenstrup 1999)

There are several types of combination sensitive neurons in the mustache bat's cortex (Suga et al. 1983). The most thoroughly studied are the so-called FM-FM neurons, which are topographically segregated from other types and are localized in the FM-FM region that is just adjacent to the primary auditory cortex. These neurons respond best to a combination of the first-harmonic, frequency-modulated (FM1) sweep in the emitted pulse and the frequency-modulated component of a higher harmonic (FM2, FM3, or FM4) in the returning echoes (Figure 8.12). In other words, FM-FM neurons are facilitated by the FM1 component in the simulated pulse and a higher harmonic FM component in the simulated echoes but only when there is a specific delay between the two signals.

One of the functional attributes assigned to FM-FM neurons is the coding of range or distance between the bat and its target (O'Neill and Suga 1979; Portfors and Wenstrup 1999). Target range is conveyed by the time interval between the FM of the emitted pulse and the FM of the returning echo (Simmons et al. 1974). The rationale follows from the finding that each FM-FM neuron has a best delay, the delay that produces the largest facilitation. The best delays of FM-FM neurons vary from about 1–20 ms, which correspond to the timing differences between the pulses and echoes that the bats receive during echolocation. There is then a striking concordance between the highly specified spectral and temporal requirements of the signals that drive these neurons optimally, and the spectral and temporal features of the biosonar signals that mustache bats emit and receive. The combinatorial proper-

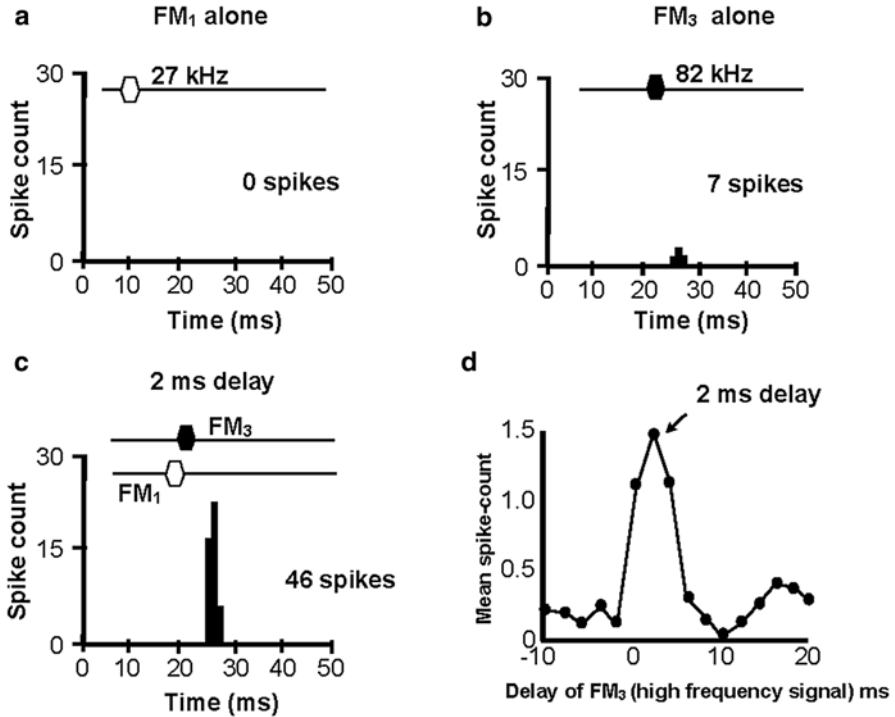


Fig. 8.12 Responses of a facilitated FM-FM neuron to single tone bursts and to the combination of a low- and high-frequency tone. (a) The neuron did not respond to a 27.0 kHz (FM₁) tone burst presented alone. (b) The neuron displayed a weak response to an 82.7 kHz (FM₃) tone (time of stimulus onset was 20 ms, and the response latency was 4 ms). (c) A facilitated response was elicited by delaying the onset of the high-frequency sound from the onset of the low-frequency sound by 2 ms. Response latency of the facilitated response was 7 ms from the onset of the low-frequency signal (5 ms from the onset of the high-frequency signal). (d) Delay tuning curve obtained by presenting two tone bursts at best facilitating frequencies (82.7 and 27.0 kHz), then varying the relative timing between the two tones. Zero milliseconds represents simultaneous presentation of the two signals. The best delay for this neuron was 2.0 ms (Adapted from Portfors and Wenstrup 1999)

ties of these neurons are tailored to the mustache bat's biosonar signals, and thus their particular features are unique to this animal.

This concordance led to the implicit hypothesis that the mustache bat's cortex is super-specialized for processing biosonar signals. It, therefore, came as a surprise when, in later studies, neurons in the FM-FM area were found to respond vigorously to a variety of signals used by mustache bats for social communication (Kanwal et al. 1994). Indeed, the neurons were selective in that they responded best to calls having particular spectral and temporal features that were similar to the spectral and temporal features of the FM-FM neurons, and they responded poorly or not at all to other signals. Thus, FM-FM neurons, as well as the other combination-sensitive types, express combinatorial properties that impart selectivity for both biosonar signals and for communication signals.

8.8.1 Combination Sensitive Neurons Are Created in the IC

Although combination-sensitive neurons were first discovered in the cortex and thus were thought to be an emergent or new response property resulting from cortical processing, subsequent studies by Wenstrup and his colleagues showed that combinatorial neurons are initially created in the mustache bat's IC (Mittmann and Wenstrup 1995; Portfors and Wenstrup 1999). Indeed, the combination-sensitive neurons in the mustache bat's IC have properties strikingly similar to those that Suga and his colleagues found in the cortex. In an elegant series of studies, Wenstrup and his colleagues worked out the mechanisms that generate combinatorial facilitation (Sanchez et al. 2007, 2008; Wenstrup et al. 2012). Remarkably, the mechanism is due entirely to inhibition rather than a summation of excitation evoked by the two frequencies. The role of inhibition in creating combinatorial facilitation once again underscores the importance of inhibition for generating selective response properties in the IC.

The evidence for the role of inhibition is that when the appropriate stimuli are presented and generate facilitation, blocking all excitatory receptors (both AMPA- and NMDA- type receptors) eliminates the response to a best frequency tone but has no effect at all on facilitation (Sanchez et al. 2008; Wenstrup et al. 2012). Indeed, even blocking GABAergic inhibition has no influence on facilitation. However, facilitation is eliminated completely when glycinergic inhibition is blocked by the iontophoretic application of strychnine. This led to the hypothesis that an initial low-frequency tone activates a glycinergic inhibitory input that produces an inhibition. The hypothesis proposed by Wenstrup et al. (2012) is shown in Figure 8.13. The key feature is that the inhibition evoked by the first low-frequency tone is followed by a post-inhibitory rebound that is excitatory but sub-threshold. The following higher frequency tone at the cell's BF also generates a glycinergic inhibition but one that is brief and has a fixed latency. Thus the BF also generates a short inhibition with a sub-threshold post-inhibitory rebound. The idea is that when the BF tone is delayed by an appropriate amount, the post-inhibitory rebound evoked by the first, low-frequency tone and the rebound evoked by the following BF tone coincide and summate to evoke a supra-threshold post-inhibitory rebound, which generates discharges and is the facilitated response. The facilitation is due entirely to glycinergic inhibition or to the rebounds from the summation of the two inhibitions. Thus, blocking AMPA or NMDA receptors has no influence on the facilitation, whereas blocking glycinergic inhibition completely eliminates it.

8.8.2 Combination Sensitivity also Imparts Selectivity for Communication Calls in the IC

One of the features that originally led to the idea that combination-sensitive neurons were specialized for echolocation was that the frequencies that evoked the facilitated responses correspond to the frequencies in the first and higher

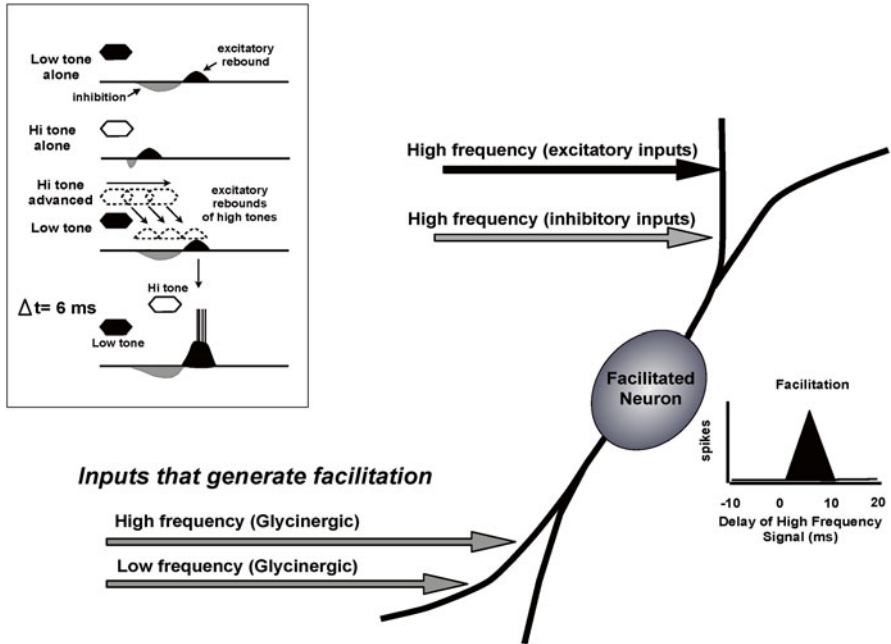


Fig. 8.13 Schematic diagram of mechanisms and circuitry underlying combination-sensitive facilitation in an IC neuron. Inset shows mechanism of post-inhibitory rebound. The high-frequency tone alone evokes a short duration inhibition of fixed latency followed by the post-inhibitory rebound. The low-frequency tone evokes a more prolonged inhibition with a longer latency followed by a rebound. When the high-frequency tone (representing the FM of the echo) follows the low-frequency tone (representing the FM of the emitted pulse) at an appropriate delay, the inhibitory rebounds evoked by both tones coincide in time and summate to generate the facilitation. The neuron receives a variety of high-frequency excitatory and inhibitory inputs tuned to its best frequency (*upper right*) that do not interact with the glycinergic inputs related to facilitation (*lower left*) (Adapted from Wenstrup et al. 2012)

harmonics of the mustache bat's echolocation calls. However, subsequent studies by Wenstrup, Portfors, and their colleagues (Portfors and Wenstrup 2002; Portfors 2004; Holmstrom et al. 2007) showed that many of the combination-sensitive neurons in the IC were tuned to non-echolocation frequencies, frequencies that were not contained in either the emitted calls or the echoes. Moreover, the best delays of many of the "non-echolocation" neurons were at or around 0 ms, intervals so short that they would never occur between the emitted pulse and echo.

These mismatches between the combinatorial properties of some IC neurons and the spectrotemporal features of their echolocation calls indicate that the non-echolocation neurons, as well as the other combinatorial neurons whose features correspond to the echolocation calls, may be important for encoding communication signals. Thus, Portfors (2004) suggested that since bats evolved from a non-echolocating ancestor and those ancestors most likely used acoustic signals for

communication, it seems reasonable to suppose that similar neural mechanisms underlie the processing of echolocation and communication calls in the auditory midbrain. To be more specific, it seems reasonable to suppose that combinatorial neurons could also generate response selectivity for communication calls.

In subsequent studies, Portfors and her colleagues showed that combination sensitivity imparts response selectivity for communication calls in the IC, allowing the cells to respond only to some calls but not to others (Portfors 2004). Indeed, the role of combination sensitivity for the processing of communication calls in the IC cells is similar to the role combination sensitivity plays in creating response selectivity for communication calls in the cortex.

8.8.3 Combination Sensitivity also Occurs in the Auditory Systems of Other Animals

Combination-sensitive neurons have been intensively studied in mustache bats, but whether they also occur in other bats, or the degree to which they occur in other bats, is unclear, mainly because it has not been studied in other species. It is significant, however, that combinatorial neurons also have been found in the IC of mice (Portfors and Felix 2005) and songbirds (Schneider and Woolley 2011), suggesting that such neurons are likely to occur in the IC of other bats as well. It should be noted, however, that there are fewer combination-sensitive neurons in the IC of mice and birds than there are in the IC of mustache bats. Additionally, and importantly, in both mice and birds the combination-sensitive neurons have been shown to be important for creating response selectivity for communication calls, as was shown previously in mustache bats.

8.9 Summary and Concluding Thoughts

The results of the studies reviewed here illustrate at least three general features of processing in the IC. The first is the dominant role that inhibition plays in shaping the responses of IC neurons. The roles of inhibition are illustrated by the marked change in the response selectivity for communication calls when inhibition is blocked and by the prominent roles of sideband inhibition for shaping FM directionality. The second feature is the heterogeneity of mechanisms that shape the response properties. There is not a single mechanism that the IC employs to form a given response property, but rather there are multiple ways in which the same response property is formed among the IC population. The third feature is the close correspondence between neural tuning and acoustic properties of conspecific communication signals. In bats at least, this correspondence suggests that IC neurons are specifically encoding features of these signals through the neural computations that generate FM selectivity. Moreover, it is clear that the various selectivities expressed by IC neurons for communication

calls are a consequence of the multiple ways in which their selective responses for features of acoustic signals, such as the direction of FM sweeps, are created.

These findings provide further support for the idea presented in the beginning of this chapter: The auditory systems of bats are not distinguished by novel mechanisms but rather that common mechanisms and features are far more pronounced in their auditory systems than in other mammals. This is also well illustrated by several higher order features that were first reported in the IC of bats, FM directional selectivity (Suga 1965), combination sensitivity (Mittmann and Wenstrup 1995), and duration tuning (Casseday et al. 1994), features that were subsequently seen in the auditory systems of other mammals (Brand et al. 2000; Portfors and Felix 2005). While all of these features are seen in the auditory systems of other mammals, they occur in different proportions than occur in bats. There is, for example, a pronounced over-representation of directional selectivity for downward FMs in the auditory systems of bats (Andoni and Pollak 2011), whereas in rats, upward and downward FMs are about equally prevalent (Zhang et al. 2003; Kuo and Wu 2012). Similarly, combination sensitivity is seen in about half of the auditory neurons in the mustache bat's IC (Wenstrup et al. 2012), whereas combination sensitivity occurs in only about 15 % of neurons in the mouse IC (Portfors and Felix 2005).

In short, the adaptations in the brain stem auditory nuclei are primarily, although not exclusively, a matter of quantity, where a species expresses certain features that are shared by other species but to a greater degree or in a more pronounced form rather than expressing wholesale qualitative changes in the mode of processing. The principal mechanisms for processing acoustic information are conserved among mammals, and thus the processing of sound in the brain stem auditory nuclei in a bat that is listening passively, as far as I can tell, is accomplished in the same way as it is in other mammals.

If the brain stem auditory systems of bats are so similar to other mammals, why don't all mammals echolocate? Perhaps the answer is that they possess the capability for echolocation, as do humans, as exemplified by Daniel Kish and his students, but they have never learned to express that ability. The argument is that the processing of sound in an animal that is passively listening must be different from one engaged in the active process of echolocation. Indeed, there is evidence that shows that sound-evoked responses change markedly when a bat is passively listening to a sound compared to when it hears the same sound during echolocation (Schuller 1979; Rubsamen and Betz 1986).

Exactly why and how active echolocation changes the responses evoked by sound is unclear. It is significant, in this regard, that there are massive descending projections from the auditory cortex that not only innervate the IC but also extend to almost all lower nuclei (Weedman and Ryugo 1996; Winer et al. 1998, 2002), and descending projections can have a profound influence on acoustically evoked responses (Suga et al. 2002; Xiao and Suga 2002). In addition, there is strong innervation by serotonergic as well as other neuromodulatory systems along the entire auditory system (Kossel and Vater 1989; Motts and Schofield 2009; Hurley and Sullivan 2012). Those systems are almost surely activated during vocalizations and during the active listening associated with vocalizations, as occurs during echolocation.

These features do not explain echolocation. But whatever role these systems play that enable echolocation, they may well function in an analogous way in most mam-

mals. If so, the operation of these systems, together with the basic mechanisms of acoustic processing in the ascending auditory system, would provide the framework for echolocation, not only in bats and cetaceans, but in humans as well.

Acknowledgments I would like to express my sincerest appreciation to the National Institutes of Health for their generous support of my research over the past 40 years. I also thank Carl Resler for his expertise with computers and his invaluable technical support throughout my career. Finally, I thank my graduate students, postdoctoral fellows, and all of my colleagues who have been so supportive and have taught me so much. Supported by NIH Grant DC 0078.

References

- Andoni, S., & Pollak, G. D. (2011). Selectivity for spectral motion as a neural computation for encoding natural communication signals in bat inferior colliculus. *Journal of Neuroscience*, 31(46), 16529–16540.
- Andoni, S., Li, N., & Pollak, G. D. (2007). Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. *Journal of Neuroscience*, 27(18), 4882–4893.
- Balaban, E. (1988). Bird song syntax: Learned intraspecific variation is meaningful. *Proceedings of the National Academy of Sciences of the USA*, 85(10), 3657–3660.
- Behr, O., & von Helversen, O. (2004). Bat serenades—complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioral Ecology and Sociobiology*, 56, 106–115.
- Bohn, K. M., Schmidt-French, B., Ma, S. T., & Pollak, G. D. (2008). Syllable acoustics, temporal patterns, and call composition vary with behavioral context in Mexican free-tailed bats. *Journal of the Acoustical Society of America*, 124(3), 1838–1848.
- Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., & Pollak, G. D. (2009). Versatility and stereotypy of free-tailed bat songs. *PLoS ONE*, 4(8), e6746.
- Boughman, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1392), 227–233.
- Brand, A., Urban, R., & Grothe, B. (2000). Duration tuning in the mouse auditory midbrain. *Journal of Neurophysiology*, 84(4), 1790–1799.
- Casseday, J. H., Ehrlich, D., & Covey, E. (1994). Neural tuning for sound duration: Role of inhibitory mechanisms in the inferior colliculus. *Science*, 264, 847–850.
- Casseday, J. H., Covey, E., & Grothe, B. (1997). Neural selectivity and tuning for sinusoidal frequency modulations in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. *Journal of Neurophysiology*, 77(3), 1595–1605.
- Casseday, J. H., Fremouw, T., & Covey, E. (2002). The inferior colliculus: A hub for the central auditory system. In D. Oertel, A.N. Popper, & R. R. Fay (Eds.), *Integrative functions in the mammalian auditory pathway* (pp. 238–318). New York: Springer.
- Catchpole, C. K., & Slater, P. J. B. (1995). *Bird song: Biological themes and variations*. Cambridge: Cambridge University Press.
- Covey, E., & Casseday, J. H. (1999). Timing in the auditory system of the bat. *Annual Review of Physiology*, 61, 457–476.
- David, S.V., Mesgarani, N., Fritz, J.B., & Shamma, S.A. (2009) Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by natural stimuli. *Journal of Neuroscience*, 29, 3374–3386.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631.
- Fuzessery, Z. M., Razak, K. A., & Williams, A. J. (2011). Multiple mechanisms shape selectivity for FM sweep rate and direction in the pallid bat inferior colliculus and auditory cortex. *Journal*

- of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 197, 615–623.
- Gill, P., Zhang, J., Woolley, S. M., Fremouw, T., & Theunissen, F. E. (2006). Sound representation methods for spectro-temporal receptive field estimation. *Journal of Computational Neuroscience*, 21, 5–20.
- Gittelman, J. X., & Pollak, G. D. (2011). It's about time: How input timing is used and not used to create emergent properties in the auditory system. *Journal of Neuroscience*, 31, 2576–2583.
- Gittelman, J. X., Li, N., & Pollak, G. D. (2009). Mechanisms underlying directional selectivity for frequency modulated sweeps in the inferior colliculus revealed by in vivo whole-cell recordings. *Journal of Neuroscience*, 29, 13030–13041.
- Gittelman, J. X., Wang, L., Colburn, H. S., & Pollak, G. D. (2012). Inhibition shapes response selectivity in the inferior colliculus by gain modulation. *Frontiers of Neural Circuits*, 6, 67.
- Griffin, D. R. (1944). Echolocation by blind men and bats. *Science*, 100, 589–590.
- Griffin, D. R. (1958). *Listening in the dark*. Ithaca, NY: Cornell University Press.
- Holmstrom, L., Roberts, P. D., & Portfors, C. V. (2007). Responses to social vocalizations in the inferior colliculus of the mustached bat are influenced by secondary tuning curves. *Journal of Neurophysiology*, 98, 3461–3472.
- Hurley, L. M., & Sullivan, M. R. (2012). From behavioral context to receptors: Serotonergic modulatory pathways in the IC. *Frontiers of Neural Circuits*, doi: [10.3389/fncir.2012.00058](https://doi.org/10.3389/fncir.2012.00058)
- Kanwal, J. S. (1999). Processing species-specific calls: Combination-sensitive neurons in an echolocating bat. In M. D. Hauser & M. Konishi (Eds.), *The design of animal communication* (pp. 133–156). Cambridge, MA: The MIT Press.
- Kanwal, J. S., Matsumura, S., Ohlemiller, K., & Suga, N. (1994). Analysis of acoustic elements and syntax in communication sounds emitted by mustached bats. *Journal of the Acoustical Society of America*, 96, 1229–1254.
- Klein, D. J., Depireux, D. A., Simon, J. Z., & Shamma, S. A. (2000). Robust spectrotemporal reverse correlation for the auditory system: Optimizing stimulus design. *Journal of Computational Neuroscience*, 9, 85–111.
- Klug, A., Bauer, E. E., Hanson, J. T., Hurley, L., Meitzen, J., & Pollak, G. D. (2002). Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition. *Journal of Neurophysiology*, 88, 1941–1954.
- Knornschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat, *Saccopteryx bilineata*. *Naturwissenschaften*, 93, 451–454.
- Koch, U., & Grothe, B. (1998). GABAergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat. *Journal of Neurophysiology*, 80, 71–82.
- Kossel, M., & Vater, M. (1989). Noradrenaline enhances temporal auditory contrast and neuronal timing precision in the cochlear nucleus of the mustached bat. *Journal of Neuroscience*, 9, 4169–4178.
- Kowalski, N., Depireux, D. A., & Shamma, S. A. (1996). Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. *Journal of Neurophysiology*, 76, 3524–3534.
- Kuo, R. I., & Wu, G. K. (2012). Generation of direction selectivity in the auditory system. *Neuron*, 73, 1016–1027.
- Li, N., Gittelman, J. X., & Pollak, G. D. (2010). Intracellular recordings reveal novel features of neurons that code interaural intensity disparities in the inferior colliculus. *Journal of Neuroscience*, 30, 14573–14584.
- Marler, P. (2004). Science and birdsong: The good old days. In P. Marler & H. Slabbekoom (Eds.), *Nature's music: The science of birdsong*. Vol. 1. Amsterdam: Academic Press.
- Mayko, Z. M., Roberts, P. D., & Portfors, V. (2012). Inhibition shapes selectivity for vocalizations in the inferior colliculus of awake mice. *Frontiers of Neural Circuits*, doi: [10.3389/fncir.2012.00073](https://doi.org/10.3389/fncir.2012.00073)
- Mittmann, D. H., & Wenstrup, J. J. (1995). Combination-sensitive neurons in the inferior colliculus. *Hearing Research*, 90, 185–191.
- Motts, S. D., & Schofield, B. R. (2009). Sources of cholinergic input to the inferior colliculus. *Neuroscience*, 160, 103–114.

- Neuweiler, G. (1990). Auditory adaptations for prey capture in echolocating bats. *Physiological Reviews*, 70, 615–641.
- O’Neill, W. E., & Suga, N. (1979). Target range-sensitive neurons in the auditory cortex of the mustache bat. *Science*, 203, 69–73.
- Payne, R. S., & McVay, S. (1971). Songs of humpback whales. *Science*, 173, 585–597.
- Pollak, G. D. (2012). Circuits for processing dynamic interaural intensity disparities in the inferior colliculus. *Hearing Research*, 288, 47–57.
- Pollak, G. D. & Casseday, J. H. (1986) *The neural basis of echolocation in bats*. New York: Springer-Verlag.
- Pollak, G. D., Winer, J. A., & O’Neill, W. E. (1995). Perspectives on the functional organization of the mammalian auditory system: Why bats are good models. In A. N. Popper & R. R. Fay (Eds.), *Hearing by bats* (pp. 481–498). New York: Springer.
- Pollak, G. D., Gittelman, J. X., Li, N., & Xie, R. (2011a). Inhibitory projections from the ventral nucleus of the lateral lemniscus and superior paraolivary nucleus create directional selectivity of frequency modulations in the inferior colliculus: A comparison of bats with other mammals. *Hearing Research*, 273, 134–144.
- Pollak, G. D., Xie, R., Gittelman, J. X., Andoni, S., & Li, N. (2011b). The dominance of inhibition in the inferior colliculus. *Hearing Research*, 274, 27–39.
- Portfors, C. V. (2004). Combination sensitivity and processing of communication calls in the inferior colliculus of the mustached bat *Pteronotus parnellii*. *Anais da Academia Brasileira de Ciências*, 76, 253–257.
- Portfors, C. V., & Wenstrup, J. J. (1999). Delay-tuned neurons in the inferior colliculus of the mustached bat: Implications for analyses of target distance. *Journal of Neurophysiology*, 82, 1326–1338.
- Portfors, C. V., & Wenstrup, J. J. (2002). Excitatory and facilitatory frequency response areas in the inferior colliculus of the mustached bat. *Hearing Research*, 168, 131–138.
- Portfors, C. V., & Felix, II, R. A. (2005). Spectral integration in the inferior colliculus of the CBA/CaJ mouse. *Neuroscience*, 136, 1159–1170.
- Portfors, C. V., Roberts, P. D., & Jonson, K. (2009). Over-representation of species-specific vocalizations in the awake mouse inferior colliculus. *Neuroscience*, 162, 486–500.
- Razak, K. A., & Fuzessery, Z. M. (2006). Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex of the pallid bat. *Journal of Neurophysiology*, 96, 1303–1319.
- Razak, K. A., & Fuzessery, Z. M. (2009). GABA shapes selectivity for the rate and direction of frequency modulated sweeps in the auditory cortex. *Journal of Neurophysiology*, 102, 1366–1378.
- Rubsamen, R., & Betz, M. (1986). Control of echolocation pulses by neurons of the nucleus ambiguus in the rufous horseshoe bat, *Rhinolophus rouxi*. I. Single unit recordings in the ventral motor nucleus of the laryngeal nerves in spontaneously vocalizing bats. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 159, 675–687.
- Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, 46, 945–956.
- Sanchez, J. T., Gans, D., & Wenstrup, J. J. (2007). Contribution of NMDA and AMPA receptors to temporal patterning of auditory responses in the inferior colliculus. *Journal of Neuroscience*, 27, 1954–1963.
- Sanchez, J. T., Gans, D., & Wenstrup, J. J. (2008). Glycinergic “inhibition” mediates selective excitatory responses to combinations of sounds. *Journal of Neuroscience*, 28, 80–90.
- Schneider, D. M., & Woolley, S. M. (2011). Extra-classical tuning predicts stimulus-dependent receptive fields in auditory neurons. *Journal of Neuroscience*, 31, 11867–11878.
- Schuller, G. (1979). Coding of small sinusoidal frequency and amplitude modulations in the inferior colliculus of ‘CF-FM’ bat, *Rhinolophus ferrumequinum*. *Experimental Brain Research*, 34, 117–132.
- Schwartz, C., Tressler, J., Keller, H., Vanzant, M., Ezell, S., & Smotherman, M. (2007). The tiny difference between foraging and communication buzzes uttered by the Mexican free-tailed bat,

- Tadarida brasiliensis*. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 193, 853–863.
- Simmons, J. A., Lavender, W. A., Lavender, B. A., Doroshov, C. A., Kiefer, S. W., Livingston, R., et al. (1974). Target structure and echo spectral discrimination by echolocating bats. *Science*, 186, 1130–1132.
- Simmons, J. A., Howell, D. J., & Suga, N. (1975). Information content of bat sonar echoes. *American Scientist*, 63, 204–215.
- Suga, N. (1965) Analysis of frequency-modulated sounds by auditory neurons of echolocating bats. *Journal of Physiology*, 179, 26–53.
- Suga, N., & O'Neill, W.E. (1979) Neural axes representing target range in the auditory cortex of the mustached bat. *Science*, 206, 351–353.
- Suga, N., O'Neill, W. E., Kujirai, K., & Manabe, T. (1983). Specificity of combination-sensitive neurons for processing of complex biosonar signals in auditory cortex of the mustached bat. *Journal of Neurophysiology*, 49, 1573–1626.
- Suga, N., Niwa, H., Taniguchi, I., & Margoliash, D. (1987). The personalized auditory cortex of the mustached bat: Adaptation for echolocation. *Journal of Neurophysiology*, 58, 643–654.
- Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, 36, 9–18.
- Thaler, L., Arnott, S.R., & Goodale, M.A. (2011) Neural correlates of natural human echolocation in early and late blind echolocation experts. *PLoS ONE*, 6 (5) e20162.
- Weedman, D. L., & Ryugo, D. K. (1996). Projections from auditory cortex to the cochlear nucleus in rats: Synapses on granule cell dendrites. *Journal of Comparative Neurology*, 371, 311–324.
- Wenstrup, J. J., & Grose, C. D. (1995). Inputs to combination-sensitive neurons in the medial geniculate body of the mustached bat: the missing fundamental. *Journal of Neuroscience*, 15, 4693–4711.
- Wenstrup, J. J., Nataraj, K., & Sanchez, J. T. (2012). Mechanisms of spectral and temporal integration in the mustached bat inferior colliculus. *Frontiers of Neural Circuits*, doi:[10.3389/fncir.2012.00075](https://doi.org/10.3389/fncir.2012.00075)
- Winer, J. A., Larue, D. T., & Pollak, G. D. (1995). GABA and glycine in the central auditory system of the mustache bat: Structural substrates for inhibitory neuronal organization. *Journal of Comparative Neurology*, 355, 317–353.
- Winer, J. A., Larue, D. T., Diehl, J. J., & Hefti, B. J. (1998). Auditory cortical projections to the cat inferior colliculus. *Journal of Comparative Neurology*, 400, 147–174.
- Winer, J. A., Chernock, M. L., Larue, D. T., & Cheung, S. W. (2002). Descending projections to the inferior colliculus from the posterior thalamus and the auditory cortex in rat, cat, and monkey. *Hearing Research*, 168, 181–195.
- Woolley, S. M., Fremouw, T. E., Hsu, A., & Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience*, 8, 1371–1379.
- Xiao, Z., & Suga, N. (2002). Modulation of cochlear hair cells by the auditory cortex in the mustached bat. *Nature Neuroscience*, 5, 57–63.
- Xie, R., Meitzen, J., & Pollak, G. D. (2005). Differing roles of inhibition in hierarchical processing of species-specific calls in auditory brainstem nuclei. *Journal of Neurophysiology*, 94, 4019–4037.
- Xie, R., Gittelman, J. X., Li, N., & Pollak, G. D. (2008). Whole cell recordings of intrinsic properties and sound-evoked responses from the inferior colliculus. *Neuroscience*, 154, 245–256.
- Zhang, L. I., Tan, A. Y., Schreiner, C. E., & Merzenich, M. M. (2003). Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature*, 424, 201–205.