PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Philosophy and Stimulus Design for Neuroethology of Complex-Sound Processing

Nobuo Suga

Phil. Trans. R. Soc. Lond. B 1992 **336**, 423-428
doi: 10.1098/rstb.1992.0078

Philosophy and stimulus design for neuroethology of complex-sound processing

NOBUO SUGA

Department of Biology, Washington University, Lindell-Skinker Avenue, St Louis, Missouri 63130, U.S.A.

SUMMARY

In research on the neural mechanisms for the processing of biologically important sounds such as species-specific sounds and sounds produced by prey and predators, it is necessary to study responses of central auditory neurons to biologically important sounds, information-bearing elements (IBES) in them, and tone bursts. The tone bursts or constant-frequency (CF) components can be an IBE in many species of animals. Information-bearing parameters characterizing these sounds must be systematically varied, and tuning of neurons to individual parameters must be studied. The measurement of a tuning curve must be performed not only for excitatory responses, but also for inhibitory and facilitative responses, if any. The selectivity of a neuron to a particular type of sound must be tested for whether it is level-tolerant. Responses to complex sounds can probably be explained on the basis of those to IBES and tone bursts, so that the use of the tone bursts, even though they are not IBES, is as essential as that of the biologically important sounds.

1. INTRODUCTION

The processing of communication calls is one of the major functions of the central auditory system. However, nothing interesting has been found of the central processing of calls in mammalian species, in spite of the promising preliminary research made in the early 1970s (Wollberg & Newman 1972), and elegant measurements on frequency tuning, amplitude tuning, binaural interactions, phase-locking, etc. over the past 25 years. On the other hand, noticeable progress has been made in birds (Margoliash 1983; Müller & Leppelsack 1985; Scheich *et al.* 1979) and frogs (Mudry *et al.* 1977; Fuzessery & Feng 1983). In songbirds and frogs, it has been shown that species-specific calls are processed by neurons that are tuned to particular combinations of two signal elements. This is also true for the processing of biosonar information in bats (see Suga *et al.* 1983). Although many more analytical and comparative studies are needed, it is clear that all of these three types of animals (bats, birds and frogs) share the same principle: species-specific complex sounds, either communication calls or biosonar signals, are processed by combination-sensitive neurons. In bats (see Suga *et al.* 1983) and frogs (Hall & Feng 1987), the central auditory system creates separate clusters of neurons tuned to different types of information-bearing parameters. These findings indicate that a neuroethological approach has been successful in the exploration of neural mechanisms for the processing of species-specific complex sounds. As I have written several articles reviewing our research on the mustached bat (for examples see Suga 1984, 1988, 1990), I shall review here the philosophy and stimulus design for auditory neuroethology.

2. INFORMATION-BEARING ELEMENTS AND PARAMETERS

In higher vertebrates, communication sounds are usually complex, and the amplitude spectra of these sounds commonly change with time. For example, human speech consists of various phonemes combined in different sequences and is consequently complicated. However, the spectrograms of speech sounds exhibit three basic components (information-bearing elements: IBES): constant frequency (CF), noise burst (NB), and frequency-modulated (FM) components. For example, consider the consonant-vowel syllable of figure 1. At the beginning of the syllable, there is a

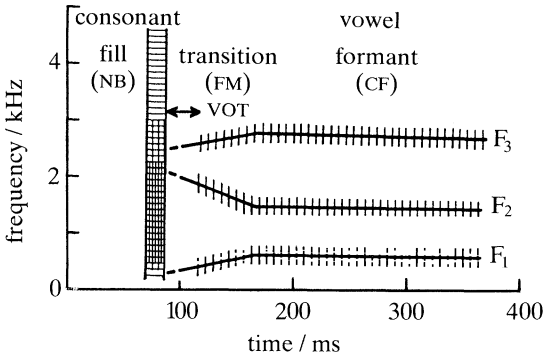

Figure 1. Information-bearing elements in human speech sounds. Schematized spectrogram of a monosyllabic sound shows four types of information-bearing elements: fill or noise burst (NB) component, transition or frequency modulated (FM) component, formant or constant frequency (CF) component, and voice 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. The short vertical bars on F_1, F_2 and F_3 indicate vocal chord activity (Suga 1984; based on Liberman *et al.* 1956).

Phil. Trans. R. Soc. Lond. B (1992) **336,** 423–428
Printed in Great Britain

vertical bar (NB component) as the air is released from the vocal tract; this corresponds to the NB portion of the consonant. The last 200 ms of the sound (the vowel portion) consists of several horizontal bars, called formants or CF components. Between the burst and vowel, there are oblique bars, called transitions or FM components (cf. Liberman *et al.* 1956).

Auditory information is carried not only by the acoustic parameters characterizing each of the above three types of IBEs, but also by parameters representing relationships among these IBEs in the frequency, amplitude and time domains. Values of a parameter comprise a continuum, but only a limited portion of the continuum is important for each species. This portion has been called the information-bearing parameter: IBP (Suga *et al.* 1981).

The three types of IBEs are also found in sounds produced by many different types of animals (Suga 1972). For example, the mustached bat, *Pteronotus parnellii*, emits complex biosonar pulses consisting of CF and FM components (figure 2*a*). These are IBEs carrying different types of biosonar information. It also emits a variety of communication sounds that differ from biosonar pulses in amplitude spectrum and duration (e.g. figure 2*b–j*). The central auditory system of different species of animals contains not only neurons selectively responding to one of the IBEs (see Suga 1969, 1973), but also neurons tuned to particular combinations of IBEs (in bats, see Suga *et al.* 1978, 1983; in birds, Margoliash 1983; in frogs, Fuzessery & Feng, 1983). In the mustached bat, arrays of combination-sensitive neurons tuned to different values of IBPs are systematically arranged in the auditory cortex and thus form computational maps (e.g. O'Neill & Suga 1982; Suga & O'Neill 1979; Suga *et al.* 1983). In the barn owl, *Tyto alba*, neurons tuned to different values of interaural time and amplitude differences are systematically arranged in the external nucleus of the inferior colliculus and thus

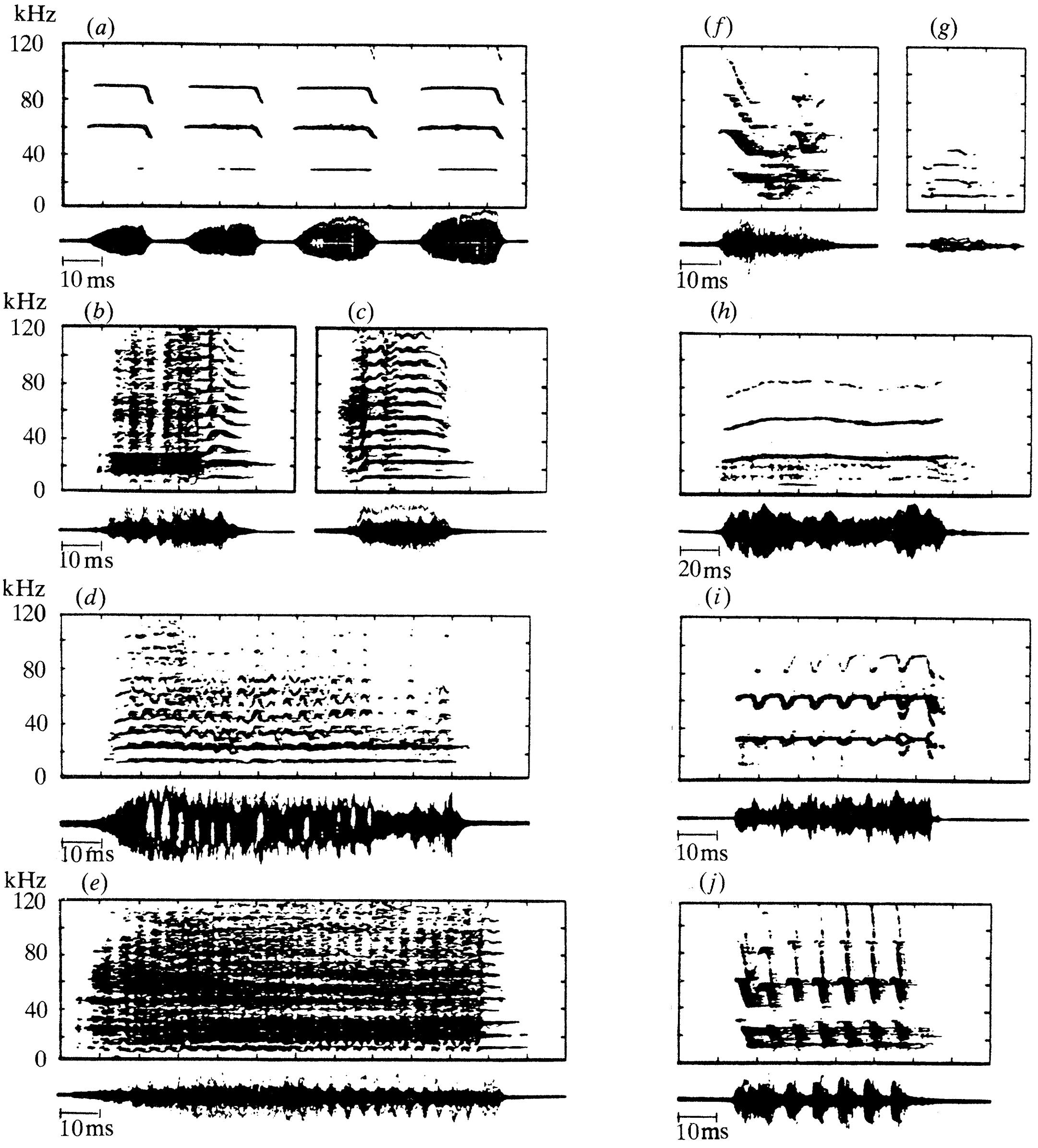

Figure 2. Biosonar pulses (*a*) and communication calls (*b–j*) of mustached bats. Note the clear differences in amplitude spectrum and temporal pattern between the pulses and the calls, except call (*j*).

form an auditory space map (Knudsen & Konishi 1978). The data obtained from the mustached bat and the barn owl clearly indicate that the central auditory system creates neurons tuned to different values of an IBP and form an IBP map. These interesting data were obtained from neuroethological studies of the auditory system.

3. PHILOSOPHY BEHIND AUDITORY NEUROETHOLOGY

Which do we want to explore, neural mechanisms for the processing of species-specific sounds or those for the processing of 'general' sounds? The auditory system has evolved for processing biologically important sounds, in particular, species-specific sounds and sounds produced by prey and predators. Therefore, it is reasonable to explore first the mechanisms for processing biologically important sounds, especially, species-specific sounds rather than those for the processing of general sounds.

In general, animals appear to have subsets of neurons specialized for processing species-specific sounds. Finding such neurons and exploring the neural mechanisms for creating them require the delivery of the species-specific sounds, individual IBEs found in them, and tone bursts. (Tone bursts can be an IBE in many species of animals.) When a neuron responds selectively to a particular complex sound, we must first confirm that no single component in the sound, whatever its amplitude, is sufficient to excite significantly the neuron, and then examine (i) which combination of two or more components is necessary for the excitation of the neuron, and (ii) how sharply the neuron is tuned to particular IBPs characterizing single IBEs or combinations of IBEs. The measurement of the filter properties of the neuron is particularly important for a quantitative description of its response properties, and also for a correlation of neural responses with biological sounds that always show some variation. Since communication, as well as biosonar, can occur over different distances, it is also important to examine whether the selectivity of the neuron to a particular sound is 'level-tolerant', i.e. does not change with stimulus level. It is particularly important to explore the neural mechanisms creating the selectivity, i.e. interaction among excitation, inhibition and facilitation.

4. STIMULUS DESIGN TO CHARACTERIZE NEURONS

In the past, species-specific calls were recorded with a tape-recorder or a computer and were played back with modifications to assess important portions of the calls to excite neurons. The modifications were made by either trimming the calls from their beginning or end, or playing them backward. Using such acoustic stimuli, some neurons were found to respond selectively to a particular 'natural' call, but not to either the reversed call or tone bursts (CF tones). Such research, however, is not analytical enough to characterize neurons. If a neuron responds only to a natural call or a particular portion of the call, its selectivity must be tested for level-tolerance; the call element or elements which excite the neuron should be identified; how sharply the neuron is tuned to a parameter or parameters characterizing the call element or elements should be measured; and the reasons why the neuron responds to call elements, but not to tone bursts must be determined. For research on the neural processing of species-specific complex calls, the calls are first categorized, and 'typical' calls are used as acoustic stimuli representing the call types. Within each call type, however, individual calls may show large variation in the frequency, amplitude and time domains. Therefore, the typical call stimuli must be varied in these three domains. To fulfill all of the above requirements, acoustic stimuli must be carefully designed.

There are two different approaches to the design of acoustic stimuli. These differ from each other with respect to which set of acoustic stimuli is to be first used, IBEs or natural calls. In one approach, IBEs and combinations of IBEs found in species-specific calls are synthesized with special-purpose instruments or a computer (figure 3), and their IBPs are systematically varied, whereas responses of single neurons are recorded. If IBEs in a call are unknown, these can be speculated by examining the spectrogram of the call. Based upon what is known about speech processing, these are usually predominant components. The IBEs speculated can then be synthesized, and their parameters can be systematically varied. Natural calls are eventually used as stimuli to examine responses of neurons.

In the other approach, typical calls are used as acoustic stimuli with or without modifications by a computer. Modifications include amplification or attenuation, filtering, shifting in frequency, cutting,

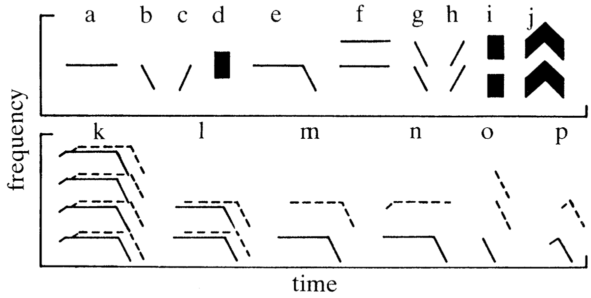

Figure 3. Schematized spectrograms of the acoustic stimuli which were used in our experiments on the biosonar of the mustached bat and the little brown bat. (*a*) CF tone (tone burst); (*b*) downward-sweeping FM sound; (*c*) upward-sweeping FM sound; (*d*) noise burst; (*e*) CF–FM sound; (*f*, *g*, *h* & *i*) correspond respectively to signals in (*a*, *b*, *c* & *d*), but with overtones; (*j*) noise burst in which the centre frequency first increases and then decreases; (*k*) a pair of a biosonar pulse (solid lines) and its Doppler-shifted echo (dashed lines); (*l*–*p*) five examples of simplified biosonar pulse (solid line) and its echo (dashed line). The delay of the echo can be varied. The duration, intensity and frequency of each signal (or component) can be varied. In FM signals, frequency can be swept linearly, exponentially, sinusoidally or trapezoidally with time. The range of the frequency sweep can also be changed.

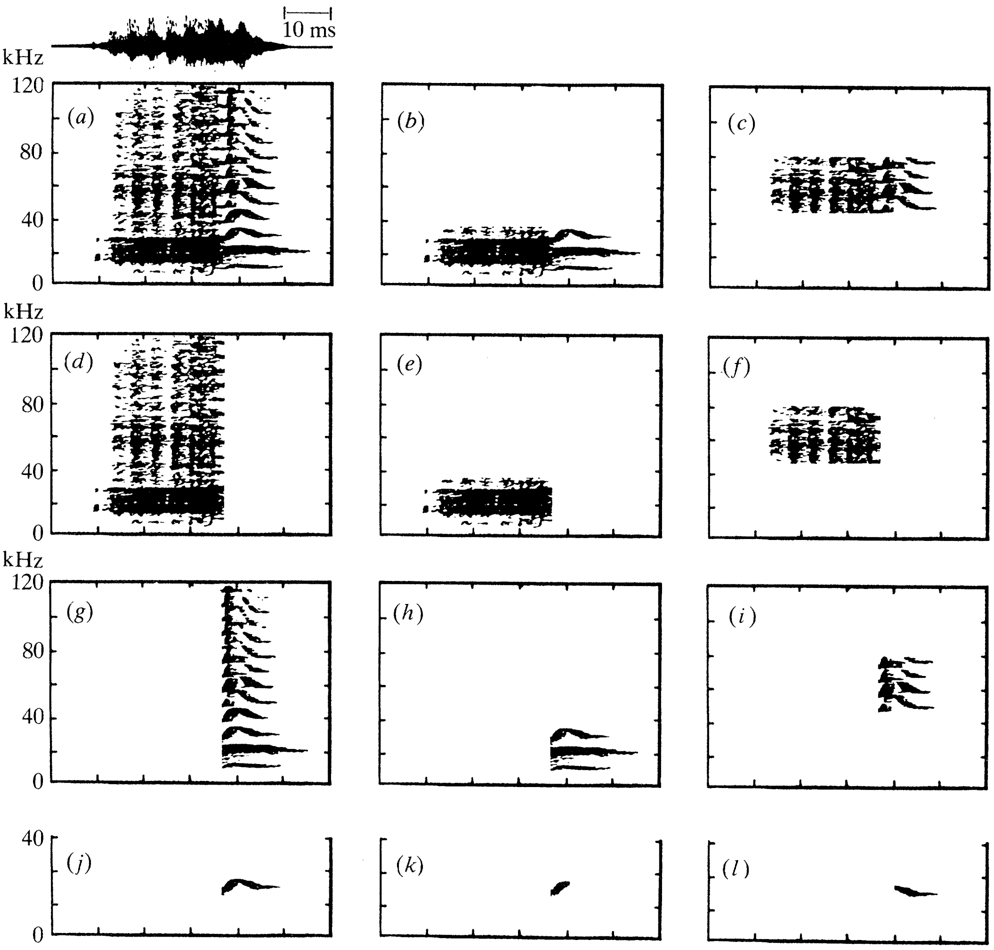

Figure 4. Modifications of a communication call to examine which part of the call is essential to excite a 'specialized' neuron. The trace at the top left shows the envelope of the call in (*a*). The normal call (*a*) is modified in the frequency domain (*b* or *c*) or time domain (*d* & *g*) or frequency-time domains (*e*, *f*, *h* & *i*). If the component shown in (*j*) excites the neuron, it will be further tested whether the neuron is sensitive to an upward-sweep (*k*) or downward-sweep (l) or their combination (*j*). If the neuron is sensitive to the combination, the tuning curve of the neuron will be measured as a function of each parameter characterizing the upward-sweep, downward-sweep, or the combination of these.

pasting, etc. For studies of single neurons, quick modifications of acoustic stimuli are required, so that filtering of the calls in the frequency and time domains can, respectively, be performed with a programmable filter which has a steep slope (e.g. 115 dB per octave) and an electronic switch which can independently control the delay, duration and rise-decay time (figure 4). Alternatively, a large file of modified acoustic stimuli can be stored in a computer. After identifying the essential signal elements for excitation of a neuron, the elements can be synthesized, and their parameters can be systematically varied.

The first approach has been taken in the studies on the auditory system of the mustached bat (e.g. Olsen & Suga 1991*a*, *b*; O'Neill & Suga 1982; Suga & Tsuzuki 1985; Suga *et al.* 1978, 1983), white-crowned sparrow (Margoliash 1983), and frog (Fuzessery & Feng 1982, 1983). All these studies have demonstrated that species-specific sounds or sequenced sounds are processed by 'combination-sensitive' neurons, and that they are clustered within certain portions of the brain.

5. EXPLORATION OF NEURAL MECHANISMS FOR PROCESSING COMPLEX SOUNDS

In the past, experiments on the auditory system of avians and mammals, communication calls of a species were recorded with a tape recorder or a computer and were played back to members of that species at, say, 60 dB SPL. Responses of single auditory neurons to the calls were compared with those to tone bursts (CF tones) at 60 dB SPL. Then, some neurons were found to be excited only by a particular type of call, but to be inhibited by the tone bursts. Therefore, it was concluded that these neurons were specialized to respond to the call, and that the responses to the call could not be explained by the responses to the tone bursts. These conclusions, however, are unacceptable because of the following two major reasons: (i) the 60 dB call (broad band signal) and 60 dB tone burst (narrow band signal) are the same in intensity, but the tone burst is much stronger than the call component alone corresponding to the tone burst (figure 5*a*); and (ii) the response properties of many central

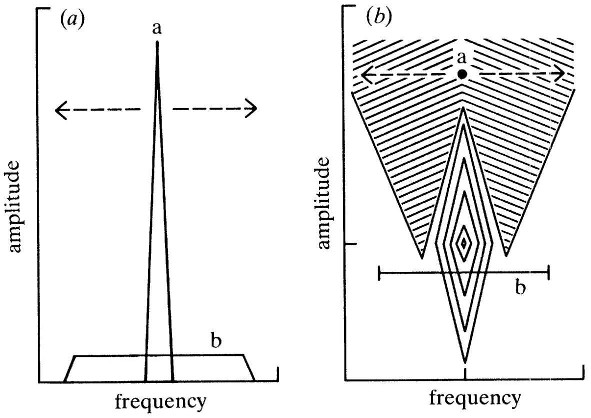

Figure 5. Amplitude spectra of two acoustic stimuli (*a*) and a response area of a central auditory neuron (*b*). (*a*) A tone burst (a) and a broadband signal (b) are the same in intensity. The frequency of the tone burst can be varied as indicated by the dashed arrows. (*b*) The excitatory and inhibitory areas of the neuron are indicated by the largest diamond and a shaded area, respectively. The four diamonds within the excitatory area indicate the iso-impulse-count contour lines. The neuron is tuned in frequency and amplitude (based upon Suga & Manabe (1982)). The neuron shows only an inhibitory response to the tone burst shown in (*a*), regardless of its frequency, because it stimulates only the inhibitory area. However, it shows excitatory response to the broadband signal shown in (*a*), because it stimulates only the excitatory area.

auditory neurons are extremely nonlinear (figure 5*b*). Therefore, it is incorrect to compare the response to the 60 dB call with that to the 60 dB tone burst. For example, some central neurons are tuned to a combination of a particular frequency and a particular amplitude and show 'upper-threshold' (in frogs, Potter 1965; in bats, Grinnell 1963; Suga 1977; Suga & Manabe 1982; in cats, Phillips & Orman 1984). 'Upper-threshold' neurons show inhibitory responses to strong tone bursts, but excitatory responses to weak tone bursts (figure 5*b*; in frogs, Fuzessery & Feng 1982; in bats, Suga 1965; Suga & Manabe 1982). Then, the response properties of these neurons characterized with tone bursts can be used to explain the observation that the 60 dB tone burst (a in figure 5*b*) does not excite but inhibits the neuron, whereas the 60 dB call (b in figure 5*b*) excites the neuron. The characterization of response properties of central auditory neurons only at a particular stimulus level is misleading, so that their response properties must be examined at different stimulus levels.

There are other examples which do not favour the statement that the responses to complex calls cannot be explained on the basis of responses to tone bursts. FM specialized neurons are inhibited by tone bursts, but excited by FM sounds which sweep across the inhibitory area. The response properties of these neurons can be explained by a disinhibition model, based upon responses to tone bursts (Suga 1965, 1968, 1969). Some cortical CF/CF and FM–FM combination-sensitive neurons do not show excitatory response to any single tone burst and FM sound, but show excitatory responses to particular combinations of a pulse and an echo (see, for example, Suga *et al.* 1983). The response properties of these CF/CF and FM–FM neurons can be explained by facilitation of responses taking place in the medial geniculate body, in other words, by responses to tone bursts (Olsen & Suga 1991*a*, *b*).

All peripheral neurons, without exception, respond to CF tones, FM sounds and noise bursts. Response properties of central auditory neurons are based upon excitatory, inhibitory and facilitatory interactions in the frequency, amplitude and time domains among neurons receiving signals directly or indirectly from these peripheral neurons. Therefore, the neural mechanisms for creating the responses of neurons selective to particular types of calls can be explored by step-by-step analysis of synaptic interactions occurring at different levels of the auditory system with the three types of IBEs (CF tones, FM sounds and noise bursts) which are commonly found in animal sounds.

To explore the neural mechanisms for the creation of 'specialized' neurons that respond selectively or preferentially to particular types of complex sounds, their response properties, in particular, tuning curves, must be extensively studied. Because responses of these neurons probably consist of excitation, inhibition and facilitation which vary in frequency, amplitude and time, the properties of these synaptic events must be studied in these three domains. Excitatory responses and tuning curves are studied with single IBEs (e.g. tone bursts). Inhibitory responses and tuning curves are studied with single, double or triple IBEs (e.g. Suga 1965; Suga & Tsuzuki 1985). When the background activity of a neuron is inhibited by single sounds, inhibition is studied with the single sounds. When it is absent or is not inhibited by single sounds, inhibition is studied with paired sounds. One sound is delivered before, without overlap with, a second sound which excites the neuron. An inhibitory tuning curve is measured as the frequency and amplitude ranges of the first sound that causes inhibition of the response to the second. The timecourse of the inhibition is measured by changing the time relationship between the two sounds. In combination-sensitive or call-sensitive neurons, the second sound in the pair is a 'combination' sound or a call that excites the neuron. Facilitative responses or tuning curves are studied with a pair of sounds (e.g. Suga *et al.* 1978, 1983). One sound in the pair is fixed in frequency and amplitude for optimum facilitation, while the other is varied in these parameters to measure a facilitative tuning curve, or one sound is delayed from the other to study the timecourse of facilitation. When a neuron is multi-combination-sensitive or call-sensitive, one of the paired sounds is a fixed complex sound or a simplified call essential for evoking facilitation, while the other is varied in frequency, amplitude and time.

Complete studies of response properties of single neurons require stable recording of action potentials at least over a few hours. Since stable intracellular recording over the long time period is hard, extracellular recording is preferred to intracellular recording. The characterization of single neurons in extracellular recording tells us the end result of neural interactions taking place at all levels interposed between peripheral neurons and a given central neuron. It does not tell us where and how particular neural interactions take place. Therefore, a top-down study must be performed step-by-step. Differences in response properties between cortical areas or subcortical nuclei and the knowledge of anatomical connections give us important insights into the neural mechanisms for processing complex sounds within and across cortical areas and sub-cortical nuclei.

For further exploration of the neural mechanisms for processing complex sounds or creating specialized neurons, synaptic mechanisms occurring in each step must be explored by micro-iontophoretic injections of various drugs with multi-barrelled electrodes. In such studies, response properties of a single neuron are first studied as described above, and drugs such as synaptic transmitters, their agonists, and their antagonists are injected. Response properties of the neuron are re-studied during and after injections. The changes in response properties evoked by drugs are evaluated in relation to working hypotheses.

Our research on the bat's auditory system has been supported by the research grants from NIDCD (DC00175), ONR (N00014-90-J-1068) and the McKnight Foundation. I thank J. S. Kanwal and K. K. Ohlemiller for their comments on the present article.

REFERENCES

Fuzessery, Z.M. & Feng, A.S. 1982 Frequency selectivity in the anuran auditory midbrain: single unit responses to single and multiple tone stimulation. *J. comp. Physiol.* A **146**, 471–484.

Fuzessery, Z.M. & Feng, A.S. 1983 Mating call selectivity in the thalamus and midbrain of the leopard frog *Rana p. pipiens*: single and multiunit analysis. *J. comp. Physiol.* **150**, 333–344.

Grinnell, A.D. 1963 The neurophysiology of audition in bat: intensity and frequency parameters. *J. Physiol., Lond.* **167**, 38–66.

Hall, J.C. & Feng, A.S. 1987 Evidence for parallel processing in the frog's auditory thalamus. *J. comp. Neurol.* A **258**, 407–419.

Knudsen, E.I. & Konishi, M. 1978 A neural map of auditory space in the owl. *Science, Wash.* **200**, 795–797.

Liberman, A.M., Delattre, P.C., Gerstman, L.J. & Cooper, F.S. 1956 Tempo of frequency change as a cue for distinguishing classes of speech sounds. *J. exp. Psychol.* **52**, 127–137.

Margoliash, D. 1983 Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci.* **3**, 1039–1057.

Mudry, K.M., Constantin-Paton, M. & Caprinica, R.R. 1977 Auditory sensitivity of the diencephalon of the leopard frog *Rana p. pipiens*. *J. comp. Physiol.* **114**, 1–13.

Müller, C.M. & Leppelsack, H.J. 1985 Feature extraction and tonotopic organization in the avian auditory forebrain. *Expl Brain Res.* **59**, 589–599.

Olsen, J.F. & Suga, N. 1991*a* Combination-sensitive neurons in the medial geniculate body of the mustached bat: encoding of relative velocity information. *J. Neurophysiol.* **65**, 1254–1274.

Olsen, J.F. & Suga, N. 1991*b* Combination-sensitive neurons in the medial geniculate body of the mustached bat: Encoding of target range information. *J. Neurophysiol.* **65**, 1275–1296.

O'Neill, W.E. & Suga, N. 1982 Encoding of target-range information and its representation in the auditory cortex of the mustached bat. *J. Neurosci.* **47**, 225–255.

Phillips, D.P. & Orman, S.S. 1984 Responses of single neurons in posterior field of cat auditory cortex to tonal stimuli. *J. Neurophysiol.* **51**, 147–163.

Potter, D. 1965 Patterns of acoustically evoked discharges of neurons in the mesencephalon of the bullfrog. *J. Neurophysiol.* **28**, 1155–1184.

Scheich, H., Langner, G. & Bonke, D. 1979 Responsiveness of units in the auditory neostriatum of the guinea fowl *Humida meleagris* to species-specific calls and synthetic stimuli. II: Discrimination of iambus-like calls. *J. comp. Physiol.* **132**, 257–276.

Suga, N. 1965 Functional properties of auditory neurons 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 neurons of bats. *J. Physiol., Lond.* **198**, 51–80.

Suga, N. 1969 Classification of inferior collicular neurons of bats in terms of responses to pure tones, FM sounds and noise bursts. *J. Physiol., Lond.* **200**, 555–574.

Suga, N. 1972 Analysis of information-bearing elements in complex sounds by auditory neurons of bats. *Audiology* **11**, 58–72.

Suga, N. 1973 Feature extraction in the auditory system of bats. In *Basic mechanisms in hearing*. (ed. A. R. Möller), pp. 675–744. New York: Academic Press.

Suga, N. 1977 Amplitude-spectrum representation in the Doppler-shifted-CF processing area of the auditory cortex of the mustache bat. *Science, Wash.* **196**, 64–67.

Suga, N. 1984 The extent to which biosonar information is represented in the bat auditory cortex. In *Dynamic aspects of neocortical function*. (ed. G. M. Edelman, W. E. Gall & W. M. Cowan), pp. 315–373. New York: John Wiley & Sons.

Suga, N. 1988 Auditory neuroethology and speech processing: complex sound processing by combination-sensitive neurons. In *Functions of the auditory system* (ed. G. M. Edelman, W. E. Gall & W. M. Cowan), pp. 679–720. New York: John Wiley & Sons.

Suga, N. 1990 Cortical computational maps for auditory imaging. *Neural Networks* **3**, 3–21.

Suga, N., Kujirai, K. & O'Neill, W.E. 1981 How biosonar information is represented in the bat cerebral cortex. In *Neuronal mechanisms of hearing* (ed. J. Syka & L. Aitkin), pp. 197–219. New York: Plenum.

Suga, N. & Manabe, T. 1982 Neural basis of amplitude-spectrum representation in the auditory cortex of the mustached bat. *J. Neurophysiol.* **47**, 225–255.

Suga, N. & O'Neill, W.E. 1979 Neural axis representing target range in the auditory cortex of the mustached bat. *Science Wash.* **206**, 351–353.

Suga, N., O'Neill, W.E., Kujirai, K. & Manabe, T. 1983 Specificity of combination-sensitive neurons for processing of complex biosonar signals in the auditory cortex of the mustached bat. *J. Neurophysiol.* **49**, 1573–1626.

Suga, N., O'Neill, W.E. & Manabe, T. 1978 Cortical neurons sensitive to particular combination of information-bearing elements of biosonar signals in the mustached bat. *Science, Wash.* **200**, 778–781.

Suga, N. & Tsuzuki, K. 1985 Inhibition and level-tolerant frequency tuning in the auditory cortex of the mustached bat. *J. Neurophysiol.* **53**, 1109–1145.

Wollberg, Z. & Newman, J.D. 1972 Auditory cortex of the squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science Wash.* **175**, 212–214.