

Chapter 23

Auditory Neuroethology and Speech Processing: Complex-Sound Processing by Combination-Sensitive Neurons

NOBUO SUGA

ABSTRACT

In auditory neuroethology, which is based upon the philosophy that the vocal and auditory systems have evolved together for acoustic communication or echolocation, biologically important sounds have been used as stimuli to explore auditory mechanisms underlying species-specific behavior. Such studies on the auditory systems of different types of animals have been successful in finding the neural mechanisms that are species specific as well as the mechanisms that are shared by different types of animals. Therefore, in this chapter I explain the neuroethological view and offer my opinion that neuroethological studies on different types of animals, including "auditorily specialized" animals, will contribute most to the understanding of speech sound processing in humans. Speech sounds are much more complex than animal sounds, but both types of sounds contain similar basic components, that is, information-bearing elements. These elements and the relationships among them are characterized by particular values of parameters, that is, information-bearing parameters (IBPs). The central mechanisms shared by different types of animals are the production of neurons tuned to IBPs, complex-sound processing by combination-sensitive neurons, and anatomical parcellation for the representation of IBPs. Since the best neuroethological data on complex-sound processing were obtained from the mustached bat, this chapter presents the essential portions of these data and offers a hypothesis as to which basic neural mechanisms are probably involved in speech-sound processing.

Auditory neurophysiologists have found neurons and functional organizations specialized for complex-sound processing in different species of animals. Their findings are interesting and have intrinsic value but contribute little to an understanding of the central mechanisms of speech processing, because speech processing must involve neural mechanisms far more elaborate than those seen in animals. It was suggested that I write a chapter for this volume describing our neuroethological data on biosonar with special reference to possible implications for understanding speech processing. Herein, therefore, I express my optimistic view: The neural mechanisms found in the central auditory system of animals must also be involved in the most fundamental portions of speech processing. I first describe a neuroethological view of the auditory system and cast doubt

on the common view that studies on the auditory system of the macaque monkey contribute the most to an understanding of basic neural mechanisms for speech processing. I then discuss several possible types of neural representation of auditory information in auditory centers and present the neuroethological data that are probably relevant to the basic mechanisms for speech-sound processing. Finally, I offer a hypothesis as to which neural mechanisms contribute to speech-sound processing.

The coding of speech sounds by the auditory periphery in animals (mammals) is probably very similar to that in humans, so that studies on coding in animals are justifiable. On the other hand, the processing of speech sounds by the human auditory center must be unique, so that studies on the processing of speech sounds by the auditory center in animals are probably not justifiable. Auditory neuroethology, the aim of which is to understand the neural mechanism for species-specific auditory behavior, may contribute to the understanding of the basic mechanisms for the processing of speech sounds, a view I substantiate in this chapter.

NEUROETHOLOGICAL VIEW OF THE AUDITORY SYSTEM

Since the auditory system has evolved together with the vocal system under the influence of the acoustic environment, it has become specialized for receiving and processing the acoustic signals that are the most important for the survival of a species. The reception and processing of signals by the auditory system should thus be considered from an evolutionary perspective. To illustrate this view, let us consider an oversimplified hypothetical forest. Initially, suppose that the hypothetical forest has only species X, which uses only one type of call for pair formation. The call is a narrow-band noise burst (X in Figure 1A). To effectively detect this call, the auditory system of species X requires only one type of neuron, which shows a frequency tuning curve matching the amplitude spectrum of the call (Figure 1B). The excitation of this neuron is directly related to the detection of this call, so that this neuron may be considered a call detector used for communication. (We ignore the temporal pattern of the call and ambient noise such as sounds produced by wind. Therefore, discrimination pressure is very low. Discrimination pressure may be determined by the number and similarity of biologically significant sounds that must be discriminated among; it produces different perceptual categories that may be coded genetically and/or learned postnatally.)

When species X is taken from the hypothetical forest to a laboratory to examine responses of its auditory neuron to various types of acoustic stimuli, it is found that the neuron responds to any stimulus that contains energy falling into its excitatory area (the area above the excitatory tuning curve). This nonspecific response does not, however, disqualify the neuron as a call detector, because in this hypothetical forest there are no sounds other than the species-specific call. Therefore, it may be stated that the qualification of a neuron as a call detector should be examined in terms of its responses to biologically significant sounds, acoustic environment, and auditory behavior. For quantitative studies of the

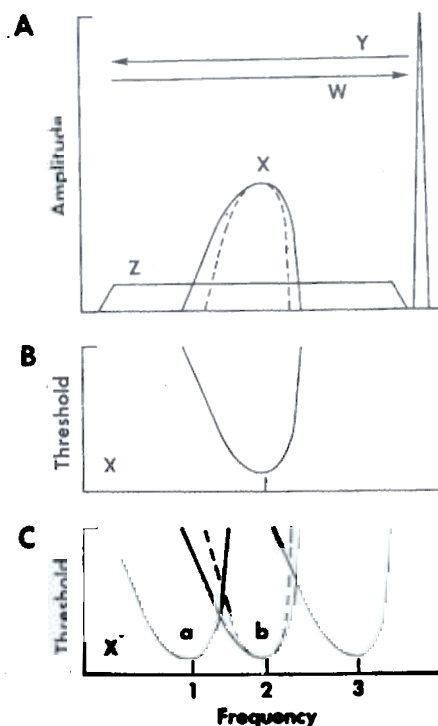


Figure 1. Properties of sounds (A) and frequency tuning (threshold) curves of peripheral auditory neurons (B and C) of animals in hypothetical forests. A: Amplitude spectra of sounds used by four different species, W, X, Y, and Z. X and Z emit narrow-band and broadband noise bursts, respectively, while W and Y produce FM sounds sweeping upward and downward as indicated by the arrows, respectively. B: Frequency tuning curve of a peripheral auditory neuron of species X in the first forest model. C: Tuning curves of three peripheral auditory neurons of species X in the second forest model. The dashed curves in A and C indicate the amplitude spectrum of a call and a tuning curve of species X with a high discrimination pressure, respectively. (From Suga, 1978.)

response properties of neurons, however, acoustic stimuli that are outside the range of variation of a biologically significant sound must also be delivered. That is, appropriate unnatural sounds should be used as stimuli.

Now consider the second hypothetical forest with three additional species, W, Y, and Z, which respectively produce an upward-sweeping frequency-modulated (FM) sound, a downward-sweeping FM signal, and a broadband noise burst (Figure 1A). The neuron of X can be excited by any of these sounds, so that now the excitation of this neuron does not necessarily mean a potential mate is calling. In this acoustic environment, this neuron is no longer a call detector. If W and Y are predators of X with different hunting strategies, and Z is the prey of X, X must be able to discriminate among all four types of sounds. Then its vocal and auditory systems should evolve to maximize the detection and discrimination of these.

Change in the Auditory and Vocal System

Animal sounds can be expressed within multidimensional continua. The higher the discrimination pressure, the more these sounds will be confined to specific, narrow ranges of the multidimensional continua. That is, the sounds will become more stereotyped and discrete. Some variation of the signals within the narrow ranges may be important for individual recognition. In our second model, properties of an acoustic signal of X may change to facilitate correct categorical perception, as shown by the dashed line in Figure 1A. In higher vertebrates, distributions of different types of acoustic signals may show significant overlap within any single parameter; that is, these may not be discrete but are probably unique in the particular combination of temporal and spectral parameters.

For discrimination among calls, X in the second model should obviously be equipped with an auditory system different from that described in the first model. That is, (1) the peripheral auditory system should have a frequency response curve broader than that matched to call X ; (2) it should have at least three peripheral neurons with different frequency tuning curves (Figure 1C); (3) it may have neurons with sharper tuning curves for the acoustic signal that is most frequently used by the animal and/or is most important for fitness (Figure 1C, dashed line); (4) it may have greater sensitivity for the more important signals; (5) it may have a higher population of sensory cells and neurons for processing the most important signals; and (6) the central auditory system may have "specialized" neurons that express the output of a neural circuit acting as a filter tuned to biologically important parameters, such as the direction and rate of frequency sweep and the center frequency and bandwidth of noise bursts (Suga, 1978). [These specializations of the auditory system are more easily discovered when the energy of a biologically important sound is concentrated at a particular narrow frequency band. This is the case in the mustached bat, *Pteronotus parnellii* (Suga, 1984; Suga and Tsuzuki, 1985).]

The selectivity of specialized neurons for an acoustic pattern should, of course, not change with stimulus level. This "level-tolerant" selectivity simply means that, for instance, neuron x , which selectively responds to call x , should not at any amplitude respond to calls other than call x . Its response magnitude may monotonically or nonmonotonically vary with the amplitude of call x , so that neuron x has a lower threshold and also may have an upper threshold. When neurons have an upper threshold in addition to a lower threshold, they are "tuned" to a particular amplitude, so that they are also specialized in the amplitude domain (Suga, 1977; Suga and Manabe, 1982; Suga et al., 1983a). Some possible neural circuits that act as filters for one of the four types of sounds in the second forest model are shown in Figure 2. Neuron x , which selectively responds to call x , can be produced by a lateral inhibition model (Figure 2A). Neurons that selectively respond to either sound y or sound w can be produced by a disinhibition model (Figure 2B). Neurons that selectively respond to sound z can be produced by a facilitation model acting as an AND gate. The neurons that express the outputs of the filters in Figure 2 have actually been found in bats (Suga, 1965a, b, 1968).

Whether the activity of specialized neurons can be directly correlated with categorical perception or individual recognition of a signal depends on whether

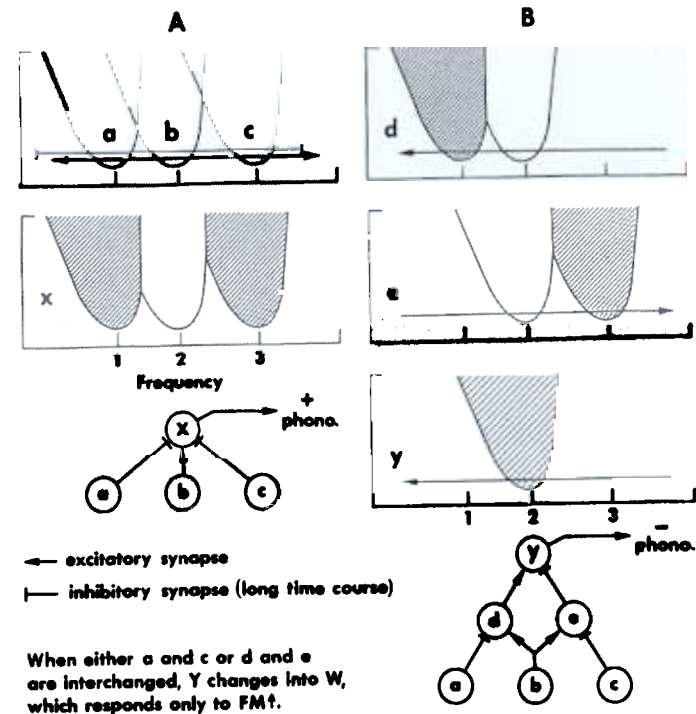


Figure 2. Neural network models, each of which acts as a filter to pass one of the four types of acoustic signals shown in Figure 1A. Neurons x and y express the output of the filter that passes only either sound x or y . A, upper graph: Tuning curves of three peripheral neurons (a , b , and c) and their thresholds for FM sounds (arrows) and the broadband noise (bar). A, lower graph: Excitatory (unshaded) and inhibitory (shaded) areas of neuron x , which are the areas above the excitatory and inhibitory tuning curves, respectively. Neuron x responds only to the narrow-band noise burst produced by species X because of the neural network mediating lateral inhibition, as shown at the bottom. B, top and middle graphs: Excitatory and inhibitory areas of asymmetrical neurons d and e , which respond asymmetrically to FM sounds because of the neural networks shown at the bottom. B, bottom graph: Inhibitory area of neuron y and its threshold for the downward-sweeping FM sound (arrow). Neuron y responds only to the downward-sweeping FM sound produced by species Y because of the neural network mediating disinhibition. (From Suga, 1978, based on Suga, 1973.)

the bandwidth of their neural filters is the same as that of a behavioral filter for categorical perception or individual recognition.

The statement that neuron x does not respond to sounds w , y , and z but only to call x is valid only when the statement is based on the data demonstrating that neuron x is "tuned" to a parameter or parameters characterizing call x . An additional necessary description of the properties of neuron x is, "sounds w , y , and z inhibit the activity of neuron x , but sound x excites it" or "potential responses of neuron x to sounds w , y , and z are 'rejected' by a neural network, but the response of neuron x to sound x is not," because all of the peripheral neurons a , b , and c in Figure 1C respond to all four types of sounds. To qualify

as a specialized neuron, a neuron must be demonstrated to have a rejection mode as a result of either inhibition or facilitation (AND gate) or both. The two neural network models in Figure 2 both have a rejection mode. This does not mean, however, that a neuron with a rejection mode is always a specialized neuron.

In lower vertebrates that are subject to very low discrimination pressure, the rejection mode to minimize responses to biologically irrelevant sounds may be found in the peripheral auditory system (Frishkopf et al., 1968; Capranica, 1977). The stronger rejection mode at the periphery is directly related to the more limited number of perceptual categories. In vertebrates, each species usually has several types of communication sounds, in which the amplitude spectrum may vary in time. The animals may also be concerned with sounds produced by other species. Discrimination pressure is thus high. Accordingly, the peripheral auditory system should be broadly tuned, without a rejection mode, and should have a large filter bank (a large array of neurons with different frequency tuning curves). The rejection mode should be incorporated in the central auditory system. This is certainly true in mammals.

Remarks. The philosophy involved in auditory neuroethology is that the vocal and auditory systems have evolved together for acoustic communication. In other words, the vocal system has adapted to produce sounds suitable for detection and processing by the auditory system, and the auditory system has evolved to detect and process these sounds. In auditory neuroethology, therefore, biologically important sounds have been used as stimuli, and the auditory system has been studied in relation to auditory behavior. This does not mean that the use of unnatural sounds, including pure tones, should be excluded. On the contrary, it is necessary to use unnatural sounds as well as natural sounds for quantitative studies, because the parameters characterizing the biologically important sounds must be varied beyond their normal range in order to study the filter properties of neurons. The statement that a neuron selectively responds to a particular type of call cannot be justified without appropriate studies of excitatory, inhibitory, and facilitative responses of the neuron to tone bursts and other stimuli. Neural mechanisms underlying the response properties of specialized neurons cannot be explored properly without tone burst stimuli.

Auditory Periphery and Center

In many species of mammals, frequency analysis and coding of complex sounds by the auditory periphery appear to be performed by nearly identical mechanisms. Therefore, insight into the analysis and coding of speech sounds by the auditory periphery of humans can be gained by electrophysiological studies on the auditory periphery of animals, in particular those showing an audiogram similar to that of humans. There is neither anatomical nor physiological reason to believe that such studies must be made with macaque monkeys. The rationale for using monkeys has only a taxonomic (phylogenetic) reason: Both humans and monkeys are primates.

The auditory periphery in certain species of bats shows anatomical and/or physiological specializations for fine frequency analysis to fulfill a species-specific

need for a velocity-sensitive echolocation system (Pollack et al., 1972; Henson, 1973; Suga et al., 1975, 1976; Bruns, 1976; Schnitzler et al., 1976; Suga and Jen, 1977). Such specialization is not shared by other species of bats from the same suborder, microchiroptera, who have a velocity-insensitive echolocation system. Therefore, the possibility that the human auditory periphery is somewhat different in stimulus coding from that of nonhuman primates cannot be denied completely.

It has been thought that the most valuable data related to the central neural mechanisms of speech processing will be obtained from physiological studies on the auditory cortex of the macaque monkey. Recent progress in auditory neuroethology, however, has raised doubts about this view. Interspecies anatomical and physiological correlations are highest between human and monkey brains. The functional organization of the cortical visual areas of monkeys is probably similar to that of those areas of humans (except the area processing written language), because monkeys and humans perceive similar visual patterns. If monkey acoustic communication were similar to human speech, the physiological correlations between monkey and human auditory centers would also be high. This is, however, not the case. Human speech is unique and highly specialized. There is an enormous difference in both quantity and quality between human speech and monkey acoustic communication, although both species belong to the same taxonomic order. Therefore, studying the monkey's auditory cortex to understand the functional organization of the human speech areas is not appropriate. Such studies would be similar to those done on monkey auditory cortex in order to understand the functional organization of the mustached bat auditory cortex. These are obviously inappropriate even if the parameters of biosonar signals are changed to stimulate the monkey's ear, because the mustached bat has an auditory cortex that is highly specialized for processing complex biosonar signals, that is, the signals for communicating with its specific environment.

Since the central auditory system, and in particular the auditory cortex, must be evolved for the processing of biologically important signals, usually species-specific sounds, we must first study how species-specific sounds are processed in the auditory system, and how the auditory cortex is functionally organized apart from tonotopic representation. To explore the functional organization of the speech areas, we obviously need a technique that allows us to examine the activity of a tiny cluster of neurons in the speech areas without inserting any foreign material into the brain. However, such a technique is not yet available. Until it is, there is some question as to the extent to which the auditory physiology of animals can contribute to the exploration of neural mechanisms of speech processing. It has been pointed out that the classical auditory physiology of neural responses to pure tones and clicks has contributed little to understanding central auditory mechanisms other than those for sound localization, and that the central auditory system should be studied through a neuroethological approach, using biologically important sounds. As a matter of fact, auditory neuroethology has been successful in the exploration of the basic mechanisms of complex-sound processing that are shared by different types of animals, as well as in the exploration of species-specific mechanisms. But neuroethology of the central auditory system of animals cannot explore the neural mechanisms of speech processing, in particular the functional organization of the speech area. However,

it can present the data that should be considered when formulating hypotheses concerning the neural mechanisms of speech processing and can clarify what kind of organization should be looked for when a new technique for studying the speech areas becomes available.

The functional organization of the auditory cortex does not necessarily follow a similarity in gross anatomy and phylogeny; rather, it follows the properties of biologically important sounds. For example, within the suborder microchiroptera, the auditory system can be functionally different even between two species belonging to the same genus when their acoustic signals differ (Grinnell, 1970; Suga et al., 1974). But it can be similar even between two species belonging to different families when their acoustic signals are similar (Suga and Jen, 1976, 1977; Suga et al., 1976, 1983a; Ostwald, 1980; O'Neill et al., 1985). The functional organization of the auditory cortex of the mustached bat is adjusted to the properties of the bat's own biosonar signal and consequently to the dialects of biosonar signals (Suga and O'Neill, 1980; Suga and Tsuzuki, 1985; Suga et al., 1987).

Since the auditory system must be functionally organized to fulfill the species-specific need, an animal's auditory center must be specialized quite differently from that of humans. Therefore, a question arises as to what kind of animals are suited for neuroethological studies for understanding the neural mechanisms of speech processing. To understand the neural mechanism of speech processing, which is highly specialized, the most suitable animals are probably those that are specialized in their auditory behavior, because the auditory systems of these animals will show us the upper limit of specialization of the neural mechanisms of that behavior. If animals and humans produce similar acoustic patterns, it is possible that they share similar neural mechanisms for processing these acoustic patterns, because there may be only a few different kinds of specialization for processing particular acoustic patterns. This does not mean that similar neural mechanisms extract similar types of auditory information, nor that similar functional organization of the auditory center indicates similar types of auditory information.

Remarks. There is no evidence to support the view that the data obtained from macaque monkeys are the most valuable for understanding the central mechanisms of speech processing in humans. The primary aim of auditory neuroethology is to explore the neural mechanisms of the auditory behavior of animals, not to obtain data valuable to understanding speech processing. However, neuroethological studies on the auditory systems of different types of animals, including monkeys, will contribute to understanding the basic mechanisms for speech-sound processing, because such studies have been and will be successful in finding the neural mechanisms that are shared by different types of animals, as well as the mechanisms that are species specific. Studies on auditorily specialized animals may contribute most to understanding speech-sound processing, which is a highly specialized function.

ACOUSTIC SIGNALS AND NEURAL REPRESENTATIONS

In higher vertebrates, communication sounds are usually complex. The amplitude spectra of these sounds commonly change with time. Human speech consists

of various phonemes combined in different sequences and is consequently very complicated. Sonograms of animal sounds, including speech sounds, exhibit three basic patterns or components: constant frequency (CF), noise burst (NB), and frequency-modulated (FM) components (Figure 3A).

For instance, the sonogram of a consonant consists of a vertical bar indicating the scatter of sound energy over many frequencies. This is called a fill, or noise burst. The fricative consonants /s/ and /sh/, for example, can be recognized by their high-energy noise bursts concentrated around 2–3 kHz and above 4.4 kHz, respectively. The sonogram of a vowel consists of several horizontal bars called formants, which are spectral peaks characterizing a vowel. The lowest is called the first formant (F_1), the second lowest is the second formant (F_2), and so on. These may be considered CF components. Vowels are identified by combinations of F_1 , F_2 , and F_3 , so that they are expressed by loci in frequency-versus-frequency coordinates (Figure 3C). When two phonemes are combined to form a monosyllabic word, a new component called a transition appears. The transition is an FM component, since its frequency changes with time (Figure 3A). This FM component is very important for the identification of consonants within words. For example, the plosive consonants /k/, /t/, /p/, /g/, /d/, and /b/ are identified by the transitions before F_1 and F_2 of a vowel (Figure 3D). Human speech sounds thus consist of three types of information-bearing elements.

These three types of information-bearing elements are also found in sounds produced by many different types of animals. For instance, the mustached bat, *Pteronotus parnellii*, emits complex biosonar signals consisting of long CF and short FM components. It also emits at least seven different types of nonorientation sounds, presumably communication sounds. Two of the seven are noise bursts, and three others are long sounds with many harmonics. The remaining two are somewhat similar to the biosonar signal.

Auditory information is carried not only by the acoustic parameters characterizing each of the above three types of information-bearing elements, but also by parameters representing relationships among these elements in the frequency, amplitude, and time domains. For instance, voice onset time, a time interval between two acoustic events (Figure 3A), is another important cue for speech recognition (Figure 3B). Target range information important for a bat is carried not by an echo but by a delay of the echo from an emitted biosonar signal. Information for sound localization is carried not by the parameters characterizing a sound but by interaural time and amplitude differences, which are created by a receiver. A parameter is a continuum and can have any value. However, only a limited part of the continuum is important for each species. This limited part of the continuum has been called the information-bearing parameter, or IBP (Suga, 1982, 1984). The same IBP may have a different biological significance for different species. For example, in the mustached bat, a parameter characterizing combinations of CF tones carries target velocity information instead of vowel information, and a parameter characterizing combinations of FM sounds carries target range information instead of information on phoneme combinations.

Representation of Auditory Information in Hypothetical Centers

In the mammalian cochlea, sensory cells may be viewed as filters arranged along the basilar membrane for frequency analysis. Their outputs are coded by primary

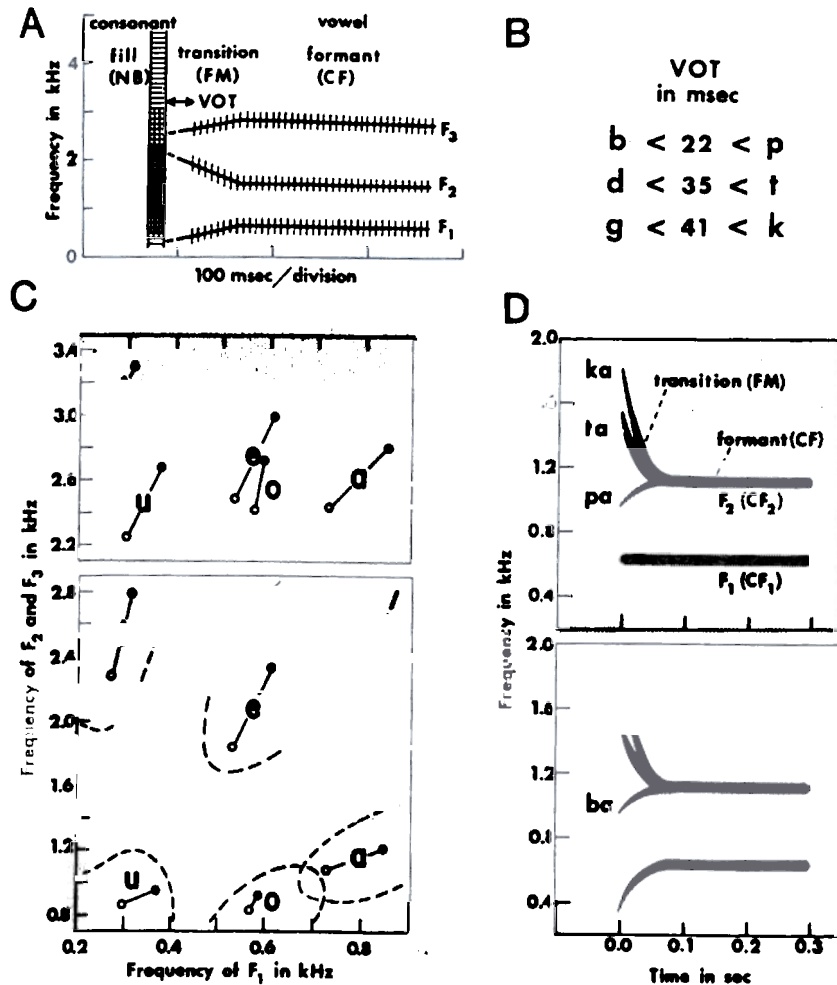


Figure 3. Information-bearing elements in human speech sounds. **A:** Schematized sonogram of a monosyllabic sound shows four types of information-bearing elements—fill or noise burst component (NB); transition or frequency-modulated component (FM); formant or constant frequency component (CF); and voice onset time (VOT) or time interval between two acoustic events. F_1 , F_2 , and F_3 are the first, second, and third formants of a vowel, respectively. **B:** The phonetic boundary of VOT is 22 msec for the consonants /b/ and /p/, 35 msec for /d/ and /t/, and 41 msec for /g/ and /k/. For example, when the VOT is shorter than 22 msec, a monosyllabic sound is recognized as /ba/. When it is longer than 22 msec, however, it is recognized as /pa/. (Based on Lisker and Abramson, 1964.) **C:** The relationship among the frequencies of the F_1 , F_2 , and F_3 of five vowels (/a/, /e/, /i/, /o/, and /u/). The average frequencies of the formants differ between male (open circles) and female (filled circles) speakers. The areas surrounded by dashed lines in the lower figure represent formant frequencies for unanimous classification of vowels by listeners. (Based on Peterson and Barney, 1952.) **D:** Schematized sonograms of monosyllabic sounds. /a/ changes into /pa/, /ta/, or /ka/ through the addition of different FM components to F_2 . /pa/, /ta/, and /ka/ become /ba/, /da/, and /ga/, respectively, through the addition of an FM component to F_1 . (From Suga, 1984, based on Liberman et al., 1956.)

auditory neurons that discharge action potentials at higher rates for larger outputs. Therefore, at the periphery the frequency of an acoustic signal is expressed by the location of activated neurons and its amplitude by their discharge rates. The duration of the signal and the interval between signals are expressed by the temporal pattern of neural activity. The peripheral auditory system has an anatomical axis for frequency only. The activity of individual peripheral neurons cannot uniquely express the properties of an acoustic signal. For instance, a peripheral neuron tuned to 40 kHz responds not only to a pure tone of 40 kHz but also to an FM sound sweeping across 40 kHz regardless of sweep direction and to a noise burst containing 40 kHz regardless of bandwidth. Therefore, the properties of an acoustic signal are appropriately expressed only by the spatiotemporal pattern of the activity of all peripheral neurons.

Action potentials sent into the brain by peripheral neurons are transmitted to many auditory nuclei and finally to the cerebral cortex, where a multiple representation of the cochlea is prominent. This multiple cochleotopic representation suggests that separate auditory areas are concerned with the representation of different types of auditory information. What kind of information is represented in each area? How is each area functionally organized? The neurophysiological data obtained thus far suggest at least four possible functional organizations of the auditory cortex. Therefore, four working hypotheses of the functional organization of a hypothetical auditory center may be proposed to explain the neural representation or recognition of acoustic signals. These hypotheses are not exclusive but are individually valid, depending upon the types of auditory information and species (Suga, 1982).

The Amplitude Spectrum Hypothesis. This hypothesis suggests that an acoustic signal is represented by the spatiotemporal pattern of activity of "nonspecialized" neurons arranged along the coordinates of frequency versus amplitude at a hypothetical auditory center and is not represented by the activity of neurons that respond selectively to the signal (Figure 4A). According to this hypothesis, each neuron has a small excitatory area tuned to a particular frequency and amplitude and responds to many different types of sounds containing a component that stimulates this area.

A frequency axis has been demonstrated in the auditory cortex of many different species of mammals, including bats (Suga, 1965b; Suga and Jen, 1976), cats (Merzenich et al., 1975), and monkeys (Brugge and Merzenich, 1973). Neurons tuned to particular amplitudes as well as frequencies are common in the particular subdivisions of the auditory cortex of bats (Suga, 1977; Suga and Manabe, 1982), cats (Phillips and Orman, 1984), and monkeys (Brugge and Merzenich, 1973), so we know that the central auditory systems of these animals have neurons necessary for amplitude spectrum representation. However, the amplitude axis (amplitopic representation) has not yet been found, except in the DSCF area of the auditory cortex of the mustached bat.

An amplitude spectrum can be expressed in the coordinates of frequency versus threshold as in the auditory periphery (Liberman, 1978, 1982). Threshold representation by neurons with a monotonic impulse count function has two disadvantages compared with amplitopic representation by neurons with a non-

Representation of acoustic signals in a hypothetical auditory center

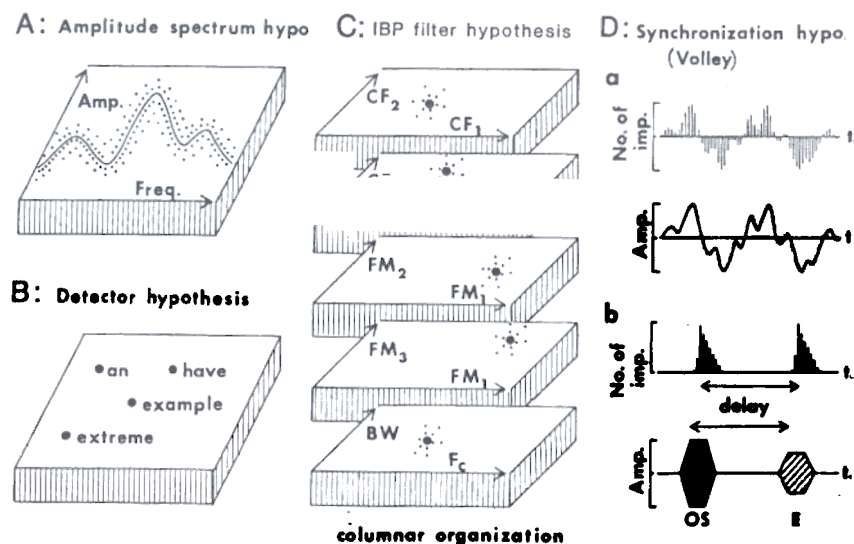


Figure 4. Four working hypotheses for the representation of auditory information by neural activity in a hypothetical center. **A:** The amplitude spectrum hypothesis, which may also be called the simple spatiotemporal pattern hypothesis. **B:** The detector hypothesis. **C:** The IBP filter hypothesis. Here coordinates are formed by neurons tuned to information-bearing parameters (IBPs) that characterize biologically important signal elements. CF_1 , CF_2 , CF_3 , FM_1 , FM_2 , and FM_3 are different signal elements in complex sounds, and the axes represent the parameters characterizing these signal elements. F_c and BW are the center frequency and bandwidth of noise bursts, respectively. In all three hypotheses, the properties of acoustic signals are represented by the spatiotemporal pattern of neural activity. The response properties of individual neurons and the interpretation of their functions are, however, different according to the hypotheses. **D:** The synchronization hypothesis, which is the extended version of Wever's volley theory. In **D_a**, the lower trace represents a sound wave; the upper trace is the compound period histogram of a single neuron response to it. In **D_b**, the lower trace represents an orientation sound (OS) and an echo (E); the upper trace is the PST histogram of a single neuron response. (From Suga, 1962.)

monotonic impulse count function. In threshold representation, (1) more neurons are involved in representing the amplitude spectrum of a stronger acoustic stimulus; and (2) the boundary between excited and nonexcited neurons is unclear and fluctuates.

The threshold representation found in the dog by Tunturi (1952) is intriguing, but it has not been accepted by auditory physiologists, probably for the following two major reasons: (1) The threshold representation was found only for the stimulation of the ipsilateral ear, which does not supply the main excitatory input to the auditory cortex. The main excitatory input comes from the contralateral ear and is uniformly lower in threshold than that from the ipsilateral ear. (2) The functional significance of the threshold representation is not clear, because most

of the cortical auditory neurons show relatively phasic on-responses to tone bursts and/or nonmonotonic impulse count functions.

The Detector Hypothesis. This hypothesis suggests that a biologically important acoustic signal is represented by the excitation of "detector" neurons that selectively respond to that particular signal (Figure 4B). Different types of detector neurons are, of course, arranged in a particular spatial pattern in the hypothetical center. The spatiotemporal pattern of neural activity in the center will thus change according to the sequence of biologically significant sounds. The essential distinction of the detector hypothesis is a one-to-one correspondence between signal and excitation of a detector neuron or neurons in a single cortical column. Neurons specialized to respond selectively to certain types of acoustic stimuli have been found in a few species of mammals (bats reviewed by Suga, 1973, 1984; monkeys reviewed by Symmes, 1981). It remains to be ascertained whether some "specialized" neurons that have been found can appropriately be called detectors or not.

The Information-Bearing Parameter Filter Hypothesis. This hypothesis falls between the previous two hypotheses. In auditory neurophysiology, the statement that neurons respond to sound x but not to other sounds is justified only when the filter properties of the neurons (tuning curves of the neurons to individual parameters characterizing sound x) are studied. When the filter properties are studied, we can treat the neurons as filters that correlate acoustic signals with their filter properties, that is, stored information. The degree of correlation is expressed by the magnitude of the output of the filters. In other words, neurons are maximally excited only when the properties of acoustic signals perfectly match their filter properties. All neurons in the auditory system, including peripheral ones, act as filters. Specialized neurons expressing the outputs of neural circuits tuned to particular IBPs or particular combinations of IBPs are called IBP filters (Suga, 1978).

Communication sounds are commonly characterized by many different parameters: frequency, FM rate, FM depth, amplitude, AM rate, AM depth, harmonics, duration, interval, and so on. Some parameters are IBPs. The IBP filter hypothesis states that the auditory center represents IBPs or combinations of IBPs by spatiotemporal patterns of activities of specialized neurons (IBP filters) acting as a kind of cross-correlator (Figure 4C). It also states that different types of IBP filters are aggregated separately in identifiable areas of the auditory center. This hypothesis is strongly supported by the data obtained from the auditory cortex of the mustached bat, which is described later.

The Synchronization Hypothesis. This hypothesis is an expanded version of the volley theory (Wever, 1949). When information-bearing elements are lower than 5 kHz, peripheral neurons produce discharges synchronized with the sound waves. The envelope of a compound period histogram of a neural response thus reproduces the stimulus waveform (Figure 4D_a). The synchronization hypothesis states that neurons in the hypothetical auditory center represent acoustic signals by synchronous discharges. In small mammals, many of the predominant components of their communication sounds are higher than 5 kHz and hence cannot

be coded by synchronous discharges, but the rate of sound emission or amplitude modulation is usually less than 1 kHz. For example, echolocating bats emit ultrasonic signals and listen to echoes. The rate of sound emission is usually less than 200/sec. The delay of an echo from the emitted sound is the primary cue for target ranging. Peripheral neurons show discharges synchronized with each emitted sound and echo. Therefore, range information is coded by the interval between a pair of grouped discharges (Figure 4D_b). The synchronization hypothesis further states that range perception is a direct consequence of paired grouped discharges but not of an excitation of neurons specialized for responding to particular echo delays.

In the cat medial geniculate body, only 3% of neurons that are activated by tone bursts up to 1 kHz show good phase-locked responses. Phase-locked responses to sounds higher than 1 kHz are hardly ever found (Rouiller et al., 1979). It is assumed that both the population of phase-locking neurons and the degree of phase locking are lower in the auditory cortex than in the medial geniculate body. Therefore, the synchronization hypothesis is limited in validity and can be true only for representation of a part of the properties of acoustic signals, for example, fundamentals.

In the FM-FM area of the auditory cortex of the mustached bat, neurons are tuned to particular time intervals between the orientation sound and echo. The time intervals are systematically represented by the loci of activated neurons in this area (Suga and O'Neill, 1979; O'Neill and Suga, 1982). Therefore, it is possible that the temporal code at the periphery is changed into a place code at the auditory center.

Synchronous (phase-locked) discharges play an important role in information processing for sound localization (Rose et al., 1966; Yin and Kuwada, 1984; Sullivan and Konishi, 1984). It is, however, possible that the location of a sound source is represented not by a magnitude of phase-locked discharges but by a locus of activated neurons that are tuned to particular interaural time or phase differences, as found in the barn owl (Knudsen and Konishi, 1978; Moiseff and Konishi, 1983).

Remarks. Human speech and animal sounds consist of three types of information-bearing elements. These elements and the relationships among them are characterized by particular values of IBPs. The neurophysiological data thus far obtained suggest at least four hypotheses as to the functional organization of a hypothetical auditory center, which may differ among subdivisions of the auditory cortex. It may also differ among species that use different types of acoustic signals and extract auditory information for different purposes. The same or a similar functional organization found among different species may have different biological significance for each species.

PROCESSING OF COMPLEX BIOSONAR SIGNALS BY COMBINATION-SENSITIVE NEURONS

In order to capture prey (flying insects) and orient itself in space, the mustached bat emits orientation sounds (biosonar signals or pulses), each of which consists of a long CF component followed by a short FM component. Since each orientation

sound contains four harmonics (H_{1-4}), there are eight components that can be defined (CF_{1-4} , FM_{1-4}). In the emitted sound, the second harmonic (H_2) is always predominant, and the frequency of CF_2 is about 61 kHz. (The frequency of the CF component is different among subspecies and also to some extent among individuals of the same subspecies.) In FM_2 , the frequency sweeps down from 61 kHz to 49 kHz (Figure 5A). H_3 is 6–12 dB weaker than H_2 , while H_1 and H_4 are 18–36 and 12–24 dB weaker than H_2 , respectively. Many species of moths have ears that are most sensitive to frequencies between 20 and 40 kHz (Suga, 1961; Fenton and Fullard, 1979) and show evasive behavior when they

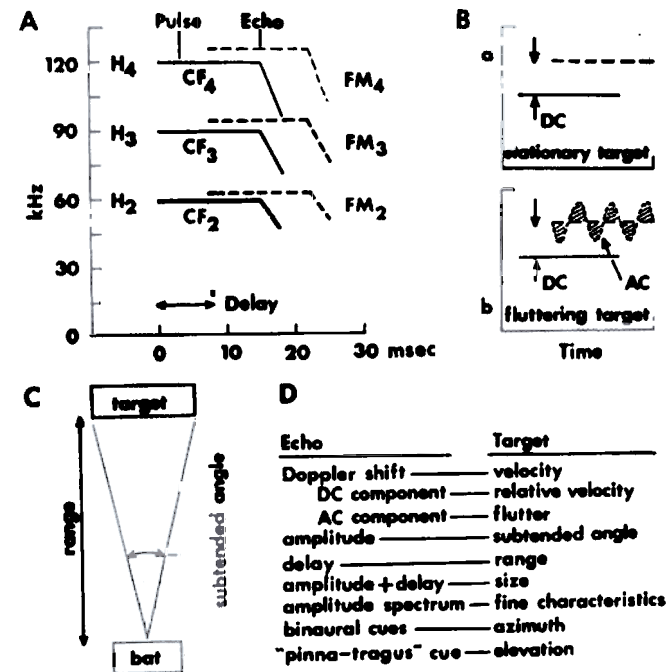


Figure 5. Biosonar signals of the mustached bat, *Pteronotus parnellii rubiginosus*, and the information carried by its signals. A: Schematized sonogram of the mustached bat orientation sound (solid lines) and the Doppler-shifted echo (dashed lines). The orientation sound is also called a pulse. The four harmonics (H_{1-4}) of both the orientation sound and the echo each contain a long CF component (CF_{1-4}) and a short FM component (FM_{1-4}). The thickness of the lines indicates the relative amplitude of each harmonic in the orientation sound: H_2 is the strongest, followed by H_3 (6–12 dB weaker than H_2), H_4 (12–24 dB weaker than H_2), and H_1 (18–36 dB weaker than H_2). B: When the mustached bat flies toward or near a stationary object, the frequency of the echo becomes higher than the emitted sound (graph a). This steady shift is called the DC component of the Doppler shift. When the bat flies toward a fluttering target—a flying moth, for example—the Doppler shift of the echo consists of a DC component proportional to the relative velocity and the periodic frequency modulation (FM) proportional to the speed of wing beat (graph b). This periodic FM is called the AC component of the Doppler shift. The AC component is complicated because the insect's four wings are moving in complex patterns and in different phase relationships relative to the bat. C: Target size is determined from both target range and subtended angle. D: Relationship between echo properties and target properties. (From Suga, 1982.)

hear the orientation sounds of bats (Roeder, 1962). Since H_1 (18–36 kHz) in the orientation sound of the mustached bat is suppressed, probably by anti-resonance of the vocal tract, the bat can approach these moths closely before being detected (Suga and O'Neill, 1979).

Echoes eliciting behavioral responses in the mustached bat always overlap temporarily with the emitted signal. As a result, biosonar information must be extracted from a complex sound containing up to 16 components. The CF component is an ideal signal for target detection and the measurement of target velocity because the reflected sound energy is highly concentrated at a particular frequency. The mustached bat uses the CF component for this purpose and performs a unique behavior called Doppler-shift compensation (see below). The short FM component, on the other hand, is more appropriate for ranging, localizing, and characterizing a target because its sound energy is distributed over many frequencies. Different parameters of echoes received by the bat carry different types of information about a target (Figure 5D).

During target-directed flight, the duration of the orientation sound shortens from 30 to 7 msec, and the emission rate increases from 5/sec to 100/sec. The shortening of the duration of the sound is mainly due to shortening of the CF component. Such changes in the signal cause a decrease in the information carried by the CF component but an increase in the information carried by the FM component. The echo amplitude and delay from the emitted sound also change systematically during target-directed flight.

One of the fascinating acoustic behaviors of the mustached bat is Doppler-shift compensation (Schnitzler, 1970). When the bat flies toward or near targets, it reacts to Doppler-shifted echoes resulting from its approach by reducing the frequency of its emitted sounds, that is, by "compensating" for Doppler-shift. This Doppler-shift compensation stabilizes the echo at a predetermined frequency. For example, the mustached bat emits an orientation sound with CF_2 at 61 kHz when there are no Doppler-shifted echoes. When a Doppler-shifted echo returns, for example, at 63 kHz, the bat reduces the frequency of subsequent orientation sounds by nearly 2 kHz, so that the Doppler-shift echo is stabilized at or near 61 kHz. Because of this compensation, the frequency of the CF_2 of the Doppler-shifted echo is kept mainly within a range of 61–62 kHz, to which the bat's cochlea is extraordinarily sharply tuned.

Parallel Hierarchical Processing

The eight components (CF_{1-4} , FM_{1-4}) of the orientation sound of the mustached bat are all different from each other in frequency, so that they are analyzed in parallel at different regions of the basilar membrane (Figure 6, bottom). Then the signals are coded and sent into the brain by peripheral neurons. In the brain, the signals are sent up to the auditory cortex through many auditory nuclei. For simplicity, we may consider that there are eight channels for processing of these signal elements: channel CF_1 , channel CF_2 , and so on. Channel CF_2 is very large in comparison with any of the other channels and is associated with extraordinarily sharply tuned local resonators in the cochlea for fine frequency analysis (Figure 6, leftmost column). Channel CF_4 is very small, if it exists at all.

In channels CF_1 , CF_2 , and CF_3 (Figure 6, left-hand columns), frequency selectivity is increased and amplitude selectivity is added by inhibition of some neurons

in the cochlear nucleus (Suga et al., 1975) and also of many neurons at higher levels (Suga and Manabe, 1982; O'Neill, 1985; Olsen, 1986). These neurons are thus tuned to particular frequencies and amplitudes. The extent to which neural sharpening occurs is different among groups of neurons tuned to different frequencies. The sharpening is most dramatic in channel CF_2 . In a certain region of the medial geniculate body, a part of channel CF_1 and a part of channel CF_2 or CF_3 are integrated, so that neurons in this region respond poorly to CF_1 , CF_2 , and CF_3 tones when delivered alone but strongly when the CF_1 tone is delivered together with the CF_2 or CF_3 tone. These CF/CF combination-sensitive neurons project to the CF/CF area of the auditory cortex (Olsen and Suga, 1983; Olsen, 1986). In the CF/CF area, two types of CF/CF neurons, CF_1/CF_2 and CF_1/CF_3 , are separately clustered and form the frequency-versus-frequency coordinates in each cluster for the representation of Doppler shifts, that is, target velocity information (Suga et al., 1981, 1983a). These neurons show sharp level-tolerant frequency tuning curves and are remarkably specialized to respond to particular frequency relationships of two CF tones (Suga and Tsuzuki, 1985). The signal processing in the CF channels is thus "parallel hierarchical."

In channels FM_1 , FM_2 , FM_3 , and FM_4 (Figure 6, center column), frequency selectivity is increased and amplitude selectivity is added by inhibition to some neurons. Interestingly, FM selectivity is also added to some neurons by disinhibition, so that these "FM-specialized" neurons respond to FM sounds but not to CF tones and noise bursts (Suga, 1965a; O'Neill, 1985). In a certain region of the medial geniculate body, a part of channel FM_1 and a part of channel FM_2 , FM_3 , or FM_4 are integrated, so that neurons in this region respond poorly to FM_1 , FM_2 , FM_3 , and FM_4 sounds when delivered alone but strongly to the FM_1 sound combined with the FM_2 , FM_3 , or FM_4 sound. The time interval between two FM sounds is a critical parameter for their facilitative responses. These FM-FM combination-sensitive neurons project to the FM-FM area of the auditory cortex (Olsen and Suga, 1983; Olsen, 1986). In the FM-FM area, three types of FM-FM neurons, FM_1-FM_2 , FM_1-FM_3 , and FM_1-FM_4 , are separately clustered and form an echo-delay, or time, axis in each cluster for the representation of target range information (Suga and O'Neill, 1979; O'Neill and Suga, 1982; Suga and Horikawa, 1986). Therefore, signal processing in the FM channels is also parallel hierarchical.

As described above, a part of one channel is integrated by a part of the other channel in the medial geniculate body. The remaining parts of these channels that are not integrated project to the auditory cortex, which is not described above. For instance, a part of the CF_2 channel projects to the DSCF area of the auditory cortex, which has the frequency-versus-amplitude coordinates to represent the amplitude spectrum of the CF_2 component of a Doppler-shifted echo. The DSCF area can be divided into two subdivisions that predominantly contain IE or EE neurons (Figure 6, right-hand column). In the cat, IE and EE neurons have been found in the pontine auditory nuclei and also at higher levels. Therefore, it is possible that some of the channels described above consist of subchannels in terms of processing of other types of auditory information. Figure 6 explains only as much parallel hierarchical processing of biosonar information as has so far been explored.

Almost all frequencies found in the biosonar signals are projected not only to the areas that appear to be important to echolocation but also to the areas

that appear not to be important to echolocation. These areas are probably important for the processing of communication sounds. Except for the CF₂ channel, which is specialized for processing biosonar information from the periphery through the auditory cortex, a clear separation of biosonar signal processing from non-biosonar signal processing takes place first in the medial geniculate body.

Functional Organization of the Auditory Cortex

The auditory cortex of the mustached bat is 0.9 mm thick and about 14.2 mm² in size, which is very large relative to the size of the brain (Figure 7). The auditory cortex shows an intriguing multiple cochleotopic (tonotopic) representation that is directly related to the representation of different types of biosonar information (Suga, 1984). Figure 7 shows several functional subdivisions explored electrophysiologically. In these areas, certain response properties of single neurons arranged orthogonally to the cortical surface are identical. In this sense there is a columnar organization. Along the cortical surface, however, the response properties vary systematically and form axes for the representation of particular types of biosonar information.

Among the several functional subdivisions, the CF/CF, FM-FM, DF, VF, and VA areas consist of combination-sensitive neurons, so that these areas are particularly interesting in terms of neural mechanisms for complex-sound processing. The following describes the response properties of neurons in the CF/CF and FM-FM areas as examples of combination-sensitive neurons. [See my review (1984) for further information on the auditory cortex of the mustached bat.]

Representation of Target Velocity in the CF/CF Area. When the mustached bat flies toward a flying insect, the echo from the insect contains two components of Doppler shift: a DC component caused by the change in distance between the bat and the insect and an AC component caused by the insect's wingbeats (Figure 5B). How do cortical auditory neurons respond to the DC and AC components of Doppler shift in processing target velocity information?

The DC component is expressed by the frequency difference between the emitted sound and the Doppler-shifted echo. The frequency information of the emitted sound is available to the bat in the form of vocal self-stimulation and perhaps efferent copy, about which little is known. The frequency information contained in the Doppler-shifted echo is available regardless of whether Doppler-shift compensation is performed. With compensation, however, the measurement of echo frequency becomes much more accurate, because the echo frequency is then analyzed by a group of filters in the cochlea that are sharply tuned to the CF₂ sound between 60 and 62 kHz (Suga et al., 1975; Suga and Jen, 1977). At the periphery, the DC component is expressed by two groups of activated neurons with different best frequencies (BFs).

To represent a Doppler shift systematically, frequency-versus-frequency coordinates are needed: one to express the frequency of the emitted sound, the other to express that of a Doppler-shifted echo. There are obviously no such coordinates in the cochlea. In the CF/CF area, however, neurons are tuned to particular combinations of two CF components (Figures 8, 9; Suga et al., 1979,

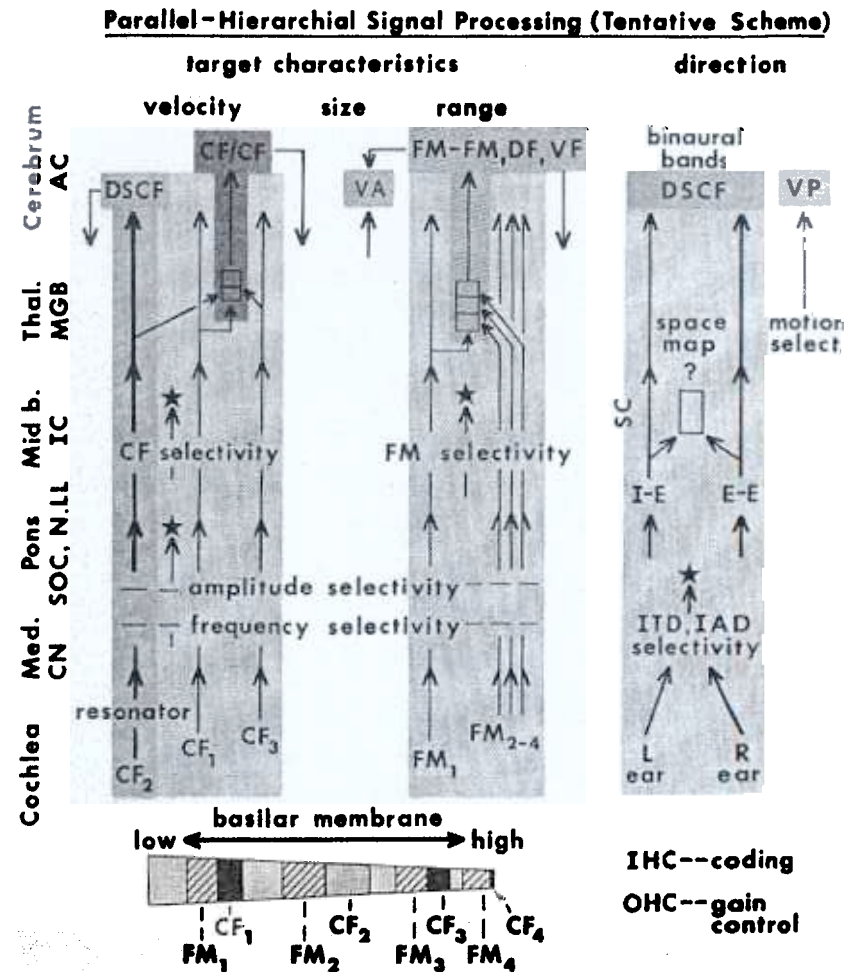


Figure 6. Parallel hierarchical processing of different types of biosonar information carried by complex biosonar signals. The CF₁₋₄ and FM₁₋₄ of the orientation sound are analyzed at different parts of the cochlear basilar membrane (bottom). Inner and outer hair cells (IHC and OHC) on the membrane are related to stimulus coding and gain control, respectively. These signal elements are separately sent up to the auditory cortex (AC) through several auditory nuclei (left margin): the cochlear nucleus (CN), superior olivary complex (SOC), nucleus of lateral lemniscus (N.LL), inferior colliculus (IC), and medial geniculate body (MGB). During the ascent of the signals, frequency, amplitude, CF, and FM selectivities are added to some neurons (arrows with a star). Each star indicates that the addition of selectivity also takes place in the auditory nuclei and cortex as well as in the nucleus where the arrow starts. The CF₂ processing channel is disproportionately large and projects to the DSCF area of the auditory cortex (leftmost column). In certain portions of the MGB, two channels processing different signal elements (e.g., channels CF₁ and CF₂ or FM₁ and FM₂) are integrated to produce combination-sensitive neurons (two left-hand columns). CF/CF and FM-FM combination-sensitive neurons project to the CF/CF and FM-FM areas of the auditory cortex, respectively, where target velocity or range information is systematically represented. Because of corticocortical connections, the DF, VF, and VA areas also consist of combination-sensitive neurons (top of center column). Target velocity and range information is thus processed in a parallel hierarchical manner. These channels may be subdivided in terms of binaural interaction. The DSCF area consists of two subdivisions mainly consisting of IE or EE neurons. The superior colliculus (SC) may have an auditory space map (right-hand column). Motion-selective neurons appear to be in the ventroposterior (VP) area of the auditory cortex.

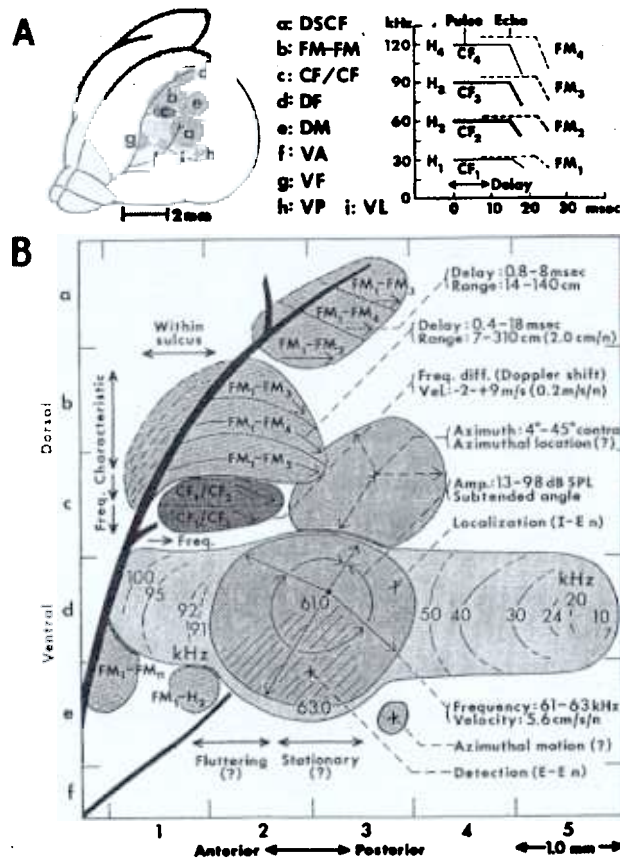


Figure 7. Functional organization of the auditory cortex of the mustached bat. **A:** Dorsolateral view of the left cerebral hemisphere. The branches of the median cerebral artery are shown by the branching lines. The longest branch is on the sulcus. **B:** Graphic summary of the functional organization of the auditory cortex. The tonotopic representation of the primary auditory cortex and the functional organization of the DSCF, FM-FM, CF/CF, DF, and DM areas are indicated by lines and arrows. The DSCF area has axes representing either target velocity (echo frequency: 61-63 kHz) or subtended target angle (echo amplitude: 13-98 dB SPL) and is divided into two subdivisions suitable for either target detection (shaded area) or target localization (unshaded area). The FM-FM area consists of three major types of FM-FM combination-sensitive neurons (FM_1 - FM_2 , FM_1 - FM_3 , and FM_1 - FM_4), which form separate clusters. Each cluster has an axis representing target ranges of 7-310 cm (echo delay: 0.4-18 msec). The dorsoventral axis of the FM-FM area probably represents fine target characteristics. The CF/CF area consists of two major types of CF/CF combination-sensitive neurons (CF_1 / CF_2 and CF_1 / CF_3), which aggregate in independent clusters. Each cluster has two frequency axes and represents target velocities of -2 to +9 m/sec (echo Doppler shift: -0.7 to +3.2 kHz for CF_2 , and -1.1 to +4.8 kHz for CF_3). The DF area and a posterior part of the VA area receive nerve fibers from the FM-FM area. The DF area consists of the three types of FM-FM neurons, but the VA area contains only FM_1 - H_2 combination-sensitive neurons. The DF area projects to the VF area, which consists of the three types of FM-FM neurons. FM-FM and CF/CF neurons are tuned to particular combinations of two signal elements in a complex sound and extract target range and velocity information, respectively. The DM area appears to have an azimuthal axis representing the azimuthal location of the target. In the VP area, neurons sensitive to azimuthal motion have been found. The functional organization of the VF, VA, and VP areas remains to be studied further. (Based on data obtained by Suga and Jen, 1976; Suga, 1977; Manabe et al., 1978; Suga and O'Neill, 1979; Suga et al., 1981, 1983a,b; O'Neill and Suga, 1982; Suga and Manabe, 1982; Asanuma et al., 1983; Kujirai and Suga, 1983; Suga and Horikawa, 1986; Edamatsu et al., 1988; Tsuzuki and Suga, 1988.)

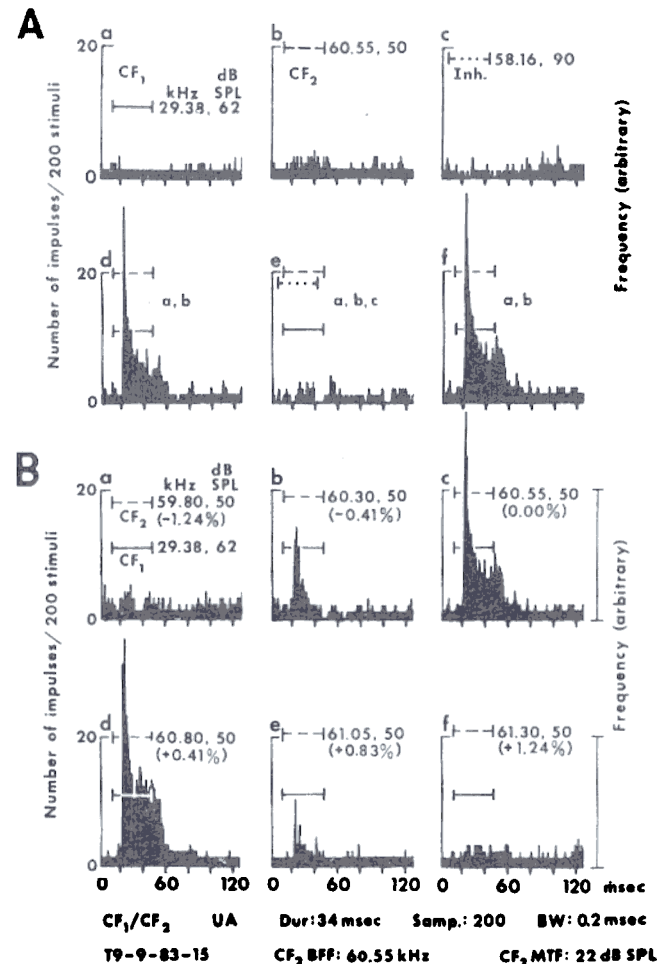
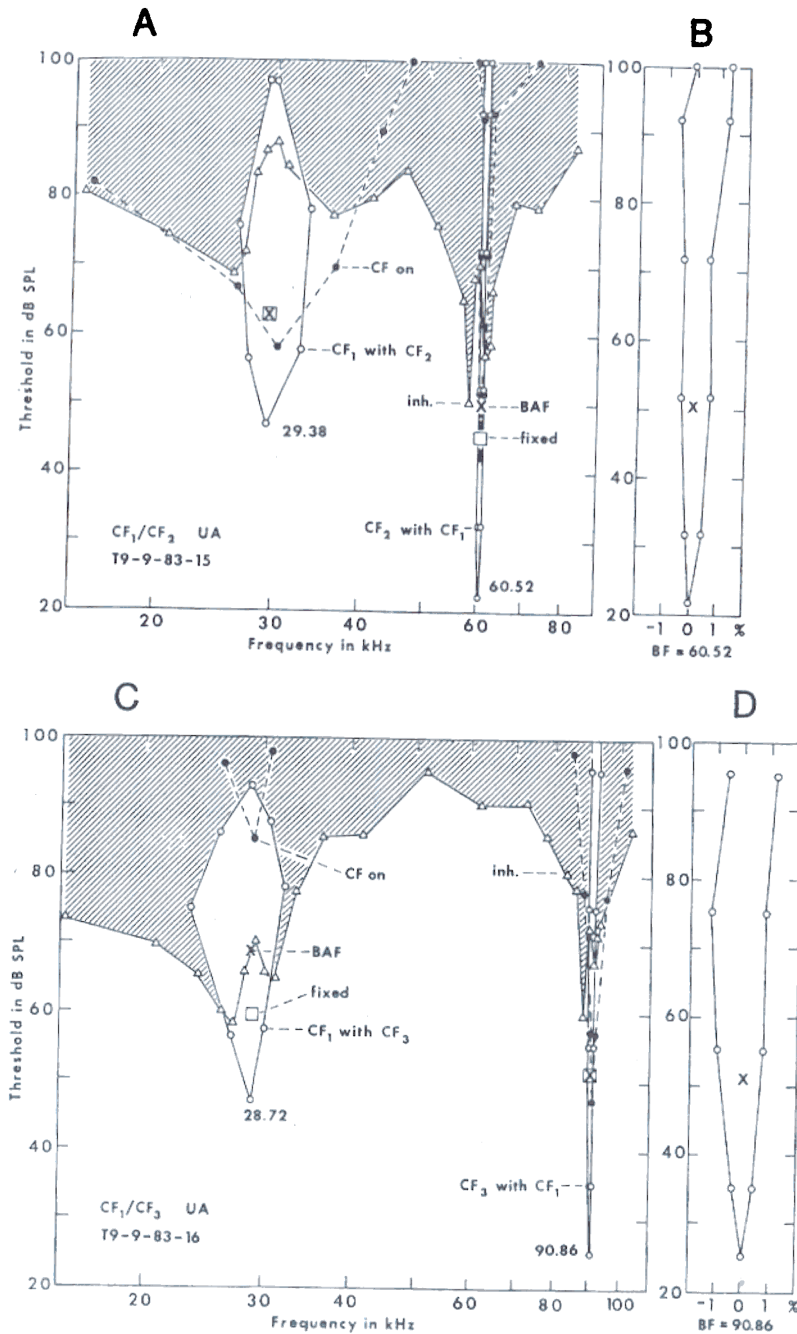


Figure 8. PST histograms displaying responses of a CF_1/CF_2 neuron to single tones (A_{a-f}) and combinations of these tones (A_{a-f} and B_{a-f}). Tonal stimuli are indicated by horizontal bars; their parameters, by numbers near bars. In A_a , for example, the CF_1 stimulus indicated by a horizontal bar is 29.38 kHz and 62 dB SPL. In A_d and A_f , CF_1 and CF_2 were delivered with a temporal overlap to demonstrate facilitative response. In A_e , an inhibitory tone (Inh), CF_1 , and CF_2 were delivered with an overlap to demonstrate inhibition of facilitation. Inhibitory sound by itself did not inhibit background discharges (A_c). In B , CF_1 and CF_2 were delivered with an overlap. CF_1 was fixed at BFF (29.38 kHz) and BAF (62 dB SPL), the best frequency and amplitude for facilitation, respectively. CF_2 was varied in frequency. The amount of frequency shift from CF_2 BFF (60.55 kHz) is expressed as a percentage of 60.55 kHz. The amplitude of CF_2 was fixed at CF_2 BAF (50 dB SPL), which was 28 dB above minimum threshold. Note the dramatic change in facilitation with a small shift in CF_2 frequency. Abbreviations: UA, unanesthetized bat; Dur, duration of tone burst; Samp., number of stimulus presentations for each PST histogram; BW, bin width of PST histograms; CF_2 BFF, CF_2 best frequency for facilitation; CF_2 MTF, CF_2 minimum threshold for facilitation. (From Suga and Tsuzuki, 1985.)



1983a; Suga and Tsuzuki, 1985) and form such frequency-versus-frequency coordinates (Figure 10; Suga et al., 1981, 1983a).

The CF/CF area contains two major types of combination-sensitive neurons, CF_1/CF_2 and CF_1/CF_3 , both of which respond poorly to individual signal elements but show remarkable facilitation when the CF_1 of the orientation sound (pulse) is combined with the CF_2 or CF_3 of an echo (Figure 8). Note that one of the essential signal elements for the facilitation of CF/CF neurons is CF_1 , the weakest component of the pulse. For maximal facilitation, a precise frequency relationship between the two combined elements is critical. Frequency tuning curves for facilitation are not particularly sharp for CF_1 but extremely sharp for CF_2 and CF_3 . The latter curves are as narrow as $\pm 1\%$ of BF, even at high stimulus levels (Figure 9). Consequently, they are called sharp level-tolerant frequency tuning curves.

Sharp level-tolerant tuning curves are sandwiched between broad inhibitory areas (Figure 9). The BFs for inhibition are slightly higher or lower than the BFs for facilitation. This leads us to the conclusion that sharp level-tolerant tuning curves are produced by inhibition. The extent to which neural sharpening occurs differs among groups of neurons tuned to different frequencies. The more important the frequency analysis of a particular component in biosonar signals, the more pronounced the neural sharpening. This is in addition to the peripheral specialization for fine frequency analysis of the component (Suga and Tsuzuki, 1985).

The difference in bandwidth or quality factor between the excitatory tuning curves of peripheral neurons and the facilitative and excitatory tuning curves of CF/CF neurons is larger at higher stimulus levels. Since the neural sharpening occurs mainly at the "skirt" of frequency tuning curves, the bandwidths at 10 dB and 20 dB above minimum threshold, Q_{-10} dB and Q_{-20} dB, are not appropriate for measuring whether the central auditory system has neural mechanisms for the sharpening of frequency tuning, but the bandwidth at 50 dB above minimum threshold, Q_{-50} dB, is (Suga and Manabe, 1982; Suga and Tsuzuki, 1985). [In the central auditory system, neurons are also produced that have a frequency tuning curve much broader than that of peripheral neurons (Erulkar, 1959; Hind, 1960; Katsuki et al., 1962; Grinnell, 1963; Suga, 1964, 1965a, b; Abeles and Goldstein, 1970; Goldstein et al., 1970). These neurons tend to be clustered at particular subdivisions of the auditory nuclei and cortex (Watanabe, 1959; Aitkin et al.,

Figure 9. Excitatory, facilitative, and inhibitory areas of a CF_1/CF_2 neuron, responses of which are shown in Figure 8. A: Excitatory area is above the excitatory frequency tuning curve (dashed lines, filled circles), which was measured with single tones. Facilitative areas are areas above or surrounded by facilitative frequency tuning curves (solid lines, open circles). The CF_1 facilitative area was measured with CF_2 fixed at 60.52 kHz and 45 dB SPL (square) and a test tone that was varied in frequency and amplitude. The CF_2 facilitative area was measured with CF_1 fixed at 29.38 kHz and 63 dB SPL (square) and a test tone that was varied in frequency and amplitude. B: The CF_2 facilitative area is so sharp that I replotted it on an expanded frequency axis that expresses frequency in percentage of BFF 60.52 kHz. BAFs are indicated by \times . The inhibitory area (shaded) was measured with a fixed CF_1/CF_2 pair, which evoked a clear facilitative response, and a test tone that was varied in frequency and amplitude. Note that the facilitative and inhibitory areas overlap at high stimulus levels. C and D: Excitatory, facilitative, and inhibitory areas of a CF_1/CF_3 neuron. Symbols are the same as in A and B. Note the very sharp CF_3 facilitative area. (From Suga and Tsuzuki, 1985.)

1975; Merzenich et al., 1975). To express the broadening of tuning curves, Q-50 dB is not an appropriate measure, but Q-10 dB and Q-30 dB are. The choice of parameters characterizing tuning curves depends on the problem to be discussed.]

The CF/CF area shows columnar organization: Neurons recorded in each orthogonal electrode penetration are characterized by an identical pair of best facilitative frequencies. The area consists of two rostrocaudally elongated bands containing either CF₁/CF₂ or CF₁/CF₃ neurons, which are located dorsally and ventrally, respectively (Figure 7). In both bands, the best facilitative frequency of CF₁ is represented along the rostrocaudal axis, while those for CF₂ and CF₃ are represented along the dorsoventral axis (Figures 7, 10A). These frequency

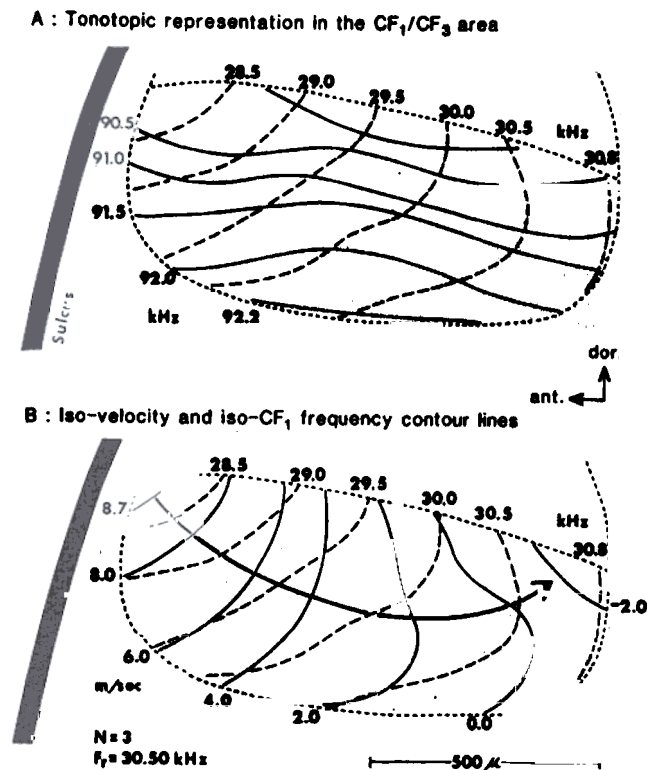


Figure 10. Functional organization of the CF₁/CF₃ area. A: Iso-best facilitative frequency contour lines for CF₁ (long dashes) and CF₃ (solid lines). These contour lines are based on data obtained from three unanesthetized mustached bats. B: Iso-velocity contour lines (solid) are shown together with the iso-best facilitative frequency contour lines for CF₁ (long dashes). The long arrow is the axis representing the Doppler shift, that is, the target velocity (8.7 to -2.0 m/sec) in the radial direction. Note the overrepresentation of speeds of 0-4 m/sec. The figure indicates, for example, that when the bat emits an orientation sound with a CF₁ of 30.5 kHz (resting frequency), neurons in the CF₁/CF₃ area are best activated by targets moving with relative velocities of -1.2 to +2 m/sec. When the frequency of the CF₁ is reduced to 29.5 kHz, however, they are best stimulated by targets moving with relative velocities of 3.4-6.1 m/sec. Iso-velocity contour lines similar to the above have also been found in the CF₁/CF₂ area. (From Suga et al., 1981.)

axes are based on the best facilitative frequencies of single neurons. When tonotopic representation in the CF/CF area was studied in the traditional way with single CF tones, the poor responses of the neurons to such stimuli misrepresented the nature of this area.

Almost all CF/CF neurons are tuned to a combination of two CF tones that are not exactly harmonically related to each other. The deviation (D) of a CF₂ or CF₃ best facilitative frequency from an exact harmonic relationship with a CF₁ best facilitative frequency is assumed to correspond to a Doppler shift in terms of the CF₂ or CF₃ of an echo, respectively. ($D = CF_2 - 2CF_1$ or $CF_3 - 3CF_1$, where CF₁, CF₂, and CF₃ are the best facilitative frequencies of CF₁/CF₂ and CF₁/CF₃ neurons.) Then D can be transformed into a relative target velocity. For example, a CF₁/CF₂ neuron tuned to a combination of 30 kHz and 61 kHz is assumed to be tuned to a 1-kHz Doppler shift, that is, a 2.84 m/sec relative target velocity.

This frequency-versus-frequency coordinate system systematically represents the DC component of a Doppler shift. Iso-Doppler shift (iso-velocity) contour lines are oblique to both frequency axes and represent velocities from 8.7 to -2 m/sec (Figure 10B). Since the bat does not fly backward, the area representing velocities of 0 to -2 m/sec probably responds to echoes from insects and conspecifics flying away from the echolocating bat. Velocities of 0-4 m/sec are overrepresented; the relative speed of a target may be predominantly within this range in the approach and terminal phases of echolocation.

Since the response properties of CF/CF neurons are very important in terms of neural mechanisms for the processing of complex sound, their response properties in the amplitude and time domain are summarized below. Most CF/CF neurons are tuned to particular amplitude relationships between the combined CF sounds, so they may play a role in target characterization, even though the CF signal is not particularly well suited to this task. In many CF/CF neurons, a decrease in the threshold of response by facilitation is prominent. It can be as large as 76 dB, which corresponds to an approximately 6300-fold increase in sensitivity (Suga et al., 1983a).

CF/CF neurons are equally sensitive to 0-10 msec echo delays, and their responses to orientation sound-echo pairs are similar, regardless of echo delay when it is within this range. Therefore, these neurons are not suitable for processing distance information, unlike FM-FM neurons, which are described later. The responses of CF/CF neurons start to deteriorate when the echo delay becomes longer than 10 msec. At a delay longer than 20 msec, the facilitation becomes very poor and the facilitation threshold becomes high, even though the two signal elements still overlap significantly (Figure 11A). Such response properties of the neurons act as a kind of time gate for echo processing that is reset by the beginning of each successive orientation sound (Suga et al., 1981, 1983a). The greater horseshoe bat, *Rhinolophus ferrumequinum*, performs Doppler-shift compensation only when the echo delay is less than 15 msec (Schuller, 1974). Our neurophysiological data show an interesting correlation with this behavior.

The AC component of a Doppler shift caused by the insect's wingbeat is expressed by synchronous discharges of each neuron and by synchronous alteration of excitation among neurons tuned to different frequencies. The peripheral neurons sharply tuned to the CF₂ sound are extremely sensitive to the AC component and can code it at as little as 0.01% with synchronous discharges (Suga and Jen,

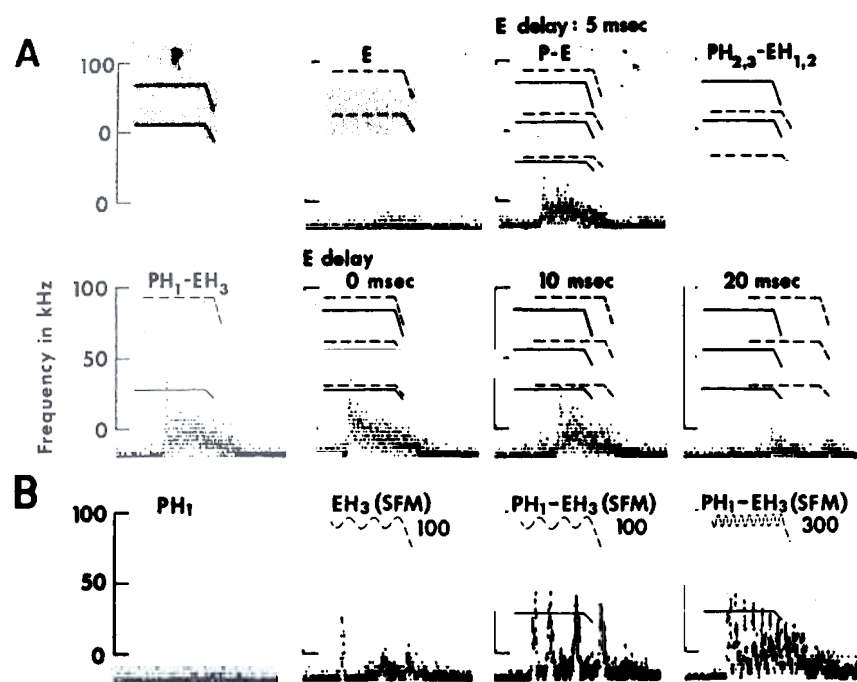


Figure 11. *A:* PST histograms displaying responses of a CF₁/CF₃ neuron to an orientation sound (pulse, P; solid lines), echo (E; dashed lines), P-E pairs, and different combinations of elements in the P-E pair. Each PST histogram is shown together with the schematized sonogram of the delivered stimulus. The durations of the CF and FM components are 30 msec and 4 msec, respectively. Facilitation of response occurs to P-E, PH₁-EH₃, or PCF₁-ECF₃ pairs, but not to P alone, E alone, or any combination of components that does not contain both PCF₁ and ECF₃. The CF₃ component must be Doppler shifted by 4.38 kHz relative to three times the CF₁ frequency to elicit the maximum facilitation in this neuron. FM components do not play a role in the facilitative response. The last three histograms show the responses to P-E pairs with different echo (E) delays of 1 msec, 10 msec, and 20 msec. (Echo delay is 5 msec for the preceding histograms in A). Facilitation is clear only when the echo delay is less than 20 msec. The latency of the response follows the shifting of echo relative to pulse. The frequencies and amplitudes of the essential signal elements for maximum response are 29.05 kHz and 67 dB SPL for PCF₁ and 91.53 kHz and 54 dB SPL for ECF₃. The number of stimulus presentations for each histogram is 200. (From Suga et al., 1983a.) *B:* PST histograms displaying responses of a CF₁/CF₃ neuron to sinusoidally frequency-modulated (SFM) echoes. From left to right, responses are shown to PH₁ alone, to EH₃ (with 100 Hz SFM) alone, to the PH₁-EH₃ (with 100 Hz SFM) pair, and to the PH₁-EH₃ (with 300 Hz SFM) pair. Note that the prominent phase-locked responses to the SFM are evoked only when EH₃ is paired with PH₁. Frequencies and amplitudes are 31.83 kHz and 66 dB SPL for PCF₁ and 91.19 kHz and 51 dB SPL for ECF₃. The SFM depth is 0.52%. (From Suga et al., 1983b.)

1977). In the CF/CF and DSCF areas of the auditory cortex, there are also neurons that are sharply tuned to CF₂ and respond well to the AC component with synchronous discharges. In the CF/CF area, 35% of the neurons show clear phase-locked responses to sinusoidally frequency-modulated (SFM) echoes, while 65% show no or poor phase-locked responses. The neurons showing clear phase-locked responses are interesting because they require the CF₁ of the orientation

sound to be paired with either the CF₂ or CF₃ of an echo (Figure 11B). The depth of the SFM echo is important for good phase locking. The best SFM depth varies from neuron to neuron, but the best SFM rate is similar for all CF/CF neurons. Their responses are somewhat better between 100 and 180 Hz, although they respond well to SFM echoes with rates between 70 and 300 Hz.

In the CF/CF areas, wingbeat information can be expressed by periodic changes in both the impulse discharges of single neurons and the location of neural activity along the frequency axis. However, the positional change in neural activity is complicated because (1) some neurons are insensitive to periodic frequency modulation; (2) individual neurons are sensitive to different modulation depths; and (3) frequency modulations evoked by beating wings are also associated with amplitude modulation (Suga et al., 1983b). Furthermore, the frequency modulations actually evoked by flying insects are far more complicated than the SFM echoes used in our experiments (Schnitzler and Ostwald, 1983).

As described below, target range information is eventually represented by the location of activated neurons tuned to particular ranges, that is, the time intervals between two acoustic events (Suga and O'Neill, 1979). Therefore, it is possible that wingbeat information is eventually expressed by a place mechanism, the location of neurons tuned to particular SFM depths and rates.

Representation of Target Range in the FM-FM Area. The primary cue for target ranging is the delay of an echo from an emitted orientation sound. At the periphery, the echo delay (target range) is coded by the interval between the responses of neurons to the orientation sound and echo. There are no anatomical components or locations representing range information. In the FM-FM area, however, most neurons respond poorly or not at all to the orientation sound and echo presented separately but respond vigorously to orientation sound-echo pairs with specific echo delays between 0.4 and 18 msec (Figures 12, 13). They are sensitive to particular target ranges between 7 and 310 cm (Suga and O'Neill, 1979; O'Neill and Suga, 1979, 1982).

There are two classes of range-sensitive neurons, tracking and range-tuned. In tracking neurons, the best echo delay (best delay), that is, the best target range (best range), for response to an echo following the orientation sound becomes shorter and delay tuning becomes narrower when the repetition rate of paired sounds increases and the duration of individual sounds becomes shorter, as in the biosonar signals emitted during target-directed flight. Because of these changes in response, tracking neurons are not suited to processing target range information. But they are suited to tracking a target and can respond continuously to an echo from an approaching target while simultaneously rejecting echoes from more distant objects that are also approaching but are not the targets of primary interest. This type of neuron is rare in the FM-FM area.

In range-tuned neurons, the best delay is constant regardless of the repetition rate and duration of a pair of sounds. Consequently, they respond to the target only when it is within a certain narrow fixed range (Figure 13). Range-tuned neurons are specialized for processing echoes from targets at particular ranges. They are predominant in the FM-FM area and are arranged systematically according to their best ranges. Thus they form a neural axis representing target ranges of 7-310 cm (Figure 14). This is called odotopic representation. Target ranges between

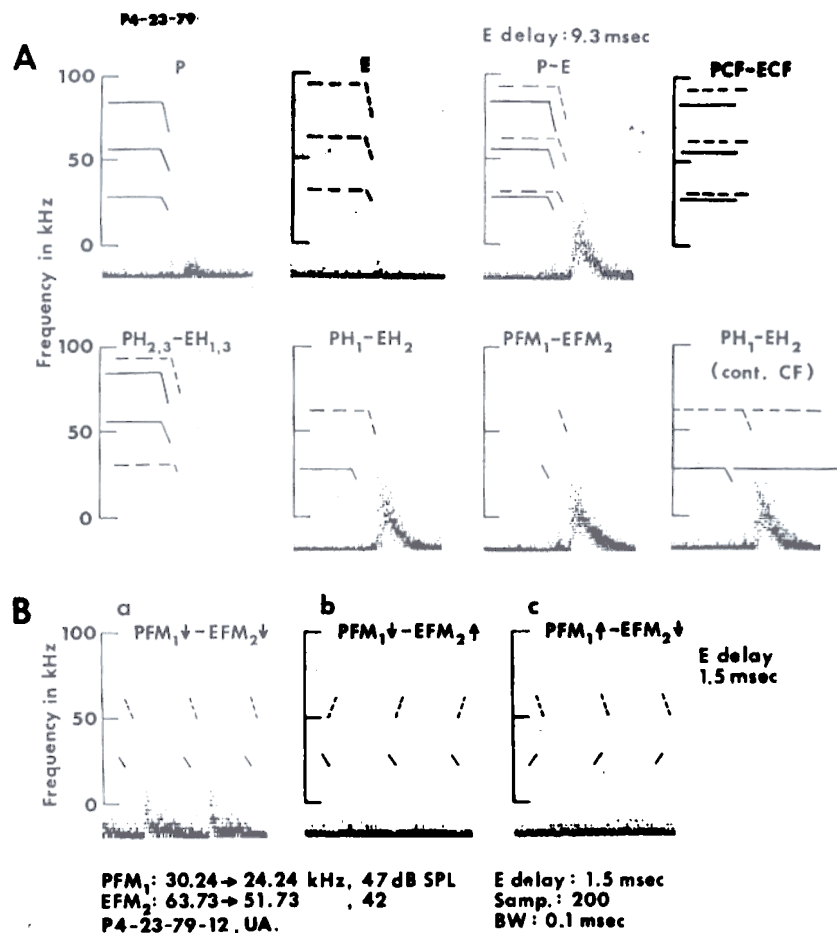


Figure 12. Response properties of FM₁-FM₂ neurons. **A:** PST histograms displaying the responses of an FM₁-FM₂ neuron to the orientation sound (pulse, P), echo (E), P-E pair, and various combinations of elements in the P-E pair. Above each PST histogram is the schematized sonogram of the delivered stimulus. The CF and FM components are 30 and 4 msec long, respectively. The essential components for facilitation are PFM₁ and EFM₂. Hence the neuron shows facilitative responses only to pairs P-E, PH₁-EH₂, and PFM₁-EFM₂. The histogram at the bottom right shows that the facilitation evoked by the PFM₁-EFM₂ pair is not masked by two continuous CF tones. The frequency sweep and amplitude are 31.2-25.2 kHz and 56 dB SPL for PFM₁ and 62.3-50.3 kHz and 38 dB SPL for EFM₂. The echo (E) delay from P is 9.3 msec, which is the best delay for this neuron. The repetition rate of the stimulus pairs is 5/sec. The number of stimulus presentations for each PST is 200. **B:** PST histograms displaying the responses of an FM₁-FM₂ neuron to pairs of FM sounds with different sweep directions. There is a strong facilitative response to pairs of PFM₁-EFM₂ when both FMs sweep down from high to low frequency, as shown in the schematized sonogram (B_a). However, there is no response when the stimulus consists of a downward-sweeping PFM₁ and an upward-sweeping EFM₂ (B_b). Note that the frequencies swept are the same as shown in B_a. In B_c there is no response when the PFM₁ sweeps upward and the EFM₂ sweeps downward. Frequency sweep and amplitude are 30.2-24.2 kHz and 47 dB SPL for PFM₁ and 63.7-51.7 kHz and 42 dB SPL for EFM₂. EFM₂ is delayed from PFM₁ by 1.5 msec, which is the best delay for this neuron. The repetition rate of the stimuli is 40/sec. The number of stimulus presentations for each PST is 200. (From Suga et al., 1983a.)

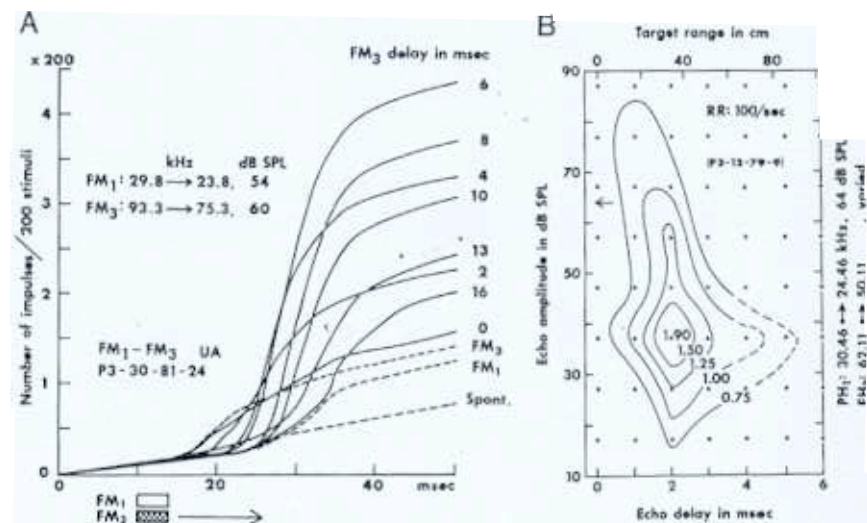


Figure 13. **A:** Cumulative histograms of responses of an FM₁-FM₃ neuron to demonstrate the importance of the delay of an echo from the pulse. The essential elements in a P-E pair for facilitation of this neuron are PFM₁ and EFM₃. The EFM₃ delay from PFM₁ is varied from 0 to 16 msec. The open rectangle at the bottom represents the actual time course of PFM₁ relative to the histograms. Maximum facilitation occurs for a 6-msec EFM₃ delay (best delay). This facilitative response is much larger than the sum of the responses to either PFM₁ or EFM₃ alone. Note the significant (6 msec) difference in latency between responses to PFM₁ and EFM₃ delivered alone (dashed lines). Spont., background discharges when no stimulus was delivered. PFM₁ and EFM₃ were fixed at the best FM sweep and best amplitude for facilitation of this neuron. The best FM sweep for facilitation was 29.8-23.8 kHz for PFM₁ and 93.3-75.3 kHz for EFM₃, which was 3.9 kHz higher than the exact third harmonic of the PFM₁. Both the best delay and the deviation of the best EFM₃ from the exact third harmonic of the best PFM₁ indicate that this neuron is not sensitive to the combination of FM₁ and FM₃ within either P or E, but to PFM₁ and EFM₃. The best amplitude for facilitation was 54 dB SPL for PFM₁ and 60 dB SPL for EFM₃. (From Suga et al., 1983a.) **B:** Iso-impulse count contours representing the response magnitude of an FM₁-FM₂ neuron plotted on the coordinates of echo amplitude versus echo delay (i.e., target subtended angle vs. target range). This neuron is tuned to a target located 34 cm in front of the bat (2-msec echo delay) that returns an echo of 37 dB SPL to the bat's ears. Since the neuron was tuned to a target at a short distance, the P-E pair for this plot was delivered at a rate of 100/sec, mimicking the terminal phase of target-directed flight. The neuron strongly responded to each paired stimulus. The essential components for excitation were PFM₁ and EFM₂ in the P-E pair. PH₁ consisted of a 5-msec CF component at 30.5 kHz followed by a 2-msec FM component sweeping from 30.5 to 24.5 kHz at an amplitude of 64 dB SPL. EH₂ consisted of a 5-msec CF component at 62.1 kHz followed by a 2-msec FM component sweeping from 62.1 to 50.1 kHz at an amplitude that was varied in increments of 10 dB. The maximum response was evoked when the echo H₂ was Doppler shifted by 1.1 kHz. This Doppler shift could be generated by the relative motion of a target approaching at a speed of 3.1 m/sec. Individual dots indicate where an average number of impulses per paired stimulus was obtained by presenting the identical paired stimulus 200 times. The contour lines are drawn on the basis of these data points. The dashed parts of the contour lines indicate where the following of the response to each paired stimulus was not reliably observed because of background noise associated with animal movement. (From Suga and O'Neill, 1979, by permission of the AAAS.)

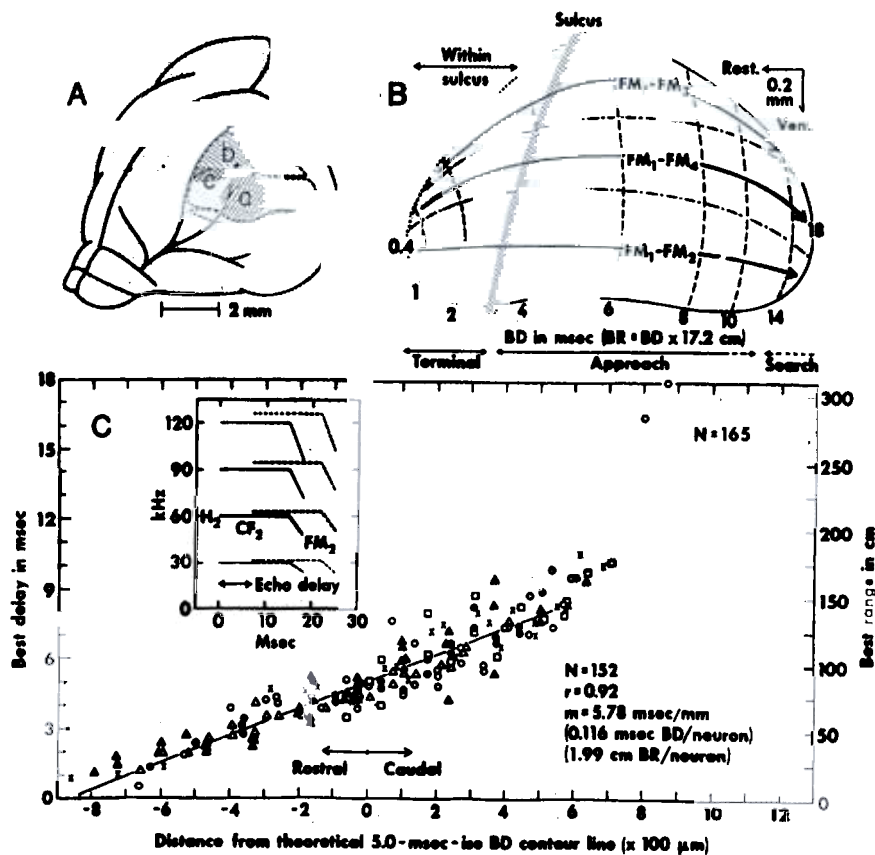


Figure 14. Odotopic representation in the FM-FM area. **A:** Left cerebral hemisphere of the mustached bat. The large auditory cortex (dotted lines) contains at least three areas specialized for processing biosonar information: a, DSCF; b, FM-FM; c, CF/CF. The branched lines are arteries. There is a sulcus below the largest branch. **B:** The FM-FM area consists of three major clusters of delay-tuned neurons: FM₁-FM₂, FM₁-FM₃, and FM₁-FM₄ neurons. Each cluster shows odotopic representation. Iso-best delay contours and range axes are schematically shown by dashed lines and solid arrows, respectively. Best delays (BDs) of 0.4 and 18 msec correspond to best ranges (BRs) of 7 and 310 cm, respectively. Range information in the search, approach, and terminal phases of echolocation is represented by neural activity at different locations in the FM-FM area. **C:** Relationship between BD and distance along the cortical surface. The data were obtained from six cerebral hemispheres, indicated by six different symbols. The regression line represents an average change in BD with distance along the cortical surface. The slope of the regression line corresponds to 0.116 msec BD per neuron, that is, 1.99 cm BR per neuron. Note that the echo-delay axis of the FM-FM area is used for representing the time interval between two acoustic events. All the data were obtained from unanesthetized bats. (From Suga and O'Neill, 1979, by permission of the AAAS.)

50 and 140 cm are overrepresented in the FM-FM area, and the best range varies at a rate of 2 cm per neuron along the range axis. Neurons tuned to distances of more than 200 cm are relatively few and appear to take a limited role in ranging because of their broad delay tuning curves. Several species of bats show the first sign of the approach phase of echolocation at a distance of about 2.3 m for a wire obstacle of 0.3 cm diameter (Grinnell and Griffin, 1958) and can discriminate among distance differences of 1.2–2.5 cm (Simmons, 1971). Our neurophysiological data thus show an interesting correlation with behavioral data.

About 25% of FM-FM neurons are tuned to echo delays shorter than 4 msec (69 cm target range) and respond vigorously to each of such orientation sound-echo pairs delivered at a rate of 100/sec. Therefore, the auditory cortex is involved in information processing even during the terminal phase of echolocation.

For the excitation of FM-FM neurons, the essential elements in the orientation sound-echo pairs are the first harmonic FM component (FM₁) in the orientation sound and one or more higher harmonic FM components (FM₂₋₄) in the echo. CF components have no significant effect on the excitation of FM-FM neurons (Figure 12). Neurons responding to FM₁-FM₃, FM₁-FM₄, or FM₁-FM₂ pairs form three separate clusters that are arranged in a dorsal-to-ventral fashion in almost all brains studied. Odotopic representation has been demonstrated in each cluster.

FM-FM neurons are "monaural" coincident detectors. When FM-FM neurons respond to an FM₁ and FM_n ($n = 2, 3, \text{ or } 4$) delivered alone, although the responses are usually poor, the latency of the response to the FM₁ is longer than that to the FM_n (Figures 12A, 13A). In each neuron, the difference in latency is the same as the best delay. This indicates that facilitation is evoked when the response to the FM_n is made to coincide with that to the FM₁ by delivering the FM_n with a certain delay from the FM₁ (Suga et al., 1983a; Olsen, 1986). The neural basis for the monaural coincident detectors has been speculated to be multiple delay lines and/or rebound from inhibition (Sullivan, 1982a,b; Olsen, 1986).

Since the response properties of FM-FM neurons are very important with regard to the overall neural mechanisms for processing complex sounds, their response properties in the frequency and amplitude domains are summarized below. Most FM-FM neurons are optimally excited by the FM_n of a positively Doppler-shifted echo following the FM₁ of the orientation sound but display broad frequency tuning to FM sounds. Therefore, these neurons have poor resolution of target velocity information. Some FM-FM neurons are excited only when the FM signals sweep downward, as in natural orientation sounds (Figure 12B; Suga et al., 1983a). The response properties of these neurons suggest that their presynaptic neurons are "FM specialized" (Suga, 1965a,b, 1968, 1969; O'Neill, 1985).

The delay tuning curves of most FM-FM neurons are elliptical because they do not show a facilitation of response to intense echoes (Figure 13B). They are best facilitated by particular amplitude relationships between the two FM sounds, so it is possible that they play an important role in target characterization. The FM₁ best amplitudes for facilitation are clustered at and around 63 dB SPL, whereas those of FM₂, FM₃, and FM₄ show large variations among neurons. In

the great majority of neurons, facilitation causes a large decrease in threshold up to 89 dB, which corresponds to an increase in sensitivity of up to 28,000-fold (Suga et al., 1983a).

Since FM-FM neurons are tuned not only to a specific echo delay but also to a particular echo amplitude (Figure 13B), they respond best to targets with particular cross-sectional areas at particular distances. Furthermore, FM-FM neurons respond best when echoes are Doppler shifted by approaching targets, as has been described. The response properties of these neurons are, as a result, quite complex (Suga et al., 1983b). FM₁-FM₄ neurons are theoretically better suited for the fine characterization of small targets than are FM₁-FM₂ neurons, because FM₄ has a much shorter wavelength and a broader bandwidth than FM₂. Therefore, it has been suggested that the distribution of neural activity perpendicular to the range axis within the cortical plane represents the amplitude spectra of all the FM components of an echo, that is, target characteristics such as shape and size.

Information Flow Beyond the FM-FM and CF/CF Areas. The FM-FM area projects to the DF and VA areas of the cerebrum as well as to other regions of the brain (Figure 7). The DF area consists of three clusters of FM-FM neurons. In each cluster, a range (echo-delay) axis is formed, along which target ranges of 7–140 cm are systematically represented (Suga and Horikawa, 1986). The DF area thus represents only the shorter half of the ranges represented in the FM-FM area. The DF area projects to the VF area, as well as to other areas in the cerebrum to which the FM-FM area does not project. The VF area also consists of three clusters of FM-FM neurons and appears to represent target ranges between 7 and 80 cm, the shorter half of the ranges represented in the DF area (Edamatsu et al., 1988). The functional significance of these multiple time axes in echolocation is not yet known. One may hypothesize that these three different areas are related to echolocation behavior at different distances to targets. The VA area contains combination-sensitive neurons that are different from FM-FM and CF/CF neurons. They show facilitative responses to the CF₂ and FM₂ of an echo when these are combined with the FM₁ of the orientation sound (Tsuzuki and Suga, 1988).

Auditory information is sent from the auditory cortex not only to the association cortex but also to the motor system. Both the FM-FM and CF/CF areas project to the pontine motor nuclei, which in turn project to the cerebellum. In the cerebellar vermis, there are tiny clusters of FM-FM and CF/CF neurons. In addition to these, there is a cluster of noise burst-sensitive neurons. In different cerebellar lobules there is a representation of the different harmonics of the orientation sound, but there is no systematic representation of frequency in each lobule (Horikawa and Suga, 1986). Biosonar information is also sent to the vocal system. Some neurons in the periaqueductal gray, for instance, become active prior to vocalization and respond to acoustic stimuli delivered from a loudspeaker (Suga and Yajima, 1988).

The projections of the CF/CF area do not overlap those of the FM-FM area. All the data thus far obtained indicate that complex acoustic signals are processed in a parallel hierarchical way in the ascending auditory system and beyond the auditory cortex.

Important Principles for the Processing of Biologically Important Sounds. The data obtained from the auditory system of the mustached bat indicate not only the specialization of the bat's auditory system for echolocation, but also the neural mechanisms that are shared or probably shared with different types of animals. In the following, these mechanisms are listed as principles for the processing of biologically important sounds.

1. The peripheral auditory system has evolved not only for the detection of biologically important sounds but also for frequency analysis of these sounds according to species-specific requirements. Therefore, the quality factor of a frequency tuning curve can be higher for peripheral neurons tuned to frequencies of sounds most important to the species.

2. Frequency tuning of central neurons can be sharpened by lateral inhibition, which eliminates mainly the "skirt" of a frequency tuning curve. The more important the frequency analysis of a particular component of sounds, the more pronounced is the neural sharpening for neurons tuned to that component.

3. The central auditory system can create neural filters tuned to IBPs other than frequency. These IBP filters can also be sharpened by lateral inhibition. The IBP filters act as cross-correlators, which correlate incoming signals with their filter properties, that is, stored information.

4. Complex sounds can be processed by IBP filters tuned to different combinations of signal elements.

5. Different types of IBP filters are aggregated separately at particular locations of the central auditory system. In other words, the system contains functional subdivisions specialized for processing different types of auditory information essential to a species.

6. In each aggregate or subdivision, IBP filters are arranged along axes for the systematic representation of IBPs, that is, variation of biologically important signals.

7. Each axis—or population of neurons—representing an IBP is apportioned according to the biological importance of the IBP.

8. The functional organization of the auditory system can be quite different among different species, reflecting differences in the properties of their acoustic signals and/or species-specific auditory behavior. Organization can also be different among individuals within the same species when their biologically important acoustic signals are different in property.

The data obtained from the barn owl (Konishi et al., this volume), for example, indicate that neural mechanisms for sound localization are based upon principles 3, 5, 6, 7, and 8. The processing of visual information is also based upon most of these principles.

Remarks. Processing of target velocity and range information by combination-sensitive neurons is the best example of parallel hierarchical processing of complex acoustic signals. The signal processing beyond the auditory cortex is anatomically parallel hierarchical. Therefore, one important topic to be studied is the upper limit of the hierarchical component in signal processing, that is, the upper limit

of specialization of single neurons. It has long been questioned whether a pontifical neuron or cardinal neurons are the eventual outcome of such parallel hierarchical processing. What we have found thus far are several groups of combination-sensitive neurons that are tuned to particular values of IBPs characterizing combinations of two signal elements. They are obviously neither cardinal nor pontifical neurons. Higher-order combination-sensitive neurons that are tuned to combinations of more than two signal elements have not yet been found, although the auditory system has a capacity to produce them quite easily. Therefore, a hypothesis based upon the data available is that, as the IBP filter hypothesis says, recognition of an overall target image is somehow directly related to a spatiotemporal pattern of neural activity over the several areas specialized for systematic representation of IBPs.

In the auditory cortex of the mustached bat, different kinds of IBPs characterizing complex acoustic signals are separately and systematically represented. The size of neural representations or maps is apportioned according to the importance of the IBPs, as well as the importance of individual values of an IBP. This does not mean that auditory information can be processed only with a neural map. A neural map representing an acoustic parameter other than frequency is probably formed for the following three reasons: (1) a variation in a parameter is biologically important; (2) a mechanism to produce an array of neurons tuned to different values of IBP operates systematically according to anatomical locations; and (3) an IBP map is advantageous for the sensorimotor interface (Suga et al., 1983b). Any parameter characterizing an acoustic signal shows some variation. If the variation is biologically very important, the auditory system may develop a subdivision for systematic representation of the variation (IBP map). If the variation is not biologically important, there may be no map, even if there are neurons tuned to different values of the parameter. The best example related to this problem is amplitopic representation, which has so far been found only in the DSCF area of the auditory cortex of the mustached bat (Figure 7). In the mustached bat, the echo amplitude carries the information about the subtended angle of a target, so that a variation in amplitude itself is biologically important. In cats and monkeys, neurons tuned to particular amplitudes of tone bursts have been found (Brugge and Merzenich, 1973; Phillips and Orman, 1984), but amplitopic representation has not yet been found. In these animals, a variation in amplitude is not in itself an IBP for communication, so that there may be no particular necessity to express it separately from other acoustic parameters.

It is worth listing the functional organizations beyond tonotopic representation that have been found in the auditory system of different species of animals: amplitopic representation (Suga, 1977; Suga and Manabe, 1982), odotopic (echo-delay) representation (Suga and O'Neill, 1979; O'Neill and Suga, 1982; Horikawa and Suga, 1986), Doppler-shift axis or frequency-versus-frequency coordinates (Suga et al., 1981, 1983a), azimuth representation (Kujirai and Suga, 1983), and binaural bands (Manabe et al., 1978), found in the mustached bat; binaural bands (Imig and Adrian, 1977), auditory space map (Middlebrooks and Knudsen, 1984), and representation of AM rate (Langner, 1985; Schreiner and Langner, this volume), found in the cat; auditory space map, found in the barn owl (Knudsen and Konishi, 1978) and guinea pig (King and Palmer, 1983).

NEURAL MECHANISMS FOR SPEECH-SOUND PROCESSING (SPECULATION)

The synchronization hypothesis of speech-sound processing by the auditory cortex has the serious limitation that only low-frequency information such as pitch can be expressed by synchronized discharges at cortical neurons. The amplitude spectrum hypothesis appears to be too simple to explain neural mechanisms for speech-sound processing. The frequency-versus-amplitude coordinates found in the DSCF area of the primary auditory cortex of the mustached bat fit this hypothesis, but they may be unique to that species. The detector hypothesis is attractive, because the auditory system can produce neurons tuned to particular combinations of IBPs. But the neurons found thus far are not specialized enough to be called detectors. Therefore, these three hypotheses will not be discussed.

The IBP filter hypothesis has been supported by the data obtained from different types of animals. Combination-sensitive neurons found in the mustached bat are tuned to particular values of IBPs characterizing the combinations of signal elements and are arranged in specific areas of the auditory cortex for systematic representation of biologically important signal variation (Suga and O'Neill, 1979; Suga et al., 1981, 1983a; O'Neill and Suga, 1982). Complex-sound processing by combination-sensitive neurons has been demonstrated not only in mammals (bats) but also in amphibians (Fuzessery and Feng, 1983) and avians (Margoliash, 1983). It may be one of the mechanisms of complex-sound processing in many different species of animals. Therefore, a neural mechanism for speech-sound processing may be hypothesized from these animal data.

In the dominant hemisphere of the human brain, there appear to be several areas or subdivisions related to language processing. Each area appears to have a particular function different from those of adjacent areas. For example, in a bilingual person who is competent in English and Greek, the two languages appear to be processed in separate areas (Ojemann, 1983). The parceling of each function may also be one of the bases of complex-sound processing. The functional organization of each area of speech-sound processing is probably more elaborate than can be deduced from the animal data. The animal data, therefore, suggest only a part of the basic mechanisms for the processing of speech sounds.

Processing of Formant Combinations (Vowels) and Fills (Consonants)

According to a simplified scenario, vowels are recognized by combinations of F_1 , F_2 , and F_3 . The formant frequency varies among speakers to some extent. This variation is one of the cues for identification of speakers, so that it is biologically important. Vowels are expressed by identifiable loci on the coordinates of F_1 -versus- F_2 frequencies and F_1 -versus- F_3 frequencies (Figure 3C). Does the human auditory system have such coordinates, enabling it to represent vowels by the loci of activated neurons? Nothing is known about the functional organization of the human auditory system, but it has been demonstrated that the auditory system of the mustached bat has such frequency-versus-frequency coordinates (Figures 7B, 10A). Therefore, I shall apply bat data to vowel processing here and assume that CF components CF_1 , CF_2 , and CF_3 are equivalent to formants

F_1 , F_2 , and F_3 , respectively. In terms of complex-sound processing by combination-sensitive neurons, it is not at all important whether these CF signals are harmonically related or not or whether their frequencies are ultrasonic or not, because the question is whether the auditory system contains neurons examining or tuned to particular combinations of two activated locations along a tonotopic axis.

At the periphery there is a large filter bank containing three arrays or groups of neurons, each tuned to F_1 , F_2 , or F_3 . The sharpness of their frequency tuning can differ according to the importance of the signal elements. For simplicity, Figure 15B shows the tuning curves of neurons *a* and *b* that are tuned to F_1 or F_2 . Since the frequencies of F_1 , F_2 , and F_3 differ among speakers, neuron *b* is tuned to a frequency slightly higher than the average F_2 frequency. The widths of these curves are broader at higher stimulus levels, so that the ambiguity in coding of stimulus frequency is very large when speech sounds are loud.

In the central auditory system, this ambiguity is reduced in some neurons by lateral inhibition, so that these neurons have sharp level-tolerant tuning. The extent of sharpening can differ according to the importance of the signal elements. The more extensive the sharpening, the smaller the difference between the BFs for excitation and inhibition. It should be noted that this lateral inhibition hardly sharpens the tip portion of the tuning curve but sharpens its skirt. Three arrays of neurons with sharp level-tolerant tuning are produced for fine frequency analysis of F_1 , F_2 , and F_3 in the central auditory system (Figure 15C).

For vowel coding, an important problem must be considered—the saturation in discharge rate of peripheral neurons at high stimulus levels. When vowels are not intense, the distribution of discharge rates over an array of neurons with different BFs expresses formants appropriately. But when the vowels are intense, this distribution does not express the formants because of saturation. On the other hand, a distribution of strength of phase-locked responses over the array of the neurons appropriately expresses the formants regardless of stimulus levels (Young and Sachs, 1979). In other words, the phase-lock code incorporated with the place code is far superior to the rate code incorporated with the place code.

In contrast to peripheral neurons, neurons in the primary auditory cortex are so limited in phase locking that they cannot express the formants by phase-locked discharges. For the cortical representation of the vowels, therefore, the phase-lock code must be translated into place code at a subcortical auditory nucleus. For the processing of target range information in the mustached bat, this translation occurs in the medial geniculate body. Therefore, a similar neural mechanism for such translation for the processing of vowels may be considered.

In Figure 15E, each phase-locked neuron sends impulses to a coincidence-sensitive neuron acting as an AND gate through two axonal branches. Branch 2 (dashed line), which acts as a delay line, is longer and/or thinner than branch 1 (solid line). It may send impulses to the coincidence-sensitive neuron through an interneuron or interneurons. Each delay line is adjusted to evoke the delay of an impulse which is the same as the period of the BF of each phase-locked neuron. For example, a phase-locked neuron tuned to a 1-kHz sound would have a 1-msec delay line.

When impulses of branches 1 and 2 arrive simultaneously at the coincidence-sensitive neuron, it is excited. The magnitude of excitation of coincidence-sensitive

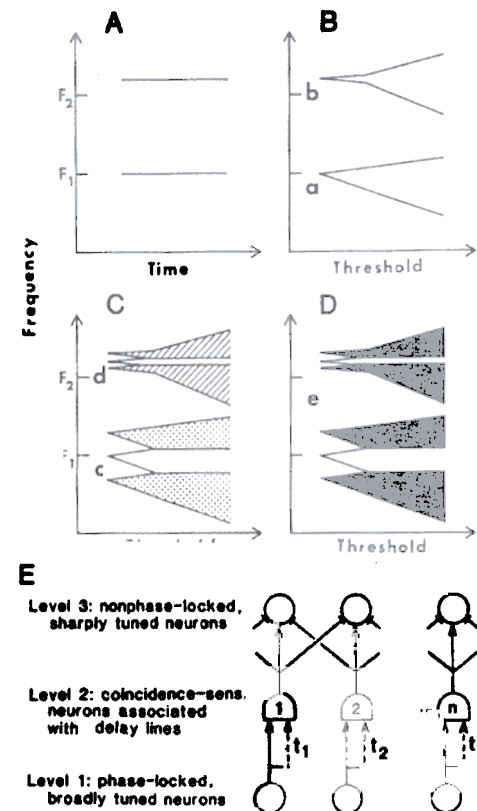


Figure 15. Parallel hierarchical processing of a complex sound consisting of two components. **A:** Schematized sonogram showing F_1 and F_2 components. **B:** Excitatory areas (tuning curves) of the peripheral neurons *a* and *b*. **C:** Level-tolerant excitatory areas of central neurons *c* and *d*. These areas are sandwiched between inhibitory areas (stippled or shaded). **D:** Facilitative areas of an F/F combination-sensitive neuron. These areas are also sandwiched between inhibitory (filled) areas. **E:** Translation of phase-lock code into place code by the circuits consisting of an array of coincidence-sensitive neurons associated with delay lines. **Level 1:** An array of phase-locked neurons with different best frequencies. The axons of these neurons form delay lines (dashed lines t_1 – t_n). t_n is the same as the period of the best frequency of neuron *n*. **Level 2:** An array of coincidence-sensitive neurons (1–*n*). **Level 3:** Tuning curves of non-phase-locked neurons are sharpened by lateral inhibition.

neurons depends upon both the extent and rate of coincidence. Their impulses are not necessarily phase-locked with stimulus waves. There is an array of coincidence-sensitive neurons associated with different delay lines. A neuron associated with a 1-msec delay line, for example, is tuned to a 1-msec interimpulse interval, that is, a 1-kHz sound. Such a neural mechanism allows the appropriate encoding of formants even at high pressure levels. The frequency tuning curves of the coincidence-sensitive neurons are not sharp but may be sharpened by lateral inhibition.

For further processing of vowels, the relationships among formants are examined by combination-sensitive neurons. In the medial geniculate body the outputs of level-tolerant neurons are integrated, and two types of combination-sensitive neurons are produced: F_1/F_2 and F_1/F_3 . These neurons act as an AND gate. Figure 15D shows the facilitative tuning curves (areas) of an F_1/F_2 neuron. These areas are sandwiched between the large inhibitory tuning curves (areas), so that this neuron can be excited only by sounds with sharp spectral peaks at two specific frequencies. It cannot be excited by a noise burst, which simultaneously and equally stimulates both the facilitative and inhibitory areas.

F_1/F_2 and F_1/F_3 neurons project separately to the auditory cortex and form two subdivisions: F_1/F_2 and F_1/F_3 areas. In each subdivision, there are frequency-versus-frequency coordinates for the systematic representation of the relationship in frequency between two formants. Therefore, the F/F area has the functional organization shown in Figure 3C.

Formant frequencies and other parameters differ among speakers and are important in identifying both speakers and vowels. Therefore, the difference in formant frequency is a biologically important variation. The frequency-versus-frequency coordinates are suited for the systematic representation of this biologically important variation. The finer analysis is required for the combination of formant frequencies that occur more frequently than the others, so that these combinations are overrepresented in large areas within the F/F area. On the other hand, the combinations of formant frequencies that rarely occur are underrepresented. Accordingly, the frequency axes of the F/F area are nonlinear and different from those in Figure 3C.

The frequency ratios between the formants of a vowel are nearly constant across speakers (Miller, 1984). Thus one may consider that a vowel is eventually processed by neurons tuned to a particular frequency ratio. Such neurons, if any, are not suitable for speaker identification, so that the auditory system must contain two groups of neurons suited for either speaker or vowel recognition. If the F/F area is assumed to be a neuronal tissue, the activity of which is somehow directly related to vowel recognition, a spatial pattern of neural activity in this area would be directly related to the recognition of both vowels and speakers. (For example, a category /i/ is represented by any spatial pattern of neural activity that occurs within a particular area, and a difference in /i/ between speakers is represented by the spatial pattern between neural activities that occurs within this area.) The formant ratios are expressed by neural activity along oblique lines in the frequency-versus-frequency coordinates in Figure 3C.

A neural network model for the integration of F_1/F_2 and F_1/F_3 neurons to produce higher-order combination-sensitive neurons that selectively respond to a combination of F_1 , F_2 , and F_3 is conceivable. We do not yet know whether or not such hierarchical signal processing takes place in the brain of animals.

In the discussion above, the role of phase-locked discharges in vowel processing was ignored because both the population of phase-locking neurons and the degree of phase-locking become smaller at levels of the auditory system beyond the colliculus, and because the mechanisms by which cortical neurons eventually encode "beautiful" phase-locked discharges occurring at the periphery remain to be studied. Phase-locked discharges are important for pitch perception. In the inferior colliculus an interesting neural mechanism for periodicity pitch has been found (Langner, 1985).

Since 1960, neurons that do not respond to pure tones but to noise bursts and/or clicks have repeatedly been reported. However, the response properties of these neurons (hereafter called NB-sensitive neurons) have not been studied well because their tuning curves to noise bursts, as functions of bandwidth and center frequency, have not been measured. Without such essential measurements, we are not sure how selective they are in their response to noise bursts. Based upon limited data (Suga, 1969), I had speculated that the auditory system contains NB-sensitive neurons. Recently, a cluster of NB-sensitive neurons was found in the vermis of the cerebellum of the mustached bat (Horikawa and Suga, 1986). Since their tuning curves to noise bursts are broad, a neural mechanism to produce NB sensitivity may be facilitation due to simultaneous excitation of presynaptic neurons with different BFs by a noise burst.

NB-sensitive neurons may be suited to respond to fills, so that we may hypothesize that the auditory system contains a cluster of NB-sensitive neurons (NB area) to process fills and that the NB area is organized in coordinates representing center frequencies-versus-bandwidths of noise bursts.

Processing of Transitions (Phoneme Combinations)

For the recognition of many nonvowel phonemes and combinations of phonemes (e.g., words), transitions (FM components) are very important. When a transition is added to the F_2 of /a/, for instance, the sound is perceived as /pa/, /ta/, or /ka/, depending on the properties of the transition. When a second transition is added to the F_1 of /a/, it is recognized as /ba/, /da/, or /ga/ (Figure 3D). This example demonstrates that combinations of transitions and, more generally, combinations of information-bearing elements are very important for speech recognition. Does the human auditory system have neurons that selectively respond to transitions as well as neurons tuned to particular combinations of transitions? We cannot answer this question, but we do know that the auditory systems of the mustached bat and the little brown bat contain FM-sensitive neurons (Suga, 1965a,b, 1968, 1969; O'Neill, 1985) and several aggregates of neurons tuned to particular combinations of FM components (O'Neill and Suga, 1979, 1982; Suga and O'Neill, 1979; Suga et al., 1983a). The functional organization of the areas containing these FM-FM combination-sensitive neurons cannot be directly applied to the processing of transitions, but it suggests an interesting mechanism for them.

Before we hypothesize neural mechanisms for the processing of transitions based upon the animal data, we have to consider what makes speech so unique in terms of acoustic pattern. Its uniqueness comes from an enormous number of different combinations of phonemes, that is, an enormous number of combinations of transitions. Therefore, the area devoted to processing of combinations of transitions, if any, may consist of many arrays of transition (T)-sensitive neurons, and its functional organization must be unique and complex.

At the auditory periphery, neurons respond to any of the formants, transitions, and fills that stimulate their excitatory response areas. In the central auditory system, T-sensitive neurons that are tuned to particular transitions are produced by a neural circuit mediating disinhibition, so that a particular portion of the IC contains this type of neuron. This region differs from that in which sharpening of frequency tuning by inhibition takes place for formant analysis.

In a particular region of the medial geniculate body, T-sensitive neurons are integrated so that the geniculate neurons in this region are tuned to particular combinations of two transitions that differ in the range and direction of frequency sweep. These neurons project to a particular part of the auditory cortex—the T/T area. The T/T area is much larger and is organized in a much more complex way than the F/F area. The T/T area consists of two major subdivisions, T_1/T_2 and T_1/T_3 , which are specialized to examine the combinations of T_1 with T_2 or T_3 . The functional organization of these subdivisions is likely to be much more complicated than that of the FM-FM area found in the mustached bat. It may be further speculated that the outputs of T_1/T_2 and T_1/T_3 neurons are integrated by neurons located in another area of the cerebrum. To date, there are no neurophysiological data to support such a speculation.

Speech Processing

Speech consists of trains of sounds. Many different phonemes are combined in different sequences. Are there neurons that respond selectively to a long sequence of sounds in the auditory and/or vocal systems of animals and humans? Combination-sensitive neurons tuned to particular sequences of two sounds or signal elements have been found in bats (Suga et al., 1978, 1983a; O'Neill and Suga, 1979, 1982; Suga and O'Neill, 1979; Sullivan, 1982a,b; Suga and Horikawa, 1986) and also in the white-crowned sparrow (Margoliash, 1983). In these animals, the sequence of sounds is nearly fixed. But in speech the sequence of sounds varies enormously, so it is unlikely that combination-sensitive neurons are produced that can examine all possible combinations of two adjacent sounds in speech. It is, however, possible that such neurons interact to examine the most common sequences.

As previously described, the neurophysiological data thus far available indicate that the upper limit of specialization of neural circuits is the production of arrays of neurons tuned to particular combinations of two signal elements. Therefore, one may hypothesize that speech recognition is somehow based upon a spatio-temporal pattern of neural activity that occurs in the F/F, NB, and T/T areas.

According to one hypothesis, speech sounds are represented by a spatiotemporal pattern of neural activity occurring in several cortical areas, each consisting of specialized neurons tuned to IBPs. There are no data from animals for any speculation of speech processing beyond this. I expect that innovative data on the neural mechanisms for speech processing will be obtained by recording neural activity from speech areas and also by electrically stimulating these areas, as Ojemann (1983) and his coworkers have been doing.

A human hand is quite different from, for instance, a batwing in shape and function. However, they share the same set of bones. Like the set of bones, the neural mechanisms listed above are probably shared by humans and other mammals. The extremely sharp, level-tolerant, frequency tuning of neurons in the mustached bat and the sharp tuning to interaural cues of neurons found in the barn owl indicate the uniqueness of their auditory systems. This uniqueness is, however, produced by lateral inhibition, which probably operates optimally in these species. In other words, the uniqueness is based upon the specialization of shared mechanisms. One important problem is to determine the extent to which speech-sound processing is based upon the specialization of shared mech-

anisms, or whether it involves mechanisms that are so unique that it is not shared at all by other animals.

ACKNOWLEDGMENTS

The work on the mustached bat was supported by a research grant from the Public Health Service, R01-NS17338 (Jacob Javits Neuroscience Investigator Award). I thank Dr. S. J. Gaioni for his comments.

REFERENCES

- Abeles, M., and M. H. Goldstein, Jr. (1970) Functional architecture in cat primary auditory cortex: Columnar organization and organization according to depth. *J. Neurophysiol.* 33:172–187.
- Aitkin, L. M., W. R. Webster, J. L. Veale, and D. C. Crosby (1975) Inferior colliculus. I. Comparison of response properties of neurons in central, pericentral, and external nuclei of adult cat. *J. Neurophysiol.* 38:1196–1207.
- Asanuma, A., D. Wong, and N. Suga (1983) Frequency and amplitude representations in anterior primary auditory cortex of the mustached bat. *J. Neurophysiol.* 50:1182–1196.
- Brugge, J. F., and M. M. Merzenich (1973) Patterns of activity of single neurons of the auditory cortex in monkey. In *Basic Mechanisms of Hearing*, A. R. Møller, ed., pp. 745–766, Academic, New York.
- Bruns, V. (1976) Peripheral auditory tuning for fine frequency analysis by the CF-FM bat, *Rhinolophus ferrumequinum*. I. Mechanical specialization of the cochlea. *J. Comp. Physiol.* 106:77–86.
- Capranica, R. R. (1977) Auditory processing of vocal signals in anurans. In *The Reproductive Biology of Amphibians*, D. H. Taylor and S. I. Guttman, eds., pp. 337–355, Plenum, New York.
- Edamatsu, H., M. Kawasaki, and N. Suga (1988) Distribution of combination-sensitive neurons in the ventral fringe area of the auditory cortex of the mustached bat. *J. Neurophysiol.* (submitted).
- Erulkar, S. D. (1959) The responses of single units of the inferior colliculus of the cat to acoustic stimulation. *Proc. R. Soc. Lond. [Biol.]* 150:336–355.
- Fenton, M. B., and J. H. Fullard (1979) The influence of moth hearing on bat echolocation strategies. *J. Comp. Physiol.* 132:77–86.
- Frishkopf, L., R. Capranica, and M. Goldstein (1968) Neural coding in the bullfrog's auditory system—A teleological approach. *Proc. IEEE* 56:969–980.
- Fuzessery, Z. M., and A. S. Feng (1983) Mating call selectivity in the thalamus and midbrain of the leopard frog (*Rana p. pipiens*): Single and multiunit analyses. *J. Comp. Physiol.* 150:333–344.
- Goldstein, M. H., Jr., M. Abeles, R. L. Daly, and J. McIntosh (1970) Functional architecture in cat primary auditory cortex: Tonotopic organization. *J. Neurophysiol.* 33:188–197.
- Grinnell, A. D. (1963) The neurophysiology of audition in bat: Intensity and frequency parameters. *J. Physiol. (Lond.)* 167:38–66.
- Grinnell, A. D. (1970) Comparative auditory neurophysiology of neo-tropical bats employing different echolocation signals. *Z. Vgl. Physiol.* 68:117–153.
- Grinnell, A. D., and D. R. Griffin (1958) The sensitivity of echolocation in bats. *Biol. Bull. (Woods Hole)* 114:10–22.
- Henson, M. M. (1973) Unusual nerve-fiber distribution in the cochlea of the bat *Pteronotus p. parnellii* (gray). *J. Acoust. Soc. Am.* 53:1739–1740.
- Hind, J. E. (1960) Unit activity in the auditory cortex. In *Neural Mechanisms of the Auditory and Vestibular Systems*, G. L. Rasmussen and W. F. Windle, eds., pp. 201–210, Charles C. Thomas, Springfield, Illinois.
- Horikawa, J., and N. Suga (1986) Biosonar signals and cerebellar auditory neurons of the mustached bat. *J. Neurophysiol.* 55:1247–1267.

- Imig, T. J., and H. O. Adrian (1977) Binaural columns in the primary field (AI) of cat auditory cortex. *Brain Res.* 138:241-257.
- Katsuki, Y., N. Suga, and Y. Kanno (1962) Neural mechanism of the peripheral and central auditory system in monkeys. *J. Acoust. Soc. Am.* 34:1396-1410.
- King, A. J., and A. R. Palmer (1983) Cells responsive to free-field auditory stimuli in guinea-pig superior colliculus: Distribution and response properties. *J. Physiol.* 342:361-381.
- Knudsen, E. I., and M. Konishi (1978) A neural map of auditory space in the owl. *Science* 200:795-797.
- Kujirai, K., and N. Suga (1983) Tonotopic representation and space map in the non-primary auditory cortex of the mustached bat. *Auris Nasus Larynx (Tokyo)* 10:9-24.
- Langner, G. (1985) Time code and periodicity pitch. In *Time Resolution in Auditory Systems*, A. Michelsen, ed., pp. 108-121, Springer-Verlag, New York.
- Liberman, M. C. (1978) Auditory-nerve response from cats raised in a low-noise chamber. *J. Acoust. Soc. Am.* 63:442-455.
- Liberman, M. C. (1982) Single-neuron labeling in the cat auditory nerve. *Science* 216:1239-1241.
- Liberman, A. M., P. C. Delattre, L. J. Gerstman, and F. S. Cooper (1956) Tempo of frequency change as a cue for distinguishing classes of speech sounds. *J. Exp. Psychol.* 52:127-137.
- Manabe, T., N. Suga, and J. Ostwald (1978) Aural representation in the Doppler-shifted CF processing area of the primary auditory cortex of the mustache bat. *Science* 200:339-342.
- Margoliash, D. (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci.* 3:1039-1057.
- Merzenich, M. M., P. L. Knight, and G. L. Roth (1975) Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38:231-249.
- Middlebrooks, J. C., and E. I. Knudsen (1984) A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.* 4:2621-2634.
- Miller, J. D. (1984) Auditory-perceptual correlates of the vowel. *J. Acoust. Soc. Am. (Suppl. 1)* 76:S79.
- Moiseff, A., and M. Konishi (1983) Binaural characteristics of units in the owl's brainstem auditory pathways: Precursors of restricted spatial receptive fields. *J. Neurosci.* 3:2553-2562.
- Ojemann, G. A. (1983) Brain organization for language from the perspective of electrical stimulation mapping. *Behav. Brain Sci.* 2:189-230.
- Olsen, J. F. (1986) *Functional Organization of the Medial Geniculate Body of the Mustached Bat*, unpublished doctoral dissertation, Washington University, St. Louis, Missouri.
- Olsen, J. F., and N. Suga (1983) Combination-sensitive neurons in the auditory thalamus of the mustached bat. *Soc. Neurosci. Abstr.* 9:768.
- O'Neill, W. E. (1985) Responses to pure tones and linear FM components of the CF-FM biosonar signals by single units in the inferior colliculus of the mustached bat. *J. Comp. Physiol.* 157:797-815.
- O'Neill, W. E., and N. Suga (1979) Target range-sensitive neurons in the auditory cortex of the mustache bat. *Science* 203:69-73.
- O'Neill, W. E., and N. Suga (1982) Encoding of target-range information and its representation in the auditory cortex of the mustached bat. *J. Neurosci.* 47:225-255.
- O'Neill, W. E., G. Schuller, and S. Radtke-Schuller (1985) Functional and anatomical similarities in the auditory cortices of the old world horseshoe bat and neotropical mustached bat for processing similar biosonar signals. *Assoc. Res. Otolaryngol. Abstr. (8th Mtg.)*, p. 148.
- Ostwald, J. (1980) The functional organization of the auditory cortex in the CF-FM bat *Rhinolophus ferrumequinum*. In *Animal Sonar Systems*, R. G. Busnel and J. F. Fish, eds., pp. 953-955, Plenum, New York.
- Peterson, G. N., and H. L. Barney (1952) Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24:175-184.
- Phillips, D. P., and S. S. Orman (1984) Responses of single neurons in posterior field of cat auditory cortex to tonal stimuli. *J. Neurophysiol.* 51:147-163.
- Pollak, G., O. W. Henson, Jr., and A. Novick (1972) Cochlear microphonic audiograms in the "pure tone" bat *Chilonycteris parnellii parnellii*. *Science* 176:66-68.

- Roeder, K. D. (1962) The behavior of free flying moths in the presence of artificial ultrasonic pulses. *Anim. Behav.* 10:300-304.
- Rose, J. E., N. Gross, C. D. Geisler, and J. E. Hind (1966) Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. *J. Neurophysiol.* 29:288-314.
- Rouiller, E., Y. de Ribaupierre, and F. de Ribaupierre (1979) Phase-locked responses to low frequency tones in the medial geniculate body. *Hear. Res.* 1:213-226.
- Schnitzler, H.-U. (1970) Echoortung bei der Fledermaus *Chilonycteris rubiginosa*. *Z. Vgl. Physiol.* 68:25-38.
- Schnitzler, H.-U., and J. Ostwald (1983) Adaptations for the detection of fluttering insects by echolocation in horseshoe bats. In *Vertebrate Neuroethology*, J. P. Ewert, R. R. Capranica, and D. J. Ingle, eds., pp. 801-827, Plenum, New York.
- Schnitzler, H.-U., N. Suga, and J. A. Simmons (1976) Peripheral auditory tuning for fine frequency analysis by the CF-FM bat, *Rhinolophus ferrumequinum*. III. Cochlear microphonics and auditory nerve responses. *J. Comp. Physiol.* 106:99-110.
- Schuller, G. (1974) The role of overlap of echo with outgoing echolocation sound in the bat *Rhinolophus ferrumequinum*. *Naturwissenschaften* 61:171-172.
- Simmons, J. A. (1971) The sonar receiver of the bat. *Ann. N.Y. Acad. Sci.* 188:161-174.
- Suga, N. (1961) Functional organization of two tympanic neurons in noctuid moths. *Jpn. J. Physiol.* 11:666-677.
- Suga, N. (1964) Single unit activity in cochlear nucleus and inferior colliculus of echolocating bats. *J. Physiol. (Lond.)* 172:449-474.
- Suga, N. (1965a) Analysis of frequency modulated sounds by neurones of echolocating bats. *J. Physiol. (Lond.)* 179:26-53.
- Suga, N. (1965b) Functional properties of auditory neurones in the cortex of echolocating bats. *J. Physiol. (Lond.)* 181:671-700.
- Suga, N. (1968) Analysis of frequency-modulated and complex sounds by single auditory neurones of bats. *J. Physiol. (Lond.)* 198:51-80.
- Suga, N. (1969) Classification of inferior collicular neurones of bats in terms of responses to pure tones, FM sounds and noise bursts. *J. Physiol. (Lond.)* 200:555-574.
- Suga, N. (1973) Feature extraction in the auditory system of bats. In *Basic Mechanisms in Hearing*, A. R. Møller, ed., pp. 675-744, Academic, New York.
- Suga, N. (1973) Feature extraction in the auditory system of bats. In *Basic Mechanisms in Hearing*, A. R. Møller, ed., pp. 675-744, Academic, New York.
- Suga, N. (1977) Amplitude-spectrum representation in the Doppler-shifted CF processing area of the auditory cortex of the mustache bat. *Science* 196:64-67.
- Suga, N. (1982) Functional organization of the auditory cortex: Representation beyond tonotopy in the bat. In *Cortical Sensory Organization*, Vol. 3, C. N. Woolsey, ed., pp. 157-218, Humana, Clifton, New Jersey.
- Suga, N. (1984) The extent to which biosonar information is represented in the bat auditory cortex. In *Dynamic Aspects of Neocortical Function*, G. M. Edelman, W. E. Gall, and W. M. Cowan, eds., pp. 315-373, Wiley, New York.
- Suga, N., and J. Horikawa (1986) Multiple time axes for representation of echo delays in the auditory cortex of the mustached bat. *J. Neurophysiol.* 55:776-805.
- Suga, N., and P. H.-S. Jen (1976) Disproportionate tonotopic representation for processing species-specific CF-FM sonar signals in the mustache bat auditory cortex. *Science* 194:542-544.
- Suga, N., and P. H.-S. Jen (1977) Further studies on the peripheral auditory system of "CF-FM" bats specialized in fine frequency analysis of Doppler-shifted echoes. *J. Exp. Biol.* 69:207-232.
- Suga, N., and T. Manabe (1982) Neural basis of amplitude-spectrum representation in the auditory cortex of the mustached bat. *J. Neurophysiol.* 47:225-255.
- Suga, N., and W. E. O'Neill (1979) Neural axis representing target range in the auditory cortex of the mustache bat. *Science* 206:351-353.

- Suga, N., and W. E. O'Neill (1980) Auditory processing of echoes: Representation of acoustic information from the environment in the bat cerebral cortex. In *Animal Sonar Systems*, R. G. Buanel and J. F. Fish, eds., pp. 589–611, Plenum, New York.
- Suga, N., and K. Tsuzuki (1985) Inhibition and level-tolerant frequency tuning in the auditory cortex of the mustached bat. *J. Neurophysiol.* 53:1109–1145.
- Suga, N., and Y. Yajima (1988) Audio-visual integration in the midbrain of the mustached bat: Periaqueductal gray and reticular formation. In *The Physiological Control of Mammalian Vocalization*, J. D. Newman, ed., Plenum, New York (in press).
- Suga, N., J. A. Simmons, and T. Shimozawa (1974) Neurophysiological studies on echolocation systems in awake bats producing CF-FM orientation sounds. *J. Exp. Biol.* 61:379–399.
- Suga, N., J. A. Simmons, and P. H.-S. Jen (1975) Peripheral specialization for fine analysis of Doppler-shifted echoes in "CF-FM" bat *Pteronotus parnellii*. *J. Exp. Biol.* 63:161–192.
- Suga, N., G. Neurweiler, and J. Möller (1976) Peripheral auditory tuning for fine frequency analysis by the CF-FM bat, *Rhinolophus ferrumequinum*. IV. Properties of peripheral auditory neurons. *J. Comp. Physiol.* 106:111–125.
- Suga, N., W. E. O'Neill, and T. Manabe (1978) Cortical neurons sensitive to particular combinations of information bearing elements of bio-sonar signals in the mustached bat. *Science* 200:778–781.
- Suga, N., W. E. O'Neill, and T. Manabe (1979) Harmonic-sensitive neurons in the auditory cortex of the mustached bat. *Science* 203:270–274.
- Suga, N., K. Kujirai, and W. E. O'Neill (1981) How biosonar information is represented in the bat cerebral cortex. In *Neuronal Mechanisms of Hearing*, J. Syka and L. Aitkin, eds., pp. 197–219, Plenum, New York.
- Suga, N., W. E. O'Neill, K. Kujirai, and T. Manabe (1983a) Specialization of "combination-sensitive" neurons for processing of complex biosonar signals in the auditory cortex of the mustached bat. *J. Neurophysiol.* 49:1573–1626.
- Suga, N., H. Niwa, and I. Taniguchi (1983b) Representation of biosonar information in the auditory cortex of the mustached bat, with emphasis on representation of target velocity information. In *Advances in Vertebrate Neuroethology*, P. Ewert, ed., pp. 829–867, Plenum, New York.
- Suga, N., H. Niwa, I. Taniguchi, and D. Margoliash (1987) The personalized auditory cortex of the mustached bat: Adaptation for echolocation. *J. Neurophysiol.* 56:643–654.
- Sullivan, W. E., III (1982a) Neural representation of target distance in auditory cortex of the echolocating bat *Myotis lucifugus*. *J. Neurophysiol.* 48:1011–1032.
- Sullivan, W. E., III (1982b) Possible neural mechanisms of target distance coding in auditory system of the echolocating bat *Myotis lucifugus*. *J. Neurophysiol.* 48:1033–1047.
- Sullivan, W. E., III, and M. Konishi (1984) Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl. *J. Neurosci.* 4:1787–1799.
- Symmes, D. (1981) On the use of natural stimuli in neurophysiological studies of audition (review). *Hear. Res.* 4:203–214.
- Tsuzuki, K., and N. Suga (1988) Combination-sensitive neurons in the ventrolateral area of the auditory cortex of the mustached bat. *J. Neurophysiol.* (submitted).
- Tunturi, A. R. (1952) A difference in the representation of auditory signal for the left and right ears in the iso-frequency contours of the right middle ectosylvian and auditory cortex of the dog. *Am. J. Physiol.* 168:712–727.
- Watanabe, T. (1959) A single neuron activity in the secondary cortical auditory area in the cat. *Jpn. J. Physiol.* 9:245–256.
- Wever, E. G. (1949) *Theory of Hearing*, Wiley, New York.
- Yin, T. C. T., and S. Kuwada (1984) Neural mechanisms of binaural interaction. In *Dynamic Aspects of Neocortical Function*, G. M. Edelman, W. E. Gall, and W. M. Cowan, eds., pp. 263–313, Wiley, New York.
- Young, E. D., and M. B. Sachs (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory nerve fibers. *J. Acoust. Soc. Am.* 66:1381–1403.

Chapter 24

Neurophysiological and Anatomical Substrates of Sound Localization in the Owl

MASAKAZU KONISHI
TERRY T. TAKAHASHI
HERMANN WAGNER
WINTHROP E. SULLIVAN
CATHERINE E. CARR

ABSTRACT

The ideas that guided the present work include the integration of behavioral, physiological, and anatomical approaches, the selection of an animal appropriate for the question to be addressed, the use of natural behavior and stimuli, and the study of successive stages of stimulus processing starting from higher-order complex neurons. The barn owl is particularly suitable for the study of sound localization because of its exquisite ability to track prey by hearing. The owl uses interaural time and intensity differences to localize sound in azimuth and elevation, respectively. The binaural time and intensity cues are processed in anatomically separate pathways that start from the cochlear nuclei.

The owl's auditory system measures interaural time difference in a noise signal by using the phase of its spectral components. The nucleus laminaris, the first site of binaural convergence in the time pathway, receives phase information in the form of phase-locked spikes from the nucleus magnocellularis. We hypothesize that the nucleus laminaris uses the principles of coincidence detection and delay lines to measure interaural time differences in different frequency bands. These mechanisms underlie the origin of neuronal selectivity for interaural time differences. When a neuron is selective for a particular time disparity independent of stimulus frequency, the neuron is said to have a characteristic delay. All binaural neurons in the time pathway presumably have a characteristic delay, although they also respond to time disparities that are removed from the characteristic delays by n periods of the neuron's best or stimulus frequency. However, at the final stage of time coding the space-specific neurons respond exclusively to the interaural time differences equivalent to their characteristic delays.

Indirect evidence supports the hypothesis that the nucleus laminaris contains a map of phase equivalents of characteristic delays. In the central nucleus of the inferior colliculus, one of the projection areas of the nucleus laminaris, the phase equivalents of characteristic delays are also mapped. Phase-sensitive neurons of the central nucleus are arranged such that a columnar array of them across isofrequency laminae represents a single characteristic delay. This column of neurons appears to project onto a space-specific neuron in the external nucleus of the inferior colliculus; the projection pattern confers upon a space-specific neuron its broad frequency tuning and ability to signal the characteristic delay unambiguously. The connectivity also explains how a nontopographic