SONGBIRDS, GRANDMOTHERS, TEMPLATES:
A Neuroethological Approach

Thesis by
Daniel Margoliash

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1984

(Submitted August 4, 1983)

to my family

ACKNOWLEDGEMENT

It is a great pleasure that I am given this opportunity to formally express my gratitude to Dr. Mark Konishi for his support and encouragement these past years. In terms of this thesis, and more broadly for my intellectual growth at Caltech, Mark has been a consistent source of invaluable contributions. The enthusiasm with which Mark approaches his life's work, and the depth and breadth of his intellectual pursuits, are standards we can all profitably attempt to adopt. He truly is my mentor, and for this I am most humbly grateful.

Dr. Derek Fender has served as my official thesis advisor for most of these years. It was on his suggestion that I pursued working in Beckman Laboratories. At all junctures in my career, Dr. Fender has willingly provided helpful advice and support. For this I thank him, as well as for his considerable attention to my thesis preparation. Dr. David Van Essen continues to be an excellent source of critical assessment, and I enjoyed wide-ranging discussions with Dr. Jim Hudspeth. In my role as a student, I would like to make special note of the excellence of teaching conducted by Drs. Hudspeth and Van Essen.

I have been at Caltech too long to possibly thank all the people who have contributed to my scientific endeavors. Certainly Drs. Erik Knudsen and Andy Moiseff have had important influences on me. I also wish to acknowledge Dale Knudson, my teacher of the art of computer programming. It has been a pleasure to learn from and work with Gene Akutagawa. My good friend Michael Patrick Walsh sweated out those early years with me as we struggled to build the synthesizer. He is one of a number of excellent Biology staff personnel whose efforts largely freed me to pursue my research. They include Bill Lease, Mike Miranda, and last but certainly not least, Nancy Gill. I thank them all for their goodwill and considerable patience.

Having spent 13 years at Caltech, I can justifiably claim some insight into

the quality of life here. I am happy to report my alma mater healthy, the Caltech community vibrant, and the bureaucracy small. It is my sincere wish that this happy state of affairs continues. I am grateful for the support of the Jean Wiegle fund, which enabled me to prepare this document on the Computing Center VAX/VMS system. The experiments described herein were supported by the Pew Memorial Trust and a National Institute of Health Training Grant (1 T32 GM07737).

ABSTRACT

Songbirds such as the white-crowned sparrow memorize the song of conspecific adults during a critical period early in life, and later in life develop song by utilizing auditory feedback. Neurons in one of the telencephalic nuclei controlling song have recently been shown to respond to acoustic stimuli. I investigated the auditory response properties of units in this nucleus using a technique that permitted great flexibility in manipulating complex stimuli such as song. A few of the units exhibited considerable selectivity for the individual's own song. In wild-caught birds, song specific units exhibited intra-dialect selectivity. In those birds that sang abnormal songs due to laboratory manipulation of song exposure during the critical period for song learning, units were selective for the abnormal songs. By systematic modification of a song, and by construction of complex synthetic sounds mimicking song, the acoustic parameters responsible for the selectivity were identified. Song specific units responded to sequences of two song parts, but not to the parts in isolation. Modification of the frequencies of either part of the sequence, or increasing the interval between the parts, varied the strength of the response. Thus, temporal as well as spectral parameters were important for the response. When sequences of synthetic sounds mimicking song were effective in evoking an excitatory response, the response was sensitive to the aforementioned manipulations. With these techniques it was possible to elucidate the acoustic parameters required to excite song specific units. All songs of the repertoire eliciting a strong excitatory response contained the appropriate parameters, which were missing from all weakly effective, ineffective, or inhibitory songs. These observations suggest that the ontogenetic modification of integrative neural mechanisms underlying song learning or song crystalization is reflected at the level of single neurons.

Table of Contents

and your bird can sing

John Lennon

Introduction

A Logical Basis for Single-Cell Neurophysiology

A fundamental problem in the study of integrative brain mechanisms is the question of the codes used by the nervous system. Behavioral phenomena such as learning, memory, and complex sensory perceptions may manifest themselves as particular states of activity in large aggregates of neurons, the activity of individual cells representing only limited perceptual information. This may be termed the population hypothesis. It contends that the neural mechanisms underlying various perceptual phenomena cannot be elucidated by recordings from individual neurons because the information is encoded only in the activity of the ensemble of neurons. Alternatively, these phenomena may be encoded in the brain by the activity of small, possibly redundant populations of relatively specialized neurons. This is called the single cell hypothesis. The hypothesis predicts that the neural bases of perceptual phenomena can be elucidated with single cell recording techniques. It does not imply that every perception has its corresponding single, highly complex neuron. As it is entirely conceivable that both types of organization coexist within the same neural system (e.g. Selverston and Miller, 1982), it is unfortunate that arguments are often couched in terms of population ***versus*** single cell hypotheses. Nevertheless, increasingly it is becoming common and even popular to despair with the use of the single cell paradigm for elucidating integrative neural processes.

In this chapter, I will discuss emerging principles for the study of the integrative action of the brain. I will show that ethological studies are a powerful tool to guide neurophysiological investigations. Within the context of behavior, I then will critically examine the theoretical and experimental evidence supporting the single unit doctrine, and in the process I will dispute the arguements against it. By discussing the logic of behavior and single cell neurophysiology, I hope to provide a rationale for single unit investigations of complex behavioral phenomena while specifically setting the stage for my own experiments.

The Neuroethological Approach

Behavior. In designing experimental paradigms, sensory neurophysiologists are faced with the critical choice of stimulus repertoire. Neuronal response properties are typically highly nonlinear; no mathematical formulation has been devised that adequately predicts responses to novel stimuli. A logically based and mathematically rigorous resolution of this problem is to test neurons with a stimulus that embodies all possible parameter combinations. Unfortunately, the parameter space over which neuronal responses can vary is truly enormous. It is an untenable proposition to systematically and thoroughly explore such a parameter space with classical sensory stimuli such as fragments of visual images or various spectral and temporal combinations of sound. Furthermore, in the absence of a theoretical basis, no classical exploration of a single neuron's response properties can justifiably conclude that additional or alternate properties have been rigorously excluded.

White noise is one of a class of random signals that parsimoniously spans an arbitrary parameter space. As such, it is potentially an attractive and powerful stimulus for neurophysiological investigations. The input/output relationship of those systems that satisfy certain requirements, including analyticity and stationarity, can be completely described by the associated Weiner kernels (Weiner, 1958; Lee and Schetzen, 1965), even if the system is highly non-linear. Many central nervous system neurons, however, do not exhibit the continuous or smooth input/output relationships optimal for the white noise technique. For example, in the bullfrog, *Rana pipens*, Frishkopf and Goldstein (1963) described a population of auditory fibers that responded only to some sound frequencies, and were inhibited by broad-band signals such as white noise. This type of non-analyticity, which may prevent the analysis from converging, is common throughout the auditory systems of virtually all animals studied to date. Similarly, it is well established that neuronal response properties can be dramatically altered by the state of attention (Mountcastle, 1976,1978). Thus, neurons often exhibit clear non-stationarities. As a result of these complexities, in many central nervous system single cell recordings white noise analysis cannot be advantageously applied (Poggio, 1981).

In lieu of a mathematically justifiable approach, an alternate logic is required to provide a theoretical basis for neurophysiological investigations. The study of natural animal behavior, ethology, provides a conceptual framework from which to develop such a neuroethological theory (Bullock, 1961; Hoyle, 1976; Konishi, 1971). The underlying tenets are:

1) Behavior is a quantifiable phenomenon.
2) Each species has a unique perceptual world.
3) The brain controls behavior.
4) Neural processing reflects an animal's behavioral repertoire.

The following discussion concentrates on those historical developments that have influenced ethology, and on ethology's relationship to modern neurophysiological investigations.

Behavior is adaptive. The behavioral repertoire of an animal is uniquely adapted to its ecological niche. Thus, each species has its own unique perceptual realm (von Uexkull, 1909, sited in Konishi, 1978), and a unique set of motor programs. It follows that the student of behavior must develop observational techniques sensitive to these evolutionary influences. These conclusions provide the underpinnings for the central organizing principle of classical ethology, that naturalistic methods are the essential starting point in the study of behavior. Although this suggests the importance of fieldwork, it should not be misconstrued that ethology is limited to fieldwork. Since field observations may suffer from uncontrollable and unmeasurable factors, it is widely accepted today that many behavioral observations require clarification in a controlled environment. The advantage accrued by field observations is precisely that they suggest specific laboratory experiments. Behavioral demonstration of microsecond sensitivity achieved by electric fish, barn owls, and bats (Langner and Scheich, 1978; Moiseff and Konishi, 1981; Simmons, 1979) are impressive examples of a precision in behavior uncovered by modern ethological investigations.

A number of organizing principles for sensory, motor, motivational and ontogenetic components of behavior have been elucidated by ethologists. Animals often recognize biologically relevant sign stimuli by attending to one or a few specific parameters within the stimulus. Sign stimuli evoke stereotypic responses (fixed action patterns - FAP) by a postulated mechanism termed the Innate Releasing Mechanism (Lorenz, 1935, sited in Tinbergen, 1951). After release, the response typically goes to completion (Tinbergen, 1951). The drive or motivational state (Lorenz, 1950) is also an important element in organizing behavior. Substantial drive for an activity, blocked by prolonged absence of the associated releaser or by a balance between opposing releasers, may result in release of unrelated activities. This is known as displacement behavior. The shaping of behavior by acquisition of learned patterns early in life is well documented for human language as well as for other systems such as bird song (Thorpe, 1961). Finally, ethologists recognize a variety of other behavioral mechanisms such as reflexive and conditioned behaviors.

Before discussing the value of the neuroethological approach, it is instructive to compare ethology with other approaches to the study of behavior. In many cases, whatever the differences in background, the methodology is similar. For example, many cognitive psychologists are essentially taking up the naturalistic approach when they investigate the psychophysics of humans and other animals. However, although the significance of evolutionary pressures on behavior may seem obvious, and in spite of the success ethology has enjoyed, the comparative naturalistic approach is not universally supported. For those psychologists who describe themselves as studying learning, a significant bias towards emphasis of conditioning paradigms persists.

> If, as animal psychologists, we wish to study behavioural capacities and competences which transcend particular adaptations, an attempt must be made to minimize the role of species-specific intelligence in the experimental task. The arbitrary nature of the psychologist's apparatus often represents an attempt to do just this. . . A similar case can be made for concentrating the major experimental effort on a limited range of species. If we are primarily interested in capacities that transcend any particular species, rather than in variations and differences between animals, comparative studies are of interest as a test whether a given capacity is general or not. However, such comparative validation can be seen as a secondary task to that of devising an adequate theory for some target species, usually the laboratory rat and pigeon, chosen in the first instance for purely technical reasons (Dickinson, 1980).

Those psychologists who chose to study inbred animals in unnatural laboratory environments performing arbitrary behavioral tasks will unavoidably minimize important aspects of behavior. Genetic and instinctive aspects of behavior, as well as the organizing role of sign stimuli, fixed action patterns, displacement behavior, and so on, are all easily demonstrable in a laboratory setting. These behavioral attributes of necessity must be included in an "adequate theory" of behavior. The notion that all or even much of vertebrate behavior can be explained by a unitary conditioning mechanism (Skinner, 1938) has been widely and vigorously criticized (Chomsky, 1959). One need search no further than the laboratory rat to demonstrate the limitations of narrowly construed ideas of behavioral organization. Maze running and bar pressing experiments ignore or minimize most aspects of a rat's natural behavior, such as ultrasonic communication, grooming, reliance on vibrissae, mother/offspring relationships, courtship behavior, and running in packs.

Ethology and Neurobiology. Perhaps the best test of a theory of behavior is the insight it provides into the neural mechanisms generating the behaviors. In this regard, ethology has contributed considerably. Many of the behavioral principles elucidated by ethologists have helped organize neurophysiological investigations. It is now possible to examine the specific neural circuits underlying several of these behavioral patterns. As one example, command neurons (Wiersma and Ikeda, 1964) are interneurons that when stimulated produce recognizable, stereotypic behaviors. Once initiated, the motor program goes to completion. Command neurons may subserve a variety of episodic and repetitive activities, including posturing, locomotion, feeding, and escape (Bentley and Konishi, 1978). Command neurons are typically embedded in a hierarchically organized system. They receive input from the brain and synapse onto pattern generator or other motor circuits. These circuits are the neural basis for the FAP. The term "command" has functional implications (Wiersma, 1952). If a command neuron is a decision-making element, it can be thought of as the neuronal embodiment of the previously postulated IRM. Although operationally defined by their response to electrical stimulation, command neurons are typically assumed to be recruited by the animal during normal execution of the behavior (Kupferman and Weiss, 1978). Usage, neccesity, and sufficiency are the criteria currently proposed for unambiguously assessing a neuron's "command" functions (Kupferman and Weiss, 1978). However, command neurons may meet all these criteria in reduced preparations, even while not meeting them for the behavior of the intact animal (Eaton, 1983). This suggests a unitary mechanism may not underlie the postulated IRM. The study of command neurons is perhaps the clearest attempt to define what it means, at a neuronal level, for a behavioral decision to be made.

As an example of a more theoretical issue in ethology that has bearing on the organization of the nervous system, I would like to examine the relative contributions of peripheral sensory filtering and central processing to brain and behavior. As originally envisioned by Lorenz, the Innate Schemata reflected central processes. This position provoked considerable controversy, as it was argued that peripheral mechanisms could account for perception of the sign stimulus (Lehrman, 1953). Indeed, peripheral specializations are now known to contribute significantly to stimulus processing of both invertebrates (e.g. Roeder and Treat, 1961) and vertebrates (Bruns, 1976; Frishkopf and Goldstein, 1963; Konishi, 1973). Selective filtering has been demonstrated for species-specific vocalizations, mate and prey recognition, and stimulus localization. All three extra-personal senses (vision, hearing, olfaction) can reflect these constraints. Peripheral filtering does not exclude stimulus-specific central processing (e.g. Knudsen and Konishi, 1978; Margoliash, 1983; O'Neill and Suga, 1979; Suga, 1977), even in the same animal. Indeed, central processing of behaviorally relevant stimuli is a well established principle (Bullock, 1977; Ewert et. al., 1983; Wordon and Galambos, 1972). A theoretical understanding of each species' choice of peripheral and/or central filtering awaits an understanding of the evolutionary pressures underlying stimulus filtering (Konishi, 1971). Presumably this understanding will be a product of ethological investigations, nevertheless neurophysiology has clearly contributed to its study.

The preceding discussion illustrates how interplay between ethology and neurophysiology can be mutually beneficial. The behavioral approach affords some sophistication to what must otherwise be regarded as the rather primitive tools of modern day neurophysiology. It is not overly optimistic to suggest that these principles can be applied to other, more generalized or more complex animals. For example, a systems analysis of visually directed flying in the domestic fly has produced an analytic description of the behavior, and suggests specific functional neural substrates (Reichardt and Poggio, 1978). In the primate, a systems analysis of the vestibulo-ocular reflex has provided a foundation with which to assess various hypotheses of the site of plasticity (Miles and

Lissberger, 1981). Neuroethologists chose to study simpler or specialized animals not because the theory is tenable only for these animals. It is simply easier. In such species the important sensory attributes and behavioral responses can more easily be identified and quantified, and the relatively large amounts of neural processing subserving these behaviors simplifies single unit investigations. Invertebrate preparations have received the majority of attention because they enjoy these and other advantages. There is no reason to despair when generalizing this approach to complicated or nonspecialized vertebrates. For these endeavors, I feel it is constructive to point out what is often missing is a solid behavioral base from which to launch neurophysiological examinations. Those complex systems for which intriguing behavioral observations exist, such as primate vocalizations (Green, 1975; Zoloth et.al. 1979), human speech (Morse, 1979), and vision (Julez, 1981) are likely candidates for expanding our knowledge of the neural basis of behavior. Assuredly the birdsong system must be counted among them.

Implications of Hierarchies and Maps for Brain Function

The strength of neuroethology lies in its ability, at least in general terms, to suggest or constrain the possible neural mechanisms responsible for integrative behavioral phenomena. The weakest aspect of the theory is that it cannot predict where in the brain the neural substrates lie, or in what form they may manifest themselves. In this section, I will attempt to establish some logic for the use of single unit electrodes in ferretting out these mechanisms, and I will give some suggestions as to what level of organization should initially be searched. These attempts can hardly be considered successful, as the arguments cannot be formalized.

Hierarchies. The brain is a computational machine made up of neurons (McCulloch and Pitts, 1943). What may be gleaned of the information processing in the brain by examining the activity at these points? Each neuron is a nodal point of convergence of its afferent input. As such, a neuron owes its response properties not only to its intrinsic properties (e.g. dendritic morphology), but also to the subset of neurons that impinges upon it. In this analysis, I ignore variations in the intrinsic properties of neurons. Thus, it is only the patterns of connectivity in a system of neurons that affects the system's functional properties. In a heterarchy (McCulloch, 1945), neurons cannot be assigned distinct levels within the system, and it is impossible *a priori* to predict anything about the constituent neurons' response properties. On the other hand, if a neuron is embedded in a simple hierarchy (without feedback), its response properties are determined solely by a subset of neurons at the same and lower levels in the hierarchy. That is, at each level in the hierarchy, a neuron's response properties reflects its intrinsic properties and the sum total of the processing exhibited by the particular subset of neurons that impinges upon it. The existence of feedback *per se* does not destroy the concept of hierarchy (Szentagothai and Arbib, 1975), although it may complicate the analysis. By manipulating the inputs to the hierarchy, and by examining the responses of single neurons at different levels of the hierarchy, one may systematically sample the information processing exhibited by various populations of subordinate neurons. In such a system, some of the information processing *must* be reflected in the activity of single neurons, and some of the processing *may* be encoded in the activity of populations (two or more) neurons. Thus, the existence of hierarchical organization in the brain (Stone et. al., 1979; Van Essen and Maunsell, 1983; Moiseff and Konishi, 1983) provides some rationale for the single neuron paradigm.

Hierarchical organization is a good design principle for an information processing system. Memory space is conserved in that different systems may share similar elements. This is exemplified by several computer processes utilizing a single subroutine. Decision making at the local level is possible, and results in minimizing information flow (Szentagothai and Arbib, 1975). The concept of a goal oriented, hierarchically organized system can be formalized, and it can be shown that an optimal level of interaction between subsystems of a given level exists (Mesarovic, 1968). This optimal level of information flow assures the maximal overall performance of the system. The amount of (local) information flow may be adjusted as (global) priorities change. Thus, a hierarchical system is well suited to adapt to a changing enviornment. Finally, hierarchical organization may be a potent evolutionary strategy (Dawkins, 1976). In a complex hierarchical system, various tasks may be distributed to various levels. As selective pressures demand new processing capabilities, the system need only add new elements at a given level of processing, or the system may add elements at a new, higher level. Either modification is a minor adjustment in connectivity compared with the complete reorganization required by a distributed processing system that adds new functions. The evolution of the primate visual system may reflect this advantage (Allman, 1977).

When considered in light of the behavioral approach, a surprising conclusion derives from the observation that the brain is hierarchically organized. Since behavior is the output or product of the nervous system, it follows that neural correlates to a behavior are most accurately predicted for the higher levels of the hierarchy. Of course, if the behavioral analysis is so complete as to suggest a specific functional organization, or if peripheral specializations suggest clear neural pathways associated with the behavior, then in theory a "bottom up" approach may be valid. It is a fact, however, that to date all the highly successful neuroethological systems have benefited from a "top down" approach. More than any other, this single observation precludes a rigorous codification of neuroethological thinking. One simply has to smell out the top.

Maps. One important contribution of integrative neurophysiology is the observation that sensory systems preserve throughout much of their extent maps of their corresponding sensory epithelia. For these purposes, a map is defined as a topographic arrangement of like elements, the elements systematically varying in one or several parameters. There are well documented examples of sensory maps derived from internal computation (Hubel and Wiesel, 1962; Knudsen and Konishi, 1978; Suga, 1977), some of which cannot be explained on simple bases such as closest neighbor wiring (Konishi and Knudsen, 1982). This general feature of nervous system organization provides further constraints on the brain's information processing scheme. Since each mapped element is similar, a complete description of the map does not require a complete description of each element of the map. Rather, it suffices to describe in detail several elements, and then to describe how the elements are mapped. Clearly, a mapped system is an attractive candidate for single cell neurophysiology.

When a map is embedded in a hierarchical organization, it may reasonably be assumed that the system contains multiple parallel channels, each channel reflecting the same encoding scheme while subtending incrementally different inputs. In the case of simple sensory maps, the "incrementally different inputs" can eventually be traced to incrementally displaced, contiguous sections of sensory epithelia. The concept of the ***hypercolumn*** supports this assumption: "a small machine that looks after all values of a given variable" (Hubel and Wiesel, 1974). In such a system, all the information processing up to the level of the map is represented by the response of the individual elements of the map. No information is represented in the ensemble of mapped elements that cannot be deduced from the activity of the individually examined elements.

It is interesting to speculate that the existence of maps within hierarchical systems is an information processing scheme the nervous system relies upon to constrain all encoded information into single channels, to the exclusion of population phenomena. This contention is supported by considering what advantages a mapped system affords the nervous system. Initially, one might conclude simply that maps derive from the organization of the periphery, which is in turn a product of interaction between the constraints on biological systems (specifically cellular organization) and the selective pressures of evolution. Given the existence of a mapped periphery, the corresponding central structure could be a product of developmental constraints such as passive guidance. Such mechanistic explanations do not really address the question of *why* there are maps however, as it is clear that maps based on internal (central) computation are common. Another popular assertion is that maps exist because they simplify interactions among functionally similar member elements. Processes such as lateral inhibition would only require nearest neighbor wiring. In the case of maps that emerge from internal computation, this presents an interesting point. Clearly, the information represented in the map exists at a lower level of the hierarchy, although at some level it is not mapped. The nervous systems must expend considerable energy (measured in terms of development and maintenance of complex connectivity) to extract the information from the subordinate population. It necessarily follows that this effort is justified on the basis of an advantageous organizational scheme. For mapped systems, the representation of information at the level of single cells must be inherently superior (for whatever reasons) to having the same information represented in a population.

A Critique of Critisisms of the Single Cell Doctrine

The preceding discussion has hopefully provided some reasons why single cell recording techniques are appropriate for the study of integrative brain function. In this section, I wish to respond to criticisms of the single unit doctrine. It is perhaps useful at this point to restate the ideas relating single cells to perceptual phenomena. Barlow (1972), in a seminal article, has articulated a central dogma of the single unit hypothesis:

> Perception corresponds to the activity of a small selection

> from the very numerous high-level neurons, each of which corresponds to a pattern of external events of the order of complexity of the events symbolized by a word.

The other great mover in the field, Bullock, offers a more ethologically based version (Bullock, 1983a):

> The proposition, as I would put it, is that some social signals and other natural stimuli contain the normally necessary and sufficient conditions for maximal response of a limited population of roughly equivalent neurons. That is to say, some neurons are so selectively sensitive as to respond well only to a set of stimuli which may appear complex to us, are probably in most cases natural and form an ethologically defined set; the population that so responds is not large or very heterogeneous.

There is a large and growing body of evidence to indirectly support this hypothesis (Ewert et. al., 1983), a paucity of data indirectly contradicting it (c.f. Partridge et. al., 1981), and in a least one system it should now be possible to directly test the hypothesis (Margoliash, 1983). Given this substantial (certainly not conclusive!) body of evidence, it is disappointing that the criticisms are of a loose, general nature, and often contain substantial factual errors (see below). Such failings may be the result of critiques by non-specialists - workers expert in a seemingly related field (e.g. psychology) who nevertheless are not competent to assess neurophysiological data. At the same time, it is difficult to assess the validity of the numerous alternative population models proposed, since experimental evidence that choses among them is scarce indeed. Given this state of affairs, one wonders if much of the attraction for population mechanisms stems not from a fundamental weakness in single cell ideas rather than the difficulty of a theoretical treatment of an information processing system comprising heterogeneous elements. One should note with interest that models of learning and memory systems (e.g. Hopfield, 1981) typically assume a large population of homogeneous elements, but which observations of neuroanatomy or neurophysiology support this notion?

One criticism of the single unit doctrine is that while it may be formulated in terms of propositions, these do not serve a powerfully predictive function. The criticism is not without merit - currently it is risky at best to attempt a prediction of the response properties to be encountered in a new area. Nevertheless, this limitation may be more of a reflection of the fact that neurobiology is a young science with relatively few well established guiding principles. Of necessity, this must limit the power of any theory. Furthermore, historically experimental science has most often led and provided the basis for the theoretical constructs that followed.

Other critisisms are rather misinformed. For example, in a chapter entitled "Mezzolog - Limits of Neuroreductionism", Uttal (1981) states:

> The brain of all vertebrates, and of man in particular, is immensely complicated . . . Yet each of these 10 trillion neurons may itself have many thousands of synaptic connections with other neurons. The totality of this maze of interactions defines a system intrinsically so complex that it may not be computable or analyzable in any practical sense. Thus even though microelectrodes may be inserted into single cells, the recorded responses can only poorly reflect the activity of the very complex network of other neurons of which that neuron is an interacting part. The responses recorded from the individual neurons can at best provide only the most incomplete indication of the true functional complexity of the neural net. In vertebrates, in particular, neurons are small and idiosyncratic, and recordings of their responses are ephemeral at best. Even if there were some specific functional role played by each neuron in vertebrate preparations, it is virtually impossible to recover a particular cell after the tip of (sic) recording microelectrode has moved on, and thus improbable within one preparation to replicate any observation. How much less likely it would be to replicate the recordings obtained from any particular neuron as one moves from one preparation to another!

While I share Uttal's awe at the complexity of the human brain, it is just so much hand-waving at this point to quote huge numbers and mutter something about non-computability. As a reasonable counterexample, the macaque brain is hardly simple, yet it has yielded many secrets under the scrutiny of the single unit electrode. There is no reason *a priori* to believe this happy state of affairs will not continue. More damning, however, the paragraph is replete with factual errors. Since when are vertebrate neurons in particular small and idyosyncratic? Some invertebrate nervous systems such as *Drosophila* have truly small neurons; even in these systems neural recordings are possible. One wonders what is idiosyncratic about the morphology of, say, cerebellar neurons. Furthermore, what is ephemeral about a recording from a sensory neuron that can be maintained for many hours, during the course of which thousands or tens of thousands of stimuli can be presented? Finally, it is nonsensical to claim that single cell observations are not reproducible. As one of many examples, topographic organization in sensory systems are delineated time and again in hundreds of laboratories and thousands of experiments.

As a final example of weak criticism of the single cell doctrine, I point out a disappointing misconception in an otherwise scholarly work by an emminent theoretician (Marr, 1982):

> Supose, for example, that one actually found the apocryphal grandmother cell (A cell that fires only when one's grandmother comes into view). Would that really tell us anything much at all? It would tell us that it existed - Gross's hand-detectors tell us almost that - but not *why* or even *how* such a thing may be constructed from the outputs of previously discovered cells. Do the single-unit recordings - the simple and complex cells - tell us much about how to detect edges or why one would want to, except in a rather general way through arguments based on economy and redundancy? If we really knew the answers, for example, we should be able to program them on a computer. But finding a hand-detector certainly did not allow us to program one.

On this point, Marr lacked imagination, or perhaps the insights of a neurophysiologist. Imagine, for a moment, if hand or face detector neurons were discovered (many physiologists dispute Gross's claims). Would not such a finding potentially speak to the critically important questions of hierarchical organization and the information codes used by the nervous system? Would these neurons exhibit stimulus equivalence? Would they be multimodal? After elucidating the cells' connectivity (the techniques exist!) we could ask extremely interesting questions. What would be the response properties of the penultimate cells? Imagine the excitement if the output of the grandmother cells synapsed onto cells at a lower level in the hierarchy! Answers to these questions would suddenly be accessible and would be an enormous contribution to our understanding of visual perception. The trouble is precisely that the evidence for hand detector neurons is unacceptably weak. Marr alludes to a good counterexample to his own critique. In the case of the edge detector neurons, the results are entirely convincing (see D. Hubel, the Nobel Lectures, 1981), and much is known about the functional hierarchy and other neural mechanisms that generate the response (Stone et. al., 1979). If one had a few million photoreceptors and perhaps 100 million neurons one could play with, surely one could build a crude approximation to cortical edge detectors. The problem with programming the process is not trivial at least in part because one is demanding real- time response from a single serial computer process that is attempting to model many millions of real-time parallel neural processes. Indeed, there is a growing awareness in the artificial intelligence community that massive parallelism is required to solve complex problems (S. Bankes, personal communication). Isn't this lesson clear from the study of the nervous system?

Chapter 1

Zero Crossing Analysis as Applied to Bird Songs

Neuroethologists incorporate species-specific tokens in stimulus repertoires designed for single cell neurophysiology experiments. These stimuli are typically complex, and often elicit complex responses. Experimental paradigms that incorporate complex stimuli but do not enjoy sufficiently powerful analytic and synthetic techniques to conveniently manipulate these stimuli inevitably produce incomplete results. For example, Gross et.al. (1972) have described visual neurons in the monkey inferotemporal cortex that are seemingly selective for monkey and human faces. Unfortunately, the inability to delineate the visual parameters essential for the response (circle, circles within circle, eyes within circle, entire face ?) leaves unresolved the question as to whether the neurons are selectively attentive to specific complex visual stimuli.

The present studies required the testing of many units with a large repertoire of species-specific bird songs. To limit the possible ambiguities in the interpretation of the neuronal response properties, it was essential to be able to delineate the acoustic parameters within the song (such as frequency, amplitude, and temporal pattern) eliciting the responses. This entailed the systematic modification of song, as well as the computation and presentation of novel stimuli chosen at the time of experimentation. For technical reasons, isolated single unit extracellular recordings typically can be maintained for minutes and only occasionally for several hours. This limited experimentation time is particularly onerous when complex stimuli elicit complex responses, requiring a barrage of tests to satisfactorily explore the response domain. Thus, the success of these experiments was in part dependent on the ability to rapidly yet flexibly modify bird songs.

The general solution to the problem of complex stimulus representation and manipulation has eluded investigators. Techniques applicable to most classes of stimuli are either extremely limited in terms of synthetic power (e.g. Watanabe and Katsuki, 1974; Wollberg and Newman, 1972), or require many minutes or hours of computation on medium sized laboratory computers (Dorrscheidt, 1978; Zoloth et.al., 1980; Dooling et.al., 1982), an unacceptable constraint on single unit recordings. My goal was to investigate the relationship between song acoustic parameters and unit response properties, hence a novel approach had to be developed, one which embodied rapid yet powerful analysis/synthesis techniques suitable for bird songs.

The technique I chose is based on the analysis of the zero crossings of a signal. An arbitrary signal has an ambiguous relationship to its zero crossings. Nevertheless, the zero crossings of a signal often encode important information. Human speech, for example, is a complex signal that comprises many formants. Surprisingly, human speech sounds retain linguistic information after being subjected to infinite peak clipping, which preserves only the zero crossings (Licklider and Miller, 1951). Greenewalt (1968) was the first to observe that many bird vocalizations, especially songs, fall into the class of signals that can be recovered within a multiplicative constant from their zero crossings. For such ``single voice'' signals, analysis is simplified in that the instantaneous frequency can be derived from the easily detected zero crossings of the signal (e.g. Staddon et.al., 1978). Accurate reproduction of the signal is also simplified since the frequency and amplitude of only a single oscillator need be controlled. Furthermore, the systematic modification of signals represented as frequency and amplitude functions of time is not computationally burdensome.

Only band-pass limited signals can be accurately modeled by their zero crossings. The sound spectrograph (sonagram) is not so limited, and in terms of spectral analysis is clearly a more powerful technique. Zero crossing analysis, however, does enjoy important advantages over sonography. The sonagraph is based on a Fourrier-type of analysis, wherein the signal is decomposed into a sequential time series prior to analysis. This results in an inherent tradeoff between the frequency and temporal resolution with which the signal can be resolved. As zero crossing analysis measures the instantaneous frequency, it does not suffer from such limitations. Therefore, those signals that can be recovered from their zero crossings are more accurately modeled by zero crossing analysis than by the sonagraph. For animal vocalizations such as bird songs, which contain rapid modulations, this is an important advantage.

DIGITAL ZERO CROSSING ANALYSIS

Zero crossing analysis (ZCA) is readily accomplished with available analog electronics, and has been described in detail (Staddon et.al., 1978). For the present study, however, a number of considerations indicated that a digital/analog hybrid implementation afforded significant advantages over a purely analog or purely digital approach. The outputs of the ZCA of a signal represent frequency and amplitude as time functions of the original input signal (hereafter referred to as an original song). The difficulty in *a priori* predicting what signal processing for parameter modification is necessary to fully elucidate unit response properties makes the digital approach, with its inherent flexibility, an attractive choice. In comparison, the design, implementation, and testing of analog hardware dedicated to various signal processing tasks as continuously modified by experimental considerations represents a significant overhead. Furthermore, the experimental paradigm necessitated a large repertoire of over 50 songs, while limitations in the ability to maintain single cell recordings necessitated that the songs be available for immediate recall. It is hard to achieve acceptable data acquisition times with serial analog storage devices, but not with random access digital devices. On the other hand, in a purely digital implementation, the final output waveform must be computed point by point from the time-varying frequency and amplitude functions. This task is memory and computationally intensive, and can not be achieved in real-time on the computer system available for these investigations (PDP 11/40, 24 Kwords available memory, no floating point processor). To avoid an unacceptable system response time, the synthesis utilizes analog hardware to produce the final output waveform. Combining ZCA of digital signals with analog synthesis thus provides for a happy marriage: the computer can manipulate signals at a conceptual level (i.e. song phrase or song element) or a parametric level (e.g. frequency or amplitude), while the analog hardware eliminates the computationally difficult task of calculating the final waveform. ZCA, as implemented with a combination of digital and analog techniques, provides a convenient facility for exploring the signal parameters underlying selective neuronal responses to complex stimuli.

Digital ZCA is a three pass procedure. During the initial pass, an analog recording of an original song is digitized. For the species used in these experiments, the highest frequencies present in the songs never exceeded 10 kHz. Thus, 10 kHz low-pass filtering combined with a 20 kHz sampling rate gave consistently excellent reproductions of the original songs. In digitizing the song, any DC offsets in the buffer amplifiers, the sample and hold amplifier, or the analog to digital converter (ADC), will affect the mean of the signal. The second pass corrects such errors and computes the zero of the signal. The zero value of the digitized signal can be traced to the resting position of the internal tympaniform membranes of the syrinx, the sound producing organ in birds. The internal tympaniform membranes vibrate symmetrically, and thus the zero-order statistics are time invariant. As the signal is stationary, a fiducial value for the zero of the signal may be estimated as the mean value of the entire digital stream. It should be noted that the instantaneous period of a cyclical signal will be accurately estimated by calculating the interval between *any* two corresponding points in consecutive cycles, not just at the zero crossings. Thus, ZCA is not critically dependant on the estimation of the zero. Rather, this value is important primarily to assure that the instantaneous frequency of the low amplitude part of a signal is recovered.

The third analytic pass can then extract frequency and amplitude information. Each crossing lies between consecutive points that straddle the fiducial value. The time of crossing may be reasonably estimated by linear interpolation. Adjacent crossings in the same direction (e.g. lower to higher: positive crossing) define the period of one cycle, and hence the instantaneous frequency:

$$f_c = \frac{1}{T} = \frac{1}{t_{n+1} - t_n} \tag{1.1}$$

If several cycles occur within a time window, the duration of which is chosen on the basis of factors related to synthesis (see below), the average frequency of the cycles within the window (f_{avg}) can be used as an estimate of the frequency for that interval. With the frequency estimate in hand, it is then possible to estimate the associated amplitude. For N data points $P_i(p_i - \bar{p}, t_i)$ within one cycle,

$$p_i - \bar{p} = a_i \sin(2\pi f_c t_i) \tag{1.2}$$

where a_i are the amplitude coefficients and $\bar{p}$ is the estimated zero. The amplitude of the cycle is estimated as the mean:

$$a_c = \frac{\sum a_i}{N} = \frac{\sum_{i=1}^{N} \frac{p_i}{\sin(2\pi f_c t_i)}}{N} \tag{1.3}$$

and the average is computed over all cycles of the interval. The successive values (f_{avg}, a_{avg}) may thus be computed and then stored on disk. As discussed in Chapter 2, the most useful manipulations of the song parameters modify selected frequencies or amplitudes but not both. Thus, it is convenient to store the frequency and amplitude streams independantly, as opposed to multiplexing them.

To synthesize a song, successive digital values representing frequency and amplitude are converted to voltages (12-bit DACs) at a rate determined by the chosen time window. The computer initiates a sequence of analog events at the start of a cycle by triggering custom designed electronics. The electronics consist of a voltage controlled oscillator (VCO - controls frequency), a voltage controlled amplifier (VCA - controls amplitude), and special linearizing circuits (see *Appendix*). At each timing pulse, the analog hardware acquires and stores the new voltages, and executes linear voltage ramps during the interval so as to achieve the new values at the end of the interval. (Note that the fixed interval eases the computational problem of analog calculation of the endpoints). In this manner successive voltage steps from the computer are converted to voltage ramps. As the voltages determine the frequency and amplitude of the signal, the synthesis technique results in a song whose frequency and amplitude modulations are approximated by piece-wise linear increments over successive time intervals.

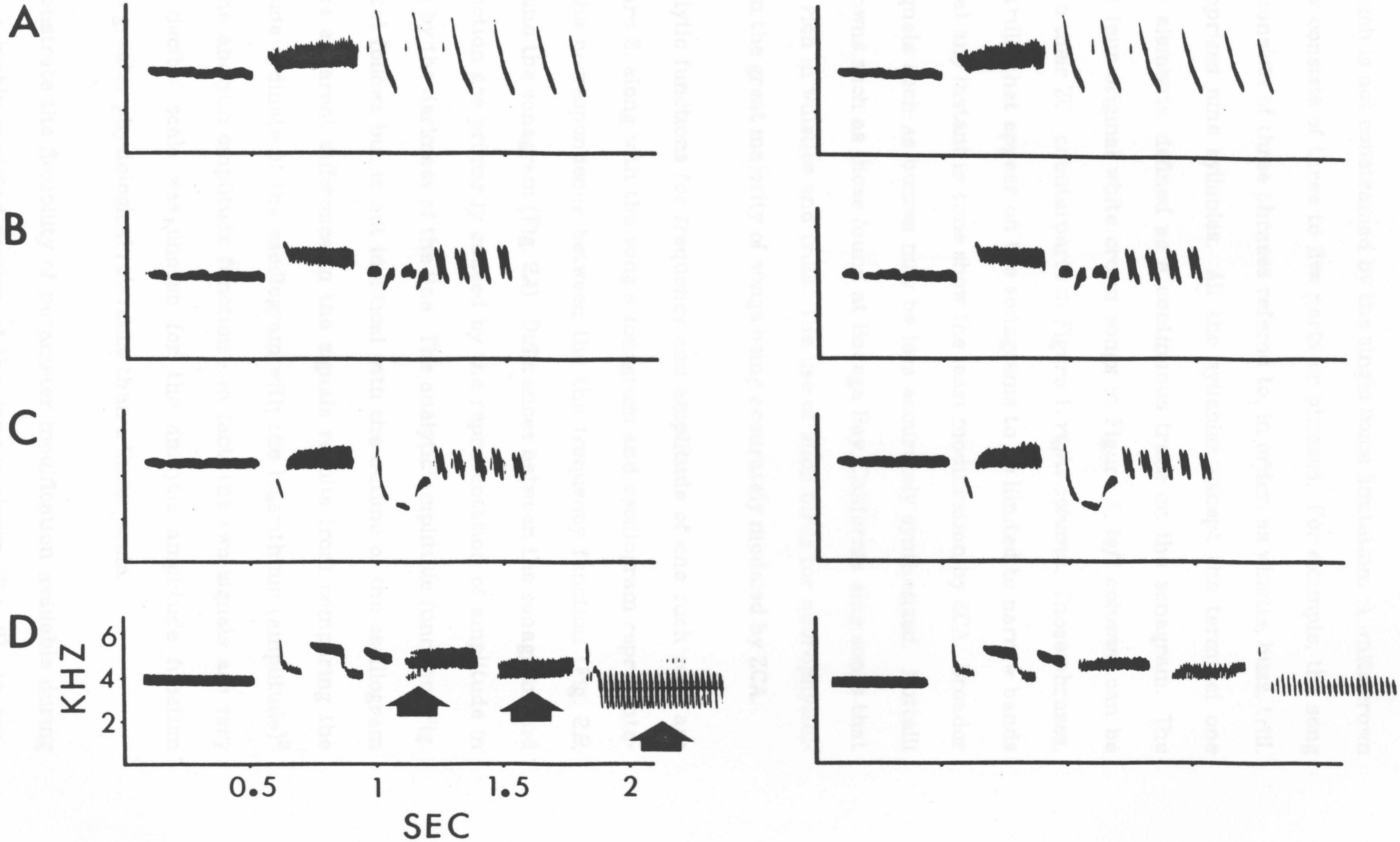
A fundamental parameter in the digital ZCA outlined above is the time window interval. A number of conflicting requirements impinge upon the choice of interval. As the memory size required is inversely proportional to the interval, it would be convenient to chose a relatively long interval. On the other hand, certain classes of artificial sounds such as clicks and bursts have short (approximately one millisecond) risetimes, and thus require a short interval. Also, a long interval can introduce analog errors caused by "droop" - diminution of signal in the sample and hold device (see Appendix A). Most significantly, however, the time window averages over several cycles, and can potentially introduce spurious results. If, over the interval, modulations of the signal approximate the duration of the time window, interactions between the carrier and the modulation will result in an inaccurate average instantaneous value. The rate of modulation within a repertoire of signals affects the accuracy of analysis for a given time window, and thus a criterion can be established.

The error caused by the piece-wise linear approximation of the signal can be assessed by comparing the computer synthesized version to the original song. To determine an acceptable time window, a survey of the natural vocalizations used in the neurophysiology experiments was undertaken. The songs of fifty-five male white-crowned sparrows were tape recorded, digitized, and subjected to ZCA. These songs include samples from four of the five extant races of white crowns. The repertoire includes the songs of 12 naturally occuring dialects, the songs of 15 birds that failed to learn the songs they were tutored in the laboratory, two samples of the songs of birds isolated from any broadcast sounds during the critical period for song learning, as well as one example of the song of a bird deafened early in life. This repertoire represents a comprehensive survey of the potential natural and laboratory-induced song parameters of the white-crowned sparrow. Within this repertoire the most rapid modulations occured within the buzz and trill phrases of the songs. The modulation period was never greater than five to ten milliseconds (c.f. Leppelsack, 1981). The use of a much shorter (one millisecond) time window constrained the frequency and amplitude modulations so as to never approximate the time window interval. With this time window, the songs were accurately reproduced.

A DEMONSTRATION OF THE TECHNIQUE

The accuracy of analysis/synthesis can be assessed by use of the

Figure 1. Synthesis of white-crowned sparrow songs utilizing zero crossing technique. Sonagrams — frequency versus time plots — in *left* column are directly from tape recordings of singing birds. Sonagrams in *right* column are computer synthesized versions using frequency and amplitude functions extracted from the original songs (see text). *A*, Nuttalli (coastal) race. *B*, *C*, Oriantha (montane) race. *D*, Gambelii (Alaskan) race. In *A* and *B* reproduction is essentially perfect; in *C* small differences somewhat obscured by the photographic process appear in the buzz (second phrase); in *D* clear differences are present in many phrases (see *arrows*).


A
B
C
D
KHZ
6
4
2
0.5
1
1.5
2
SEC

sonagram, which is not constrained by the single voice limitation. A white crown song typically consists of three to five parts or phrases. For example, the song of Figure 1*A* consists of three phrases refered to, in order, as whistle, buzz, trill. The trill comprises nine syllables. All the syllables except the terminal one comprise two elements, defined as a continuous trace on the sonagram. The sonagrams of four original white crown songs in Figure 1, *left column*, can be compared with their ZCA counterparts in Figure 1, *right column*. Those phrases, whistles and trills, that appear on the sonagrams to be limited to narrow bands of frequency at any instant in time show the least modification by ZCA. Broader bandwidth signals such as buzzes may be less accurately synthesized. Nuttalli race white crowns such as those found at Bodega Bay, California sing songs that are relatively rich in whistles and trills. The use of such birds for neurophysiology resulted in the great majority of songs being accurately modeled by ZCA.

The analytic functions for frequency and amplitude of one such song are shown in Figure 2, along with the song's sonagram and oscillogram representations. Note the correspondence between the the frequency function (Fig. 2*B*, *upper trace*) and the sonagram (Fig. 2*A*). Differences between the sonagram and frequency function are primarily caused by the representation of amplitude in the sonagram by the darkness of the line. The analytic amplitude function (Fig. 2*B*, *lower trace*) follows but is not identical with the outline of the oscillogram (Fig. 2*C*). This apparent difference in the signals results from comparing the linear amplitude ordinate of the oscillogram with the logarithmic $(\text{amplitude})^2$ ordinate of the analytic amplitude function - in fact, the two signals are very similar. The decibel scale was chosen for the analytic amplitude function because it has greater physiological relevance than a linear scale.

To demonstrate the flexibility of parameter modification available during the synthesis, a highly modified version of the song is shown (Fig. 3). In con-

Figure 2. Relationship between computer, sonagram, and oscillogram representations. The sonagram in *A* and the oscillogram in *C* are analyses of a song produced by the computer generated analytic functions of *B*. *Top trace* in *B* is frequency versus time, ***bottom trace*** is amplitude versus time. In this and all other examples in this paper of computer representations of frequency functions, the part of the otherwise continuous function corresponding to low amplitude values is blanked as a visual aid.

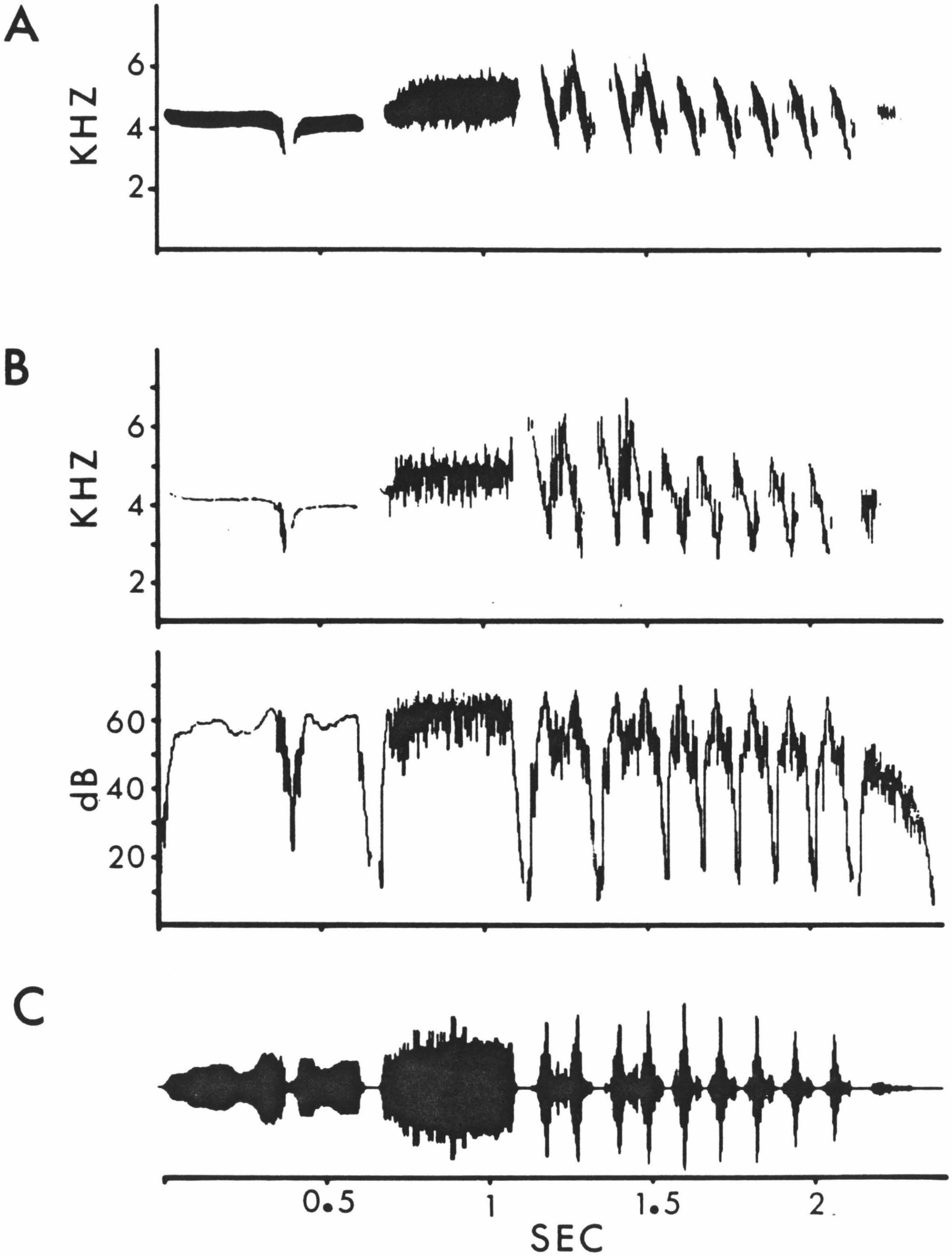
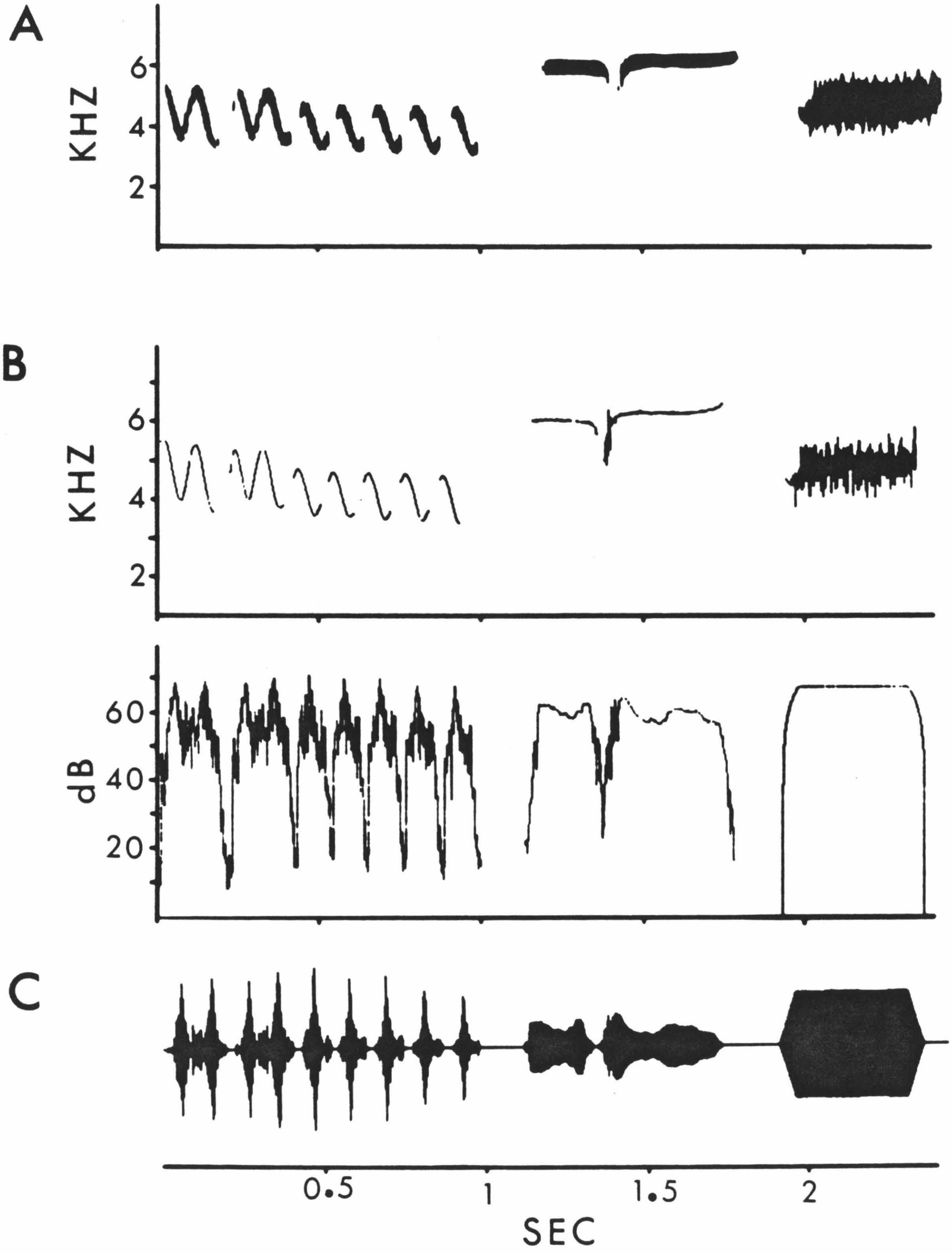

A
KHZ
6
4
2
B
KHZ
6
4
2
dB
60
40
20
C
0.5
1
1.5
2
SEC

Figure 3. Modification of computer generated analytic functions of frequency. and amplitude. The song of Figure 2 is highly modified (see text). These modifications reflect many of the parameter manipulation features available in *APES*, including phrase order rearraingment, inter-phrase interval modification, frequency and amplitude shifting, independent frequency and amplitude modification, and temporal reversal of individual phrases.



A
6
KHZ
4
2
B
6
KHZ
4
2
60
dB
40
20
C
0.5
1
1.5
2
SEC

trasting Figure 3 with Figure 2, note that phrase order and inter-phrase intervals have been changed. One phrase has been eliminated. The whistle is increased in frequency and is played in reverse, the buzz has an artificial amplitude (but not frequency) envelope, and the frequencies in the trill have been low-passed filtered (smoothed). Expressing signals according to their frequency and amplitude versus time components enables independent modification of frequency and amplitude. For example, manipulation of the computer generated frequency function (Fig. 3*B*, ***upper trace***) is reflected in the sonagram (Fig. 3*A*), but with no change in the oscillogram (Fig. 3*C*); the converse is apparently true for changes of the amplitude function. The expected results may be confounded by introducing or removing amplitude modulations that are sufficiently rapid as to induce significant frequency components. In the present experiments, such modifications were avoided. It can therefore be seen that this approach combines a reasonable degree of accuracy in representation of the original stimulus with substantial on-line flexibility for parameter modification. Sufficient analytic and synthetic power exists in this scheme to apply a systematic approach to the elucidation of the parameters underlying the responses of units to complex stimuli.

COMPUTER CONTROL OF ONLINE SYNTHESIS

To realize the potential offered by the ZCA technique, I developed a collection of programs. This system, APES (*A*uditory *P*ysiology *E*xperimentation *S*ystem), incorporates a powerful command processor to enable the user to parsimoniously describe sound stimuli and the modification of those stimuli. Other features include translation routines to convert internal format descriptions of stimuli into the analytic frequency and amplitude functions necessary for stimulus generation, time-of-event spike data acquisition routines, a module for analyzing neuronal responses and graphically displaying the stimulus and response on a storage oscilloscope, and lastly, file manipulation and other ancillary routines. To maximize speed and minimize memory requirements, the programs are written in Macro-11, the PDP-11 assembly language. Nevertheless, the system has to be implemented as an overlay structure to fit into the 24 Kwords of memory available for this project.

Three fundamentally different classes of stimuli are controlled by APES. Each class has its own internal representation. The first class is the digitized version of analog waveforms, whose internal format is a stream of consecutive digital values. This representation faithfully describes any signal, but is very difficult to manipulate parametrically (e.g. frequency shift). The utility of such stimuli for investigation of single unit response properties is therefore minimal, and they were used only as a last resort, when all other stimuli were ineffective.

The second class of stimuli are those described by their zero crossings. The internal representation for these stimuli is the previously described frequency and amplitude functions. Such functions are derived by an explicit transformation of the input data. The input to the process, however, may be a signal such as a bird song, for which no explicit description (model) exists. This latter point is important because it implies a significant constraint on the manipulation of the frequency and amplitude functions. Without an explicit model of the relevant acoustic parameters in a song, for example, no logical basis exists for changing the duration of the song (i.e. deleting or adding data). Fortunately, many modifications used in physiology experiments do not require an explicit model. Examples of these include frequency and amplitude shifting, and sequence rearrangement.

The final class is the artificial stimuli. For these stimuli, an explicit description exists, and thus any manipulation is possible. As well as the standard artificial stimuli such as noise bursts, tone bursts, and frequency modulations (FMs), APES permits the construction of complicated yet fully defined stimuli such as artificial whistles and buzzes. If necessary, it is possible to define any arbitrary frequency or amplitude function.

A system that affords generality and flexibility is potentially encumbered with a control language that is complex or difficult to use. To minimize these problems, the command language was designed to enable efficient descriptions of sounds in a natural format. It is a striking feature of most complex sounds, especially animal communication sounds, that they comprise a temporal sequence of elements (Marler, 1969). It is for this reason that sound spectrographic analysis, which results in a pictorial representation of the time-varying quantities of a signal, has been so useful in the analysis of vocalizations. APES fully exploits these observations. The description of a complex sound is built up as a series of elemental components. The temporal sequence with which these elemental sounds are reproduced is initially determined by the order in which they are specified, and may be conveniently rearranged with simple commands. Finally, the stimuli are graphically displayed on a storage oscilloscope as frequency and amplitude functions of time.

The APES command language recognizes a primitive syntax which results in command strings reading somewhat like a natural language. During execution of a command string, the command processor pushes instructions onto a stack. Stack-oriented string parsing lends itself to repetitive execution of any subset of a command string. An experimenter may wish, for example, to instruct the computer to output a sequence of two frequency modulations eight times while collecting spike data, and upon completion to display the final results as a post-stimulus time histogram. The corresponding APES command string is

CLEAR 2 * (*FM* 65 5500 3000 *INTERVAL* 50) 8 * (*COLLECT*) *PSTH* (1.4)

CLEAR is a command to initialize the sound buffer. The *2 * (. . .)* construct instructs APES to execute the string within the parentheses twice. (The ** (* is optional and is included merely for clarity). In this example, the two FMs each have a duration of 65 msec, sweep from 5.5 kHz to 3.0 kHz, and are separated by a silent period of 50 msec. Note that as the amplitude, risetime, and falltime of the FMs are not defined in this example, they take on default values. In APES, these default values are constantly updated as new sound elements are defined. *COLLECT* invokes the stimulus generation/data collection routine, and the results of eight repetitions of the stimulus are displayed by the routine *PSTH.*

Each sound element type is uniquely defined by its acronym and location in the sound element buffer. Associated with each sound element is a sequential list of parameter types that describes the element. For example, the parameter list for a tone burst is duration, frequency, amplitude, risetime, falltime. This construct enables the user to access specific sound elements or specific parameter of a sound element with English-like commands. After executing (1.4), the command string

10 * (*COLLECT FOR FM* 2 *INCREASE FREQUENCY* 100) (1.5)

will output the FM pair ten times while increasing the starting and ending frequencies of the second FM in 100 Hertz increments. Subtypes can also be defined. Thus, the user may conveniently specify any systematic modification of a stimulus as it is generated repetitively. Random sequences of stimuli may also be defined, as in

XFILE ISOLAT 5 * (*RANSNG COLLECT*) (1.6)

In this example, *XFILE* is a command to read into memory the analytic frequency and amplitude functions associated with the white-crowned sparrow song *ISOLAT*. The file header in *ISOLAT* contains a description of the starting and ending locations of the phrase(s) that constitute the song. Each time *RANSNG* is invoked, these values are accessed to define a new, pseudorandom phrase order.

In Tables 1 and 2, the various sound elements, parameter types, parameter fields, and parameter modification commands are listed. As can been seen, the APES command processor defines a diversity of stimuli, and affords a wealth of procedures for the modification of these stimuli. With this amount of flexibility, it is clear that analysis can not be limited to a previously defined protocol of standard stimuli. Thus, a complete description of the stimuli used in an experiment must be saved along with the neuronal responses. In APES, this is accomplished by saving both the initial description of the stimulus and the command string(s) that describe any stimulus modifications. The experiment may be reconstructed with this stored "telemetry". Several off-line analysis routines make use of this feature, most notably a routine that automatically analyzes data at the end of an experiment.

Table 1. Artificial stimuli currently defined under *APES*

Sound element	Description	Parameter field
NOP	silent period	dur
PT	tone burst	dur,freq,linamp,rise,fall
WN	white noise	dur,linamp,rise,fall
BN	band-pass noise	dur,freq,fmod,linamp,rise,fall
FM	frequency modulation	dur,freq_1,freq_2,linamp,rise,fall
BUZZ	buzz	dur,freq,fmod,fper,linamp,rise,fall
WHS	whistle	dur,freq,fmod,fper,linamp,rise,fall
MULTI	arbitrary FMs	freq_1,dur_1,...., freq_n,dur_n,freq_{n+1},linamp,rise,fall

Parameters	Description
dur	duration (msec)
freq	frequency (Hz)
fmod	height of frequency modulation (i.e. $f_{max} - f_{min}$) centered around ***freq*** (Hz)
fper	period of frequency modulation (msec)
linamp	maximum intensity of linear amplitude waveform (dB)
rise	risetime (msec)
fall	falltime (msec)

Table 1 (*continued*).

Parameter modifiers	Description
increase x	add x to parameter
decrease x	subtract x to parameter
incper x	add $x\frac{parameter}{100}$ to parameter
decper x	subtract $x\frac{parameter}{100}$ from parameter

Table 2. Definition and modification of zero-crossing stimuli in *APES*

Phrase sequence	Description
XFILE *filnam*	Reads *filnam* into memory; defines normal song
SONG	Defines normal song, i.e. normal phrase sequence of song *filnam*
SEQ $p_x, p_y, p_z, \cdots$	Defines any phrase sequence $p_x, p_y, p_z, \cdots$ in terms of normal phrase position
REVERSE p_x ,...	Flags each phrase p_x ,... to be played in reverse
RANSNG	Normal sequence of phrases is randomized
RANPHR	Pseudorandomly picks one phrase from *filnam*

Phrase modification	Description
ARTAMP	Assign linear amplitude waveform to phrase
ARTFREQ	Assign constant frequency to phrase
SMOOTH	Low pass filtering of frequency function
LINEARIZE	High pass filtering of frequency function
DIGITIZE	Cursor defines new phrase
CUT x,t_0,t_1	Define phrase x over time $[t_0, t_1]$ r.e. start of phrase

Chapter 2

Song Specific Units in the Song System of Adult White-Crowned Sparrows

A wealth of behavioral observations on birdsong indicates it is an attractive system for the study of neural encoding of complex stimuli (Marler and Mundinger, 1971). Song is essential for a songbird's reproductive success. In the majority of species, males sing both as an advertisement of territoriality and for mate attraction (Thorpe, 1961). For some species or subspecies of birds, songs may vary across but be similar within a locale; these song types are termed dialects (e.g. Marler and Tamura, 1962). Adult songbirds frequently make discriminations between the songs of neighbors and strangers (Brooks and Falls, 1975; Emlen, 1972); these discriminations may reflect associative learning, not habituation (Richards, 1979).

An individual's song is learned, and is sensitive to the effects of sensory deprivation during a critical period early in life (Thorpe, 1961). Most songbirds, when acoustically isolated from their species' song throughout the critical period, sing a highly abnormal song never found in the wild (Thorpe, 1961; Marler, 1970; c.f. Mulligan, 1966; Kroodsma, 1977). These songs are termed ``isolate'' songs. In some species, such as the white-crowned sparrow (*Zonotrichia leucophrys*), the critical period closes months before the bird enters a motor development phase that results in adult, ``crystallized'' song (Marler, 1970). Auditory feedback is necessary for normal motor development (Konishi, 1965). This observation led to the hypothesis that the song model learned during the critical period is stored as a neural memory trace or ``template'' that is recalled or accessed via auditory feedback during the song crystallization process (Konishi, 1965, 1978). A distinct group of nuclei involved in song production in adult birds has been described by Nottebohm et.al. (1976). Two of these nuclei, HVc (hyperstriatum ventrale, pars caudale) and RA (nucleus robustus archistriatalis), are contiguous with areas that receive projections from field L (Kelly and Nottebohm, 1979), the primary auditory telencephalic nucleus in birds. The auditory recipient zone associated with HVc forms a layer along the ventral border of HVc and has been termed the "shelf". The role of auditory feedback in song development suggests that these recently demonstrated connections between the auditory and motor systems may mediate or be involved in song learning. In an intracellular study of anesthetized birds, Katz and Gurney (1981) reported that many HVc neurons responded weakly to noise bursts. Furthermore, multiunit activity from HVc of unanesthetized and minimally restrained birds responded to recordings of species-specific (conspecific) songs, including the birds' own song, and gave weaker responses to songs played in reverse (McCasland and Konishi, 1981). Neither study, however, demonstrated that auditory responses in HVc reflect or are modified by some aspect of song learning. The present study utilized an uncommon but behaviorally relevant stimulus search paradigm, with emphasis on the bird's own song, to investigate the response properties of single HVc and shelf units recorded extracellularly. This stimulus paradigm has proven to be a powerful tool.

MATERIALS AND METHODS

White-crowned sparrows of the sedentary Nuttalli race were collected from the field either as nestlings, fledglings, or adults. Many adults and fledglings were collected at one time from one locale (Bodega Marine Laboratory, Bodega Bay, California). Most but not all of these birds sang individual versions of the Bodega Bay dialect. Nestlings were housed individually in sound attenuating boxes and exposed to various song models ("tutored") during the critical period. Some nestlings copied the song(s) with which they were tutored with varying degrees of accuracy, while others sang crystallized but isolate-like songs wholly different from either their tutor song(s) or any wild-type song.

One to several days prior to an experiment, Equithesin anesthetized birds (Jensen Salsbery: 2 ml/kg intramuscularly) were implanted with a stainless steel post. By pivoting the head around the earbars, the angle of the beak relative to horizontal was adjusted to 15°; the post was then positioned with reference to the bifurcation of the midsagittal sinus and affixed with acrylic cement. The center of the post was taken as a stereotaxic reference point. On the days of experiment, birds were acutely anesthetized with urethane (Sigma: 20%, 0.4 ml/kg *IM*). Body temperature was monitored with a lubricated (KY Jelly: Johnson and Johnson) intercloacal probe and was maintained with external heating. The recordings were stabilized by immobilizing the head via the post while restraining the body in a hammock. Glass-coated platinum/iridium electrodes were advanced with a hydraulic microdrive (Trent Wells) through small openings in the skull positioned at 2.0 mm lateral to the bifurcation of the midsagittal sinus. These procedures enabled recordings from individual units to be maintained for several hours. Electrolytic lesions were placed at the recording sites of most complex units, and at other fiduciary locations. At the end of an experiment, the animal was administered a lethal dose of Equithesin. When deeply anesthetized, the animal was exsanguinated and fixed in 10% formaldehyde via intracardial perfusion. Standard frozen section histology (30 μm sections, cresyl violet staining) enabled the reliable recovery of electrolytic lesions.

Song analyses, on-line song synthesis, and response analyses such as dot rasters and histograms were implemented on a computer based system (PDP 11/40). Birds were placed in small acoustic isolation chambers and were induced to sing by a subcutaneous testosterone implant. Songs were recorded with an analog tape recorder (TEAC). The recorded signals were then low-pass filtered at 9.5 kHz (8 pole filter) and digitized with 12-bit resolution at a 20 kHz sampling rate. Off-line programs extracted frequency and amplitude information for one millisecond intervals and established the starting and ending times for each part or phrase in the song. The amplitude was variable over 80 dB (adjusted to 0 dB = 20 μPa), and distortion products were 40 dB below signal intensity. Sound was delivered in freefield conditions by a small speaker (Lafayette: ± 2.5dB, 1.5kHz to 11 kHz) mounted 1.7 m from the bird inside a 2 X 2 X 2 m sound attenuating chamber. The double walled, fiberglass-lined chamber attenuated room noises by at least 80 dB for frequencies above 300 Hz.

In this study, units typically were tested with tone bursts, narrow-band (600 Hz bandwidth), and wide-band (10kHz bandwidth) noise bursts, the song of the experimental subject (bird's own song), and with the tutoring song(s), available only for laboratory-reared birds. Numerous units were also tested with songs played in reverse. Neuronal responses during stimulus presentations were accumulated as time-of-event data, and displayed by the computer. Graphs were generated on-line to aid in deciding if a unit's responses to songs were predictable on the basis of the responses to simple stimuli. Typically, such predictions could be made on the basis of the tone burst tests. Those units exhibiting complex response properties, namely responses to song that were not predicted by the responses to simple stimuli (see *Discussion*), were investigated further by systematic parameter modification of the songs and by synthesis of complex artificial stimuli. The responses of numerous multiunit recording sites to various songs were recorded on magnetic tape for off-line analysis.

RESULTS

Overview. The response characteristics of 931 units, isolated in HVc and surrounding areas of 33 birds, were analyzed. Of these, 477 units were documented on the computer. Most units either did not respond or responded weakly to song and other sounds. A number of units (approximately 10%) responded vigorously to song in a manner predicted by their responses to tone bursts. These response classes are dealt with in a subsequent section (see *Other response categories*).

The main emphasis of this report is on those few units that showed excitation to song and to complex artificial stimuli mimicking song, but not to simple stimuli such as tone bursts or narrow-band or wide-band noise bursts. The rationale for labeling these units as ``song specific'' is given in detail in *Discussion*. For these units simple stimuli were ineffective or inhibitory; only a subset of songs from the repertoire elicited an excitatory response. Often, but not always, songs played in reverse proved to be ineffective. By elucidating the parameters common only to those natural and modified songs that were excitatory, it was possible on occasion to construct effective artificial stimuli. As will be shown, systematic modification of acoustic parameters within these artificial stimuli varied the units' responses in a manner consistent with the responses to various song parameters within the repertoire. Song specific units were rare (27 well isolated units in 17 animals) and often habituated to stimulus repetition rates of greater than once every five to seven seconds. Thus, when necessary, randomly chosen repetition periods of seven to twenty seconds were used. These intersong intervals fall well within the range of behaviorally relevant intervals for singing white-crowned sparrows (personal observations).

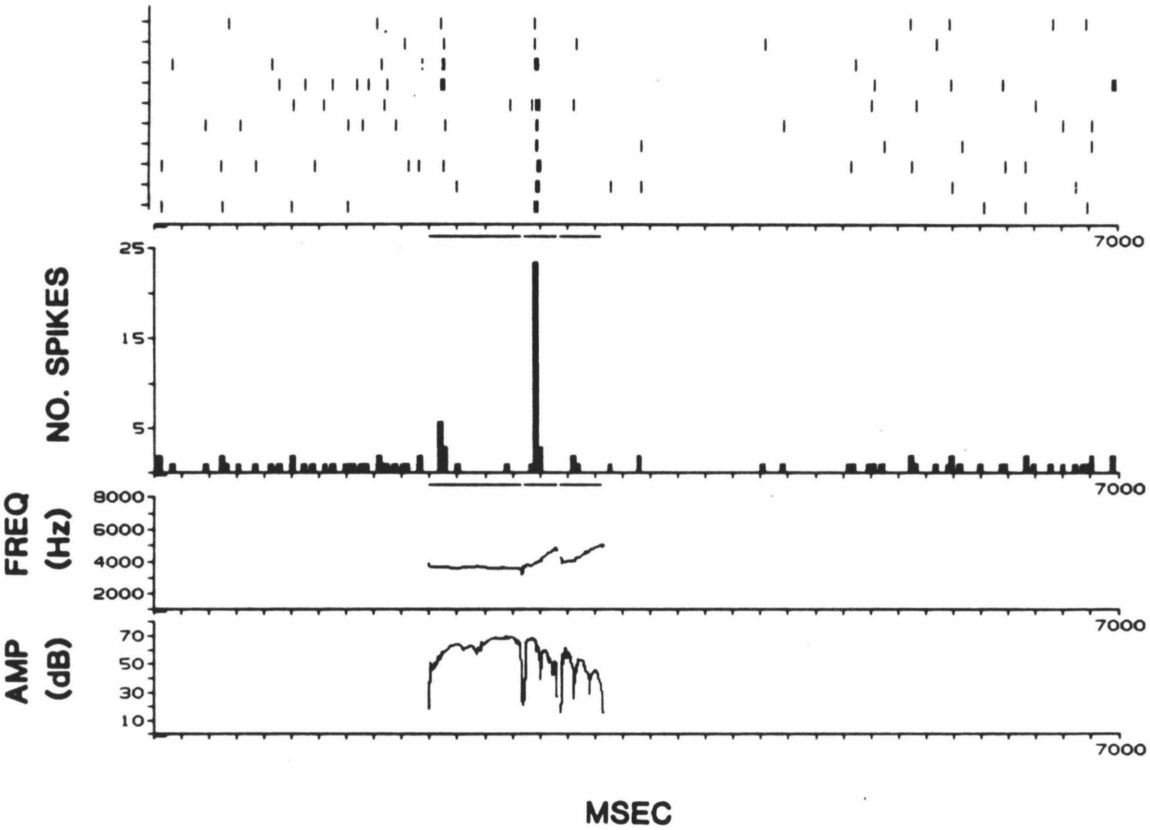
Song specific units were typically encountered within a region that exhibited strong multiunit responses to song, and weaker responses to tone bursts of frequencies contained in the songs. The multiunit responses were prominent, and of a bursting nature. Such properties made the isolation of single units difficult. Combined with the need to maintain recordings for several hours, these conditions contributed to the paucity of song specific units documented. Nevertheless, the data clearly indicate that song specific units comprise a small percentage ($<10\%$) of all robust auditory units in HVc. While sampling bias introduced by the electrode could seriously affect such estimates, it is nevertheless thought provoking to attempt to discern the total number of song specific units in the region. Since each HVc comprises 1.5×10^5 neurons (M. Gurney, personal communication), I estimate 10^3 song specific units in each HVc. These may be further subdivided into several functional classes (see below), so that each functional class comprises only several hundred song specific units.

A striking and consistent observation was that two song phrases or two or more elements in the sequence from the song were required to elicit an excitatory response. For most song specific units, any individual phrase from the appropriate sequence elicited an inhibition (that is, activity lowered below spontaneous rate), no excitation throughout the phrase, and/or weak excitation at the onset or offset. Sometime after the onset of the second phrase in the sequence, however, the unit responded vigorously. Modification of the frequency modulation or center frequency of the various phrases even without modification of the amplitude envelope could totally abolish the response. Temporal characteristics of the sequence strongly affected the response. A number of tests were designed to manipulate the temporal pattern within the song while leaving unaltered the total power spectrum. Changing the order of phrases within the sequence, or reversing the direction of play of the otherwise normal sequence typically abolished the response. Increasing the interval between phrases over hundreds of milliseconds, however, diminished the response only gradually.

The song specific units of a given bird were maximally stimulated by the song of that individual. Other songs were effective only if they contained similar phrases in the appropriate order. Thus, in birds that sang isolate-like songs - songs with abnormal morphologies never found in the wild - it was possible to find units that responded only to the bird's own song, and not to any wild-type song in the repertoire. More remarkable, however, song specific units in wild caught Nuttalli race Bodega Bay adult birds only responded to a few songs within a large repertoire of quite similar songs all of the same dialect. Thus, such units exhibited specificity for song and intra- as well as inter-dialect selectivity.

Response properties of whistle-whistle units. Song specific whistle-whistle units (n=7) responded after the onset of the second of two whistles. Such whistle pairs occur frequently in the songs of Nuttalli race white crowns, and are also common in isolate-like songs. The response of one such unit recorded from a bird that sang an isolate-like song is shown in Figures 4 and 5. The bird's own song consisted of an initial constant frequency phrase followed by two similar phrases of slowly increasing frequency, each of which started with a brief constant frequency segment (see Figure 4). The unit exhibited weak excitation at the onset of the first phrase followed by inhibition, but responded vigorously after the onset of the second phrase (Figure 5*A*). Inhibition followed that lasted throughout the final phrase and approximately two seconds beyond. Reversed song was an inhibitory stimulus (Figure 5*B*), indicating that the response was sensitive to the sequence of frequency or amplitude modulation within the song. While the unit responded *at* the second phrase in the song, it did not respond *to* the second phrase presented alone (Figure 5*C*). When the interval between the first and second phrases was increased even by as much as 300 milliseconds, the response was not abolished, but followed with constant latency the onset of the second phrase (Figure 5*D*). Thus, the unit was not responding solely to the first phrase, and instead required both phrases. The center frequency of the initial segment of the second phrase was essentially identical to the center frequency of the first phrase. With this in mind, the whistle pair was modeled as two

Figure 4. Response of song specific whistle-whistle unit 560107. *Top section* is a dot raster representation of the unit's response. Each vertical tick represents one spike; each row is marked on the ordinate by a horizontal tick and represents the response to one stimulus presentation. *Middle section* is a histogram of the raster; *ordinate*: number of spikes per bin. As for all response rasters and histograms in this paper, the horizontal lines directly underneath the plot are timing marks. *Bottom section* is computer representation - frequency and amplitude versus time - of the stimulus, the bird's own song. The time axes of all plots are aligned.



7000
25
15
5
NO. SPIKES
7000
8000
6000
4000
2000
FREQ (Hz)
7000
70
50
30
10
AMP (dB)
7000
MSEC

successive tone bursts of identical frequencies. The duration of the tone bursts and the interval between them were taken from the song. These successive tone bursts of identical frequencies elicited a weak excitation at the onset of the first tone followed by inhibition, and a strong excitation at the onset of the second tone followed by inhibition (Figure 5*E*). Note the similarity with the response to the song (Figure 5*A*). The first tone burst was by itself ineffective (Figure 5*F*). Thus, the response properties of this unit were unequivocally not predictable on the basis of its responses to single tone bursts. The duration of the first of the tone bursts was not critical for the response. For instance, the two tone bursts could have identical durations (and all other parameters) yet the unit would respond only at the second tone burst of the pair (Figure 5*G*). Note also that the ineffective single tone burst stimuli (Figure 5, *F* and *H*) were intermediate in power to effective tone burst pairs (Figure 5, *E* and *G*). In Figure 5*I*, the actual stimulus trace demonstrates the clear single unit nature of the recording.

These observations typified some features of song specific whistle-whistle unit response properties. The response could be elicited by consecutive tone bursts and thus did not require the fine frequency ``jitter'' from which a bird whistle derives its distinctive sound quality. Within the constraints of consecutive tone bursts, the timing information, namely duration of either of the tone bursts or the interval, was not critical to the unit's response. In fact, the timing information in the song was not necessarily optimal for evoking a response. Typical inter-tone interval tuning curves were broad, and were not peaked around the interval in the bird's own song (Figure 6). The maximum interval eliciting a response at the second tone could be several hundreds of milliseconds and was never less than approximately one hundred milliseconds (Figure 6). Fine features of the amplitude modulation as found in the song, as well as the overall power, were not critical.

Figure 5. *A*, As Figure 4. *B*, Histogram and reversed song stimulus only. *C*, Second phrase in isolation. *A,B,C*, Time axes 7000 msec. *D*, Histogram only for three stimuli: first and second phrase with 25, 125, 325 msec intervals. The arrows point to the onset of the second phrase. Note time axis is 2000 msec. *E* to *H*, Two tone and single tone models; time axes 4000 msec. *I*, Actual recording trace of unit response (*top trace*, arrow) to stimulus (*bottom trace*, bird's own song). Unit is well isolated. All histograms 10 repetitions, 30 msec binwidth.

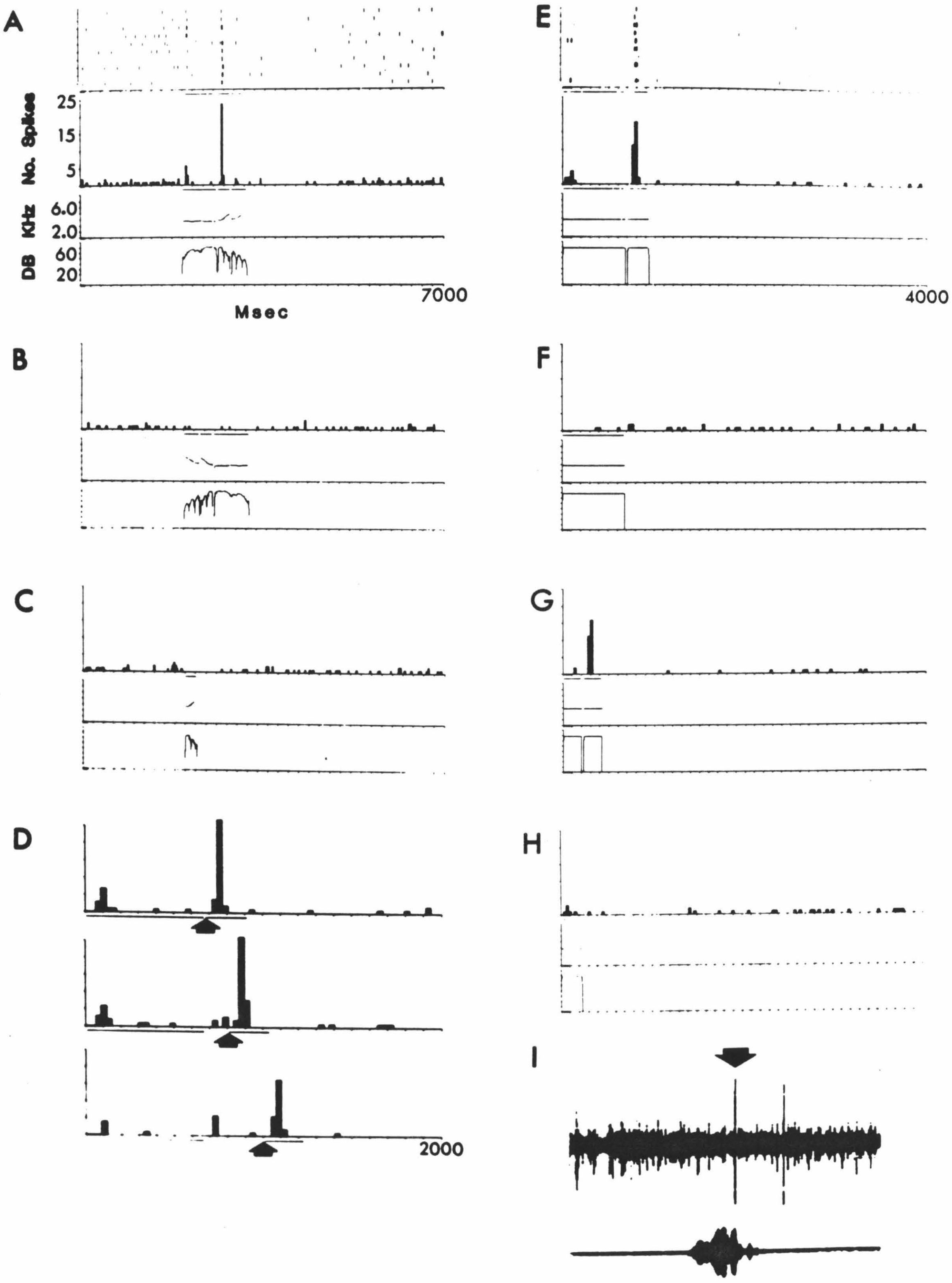
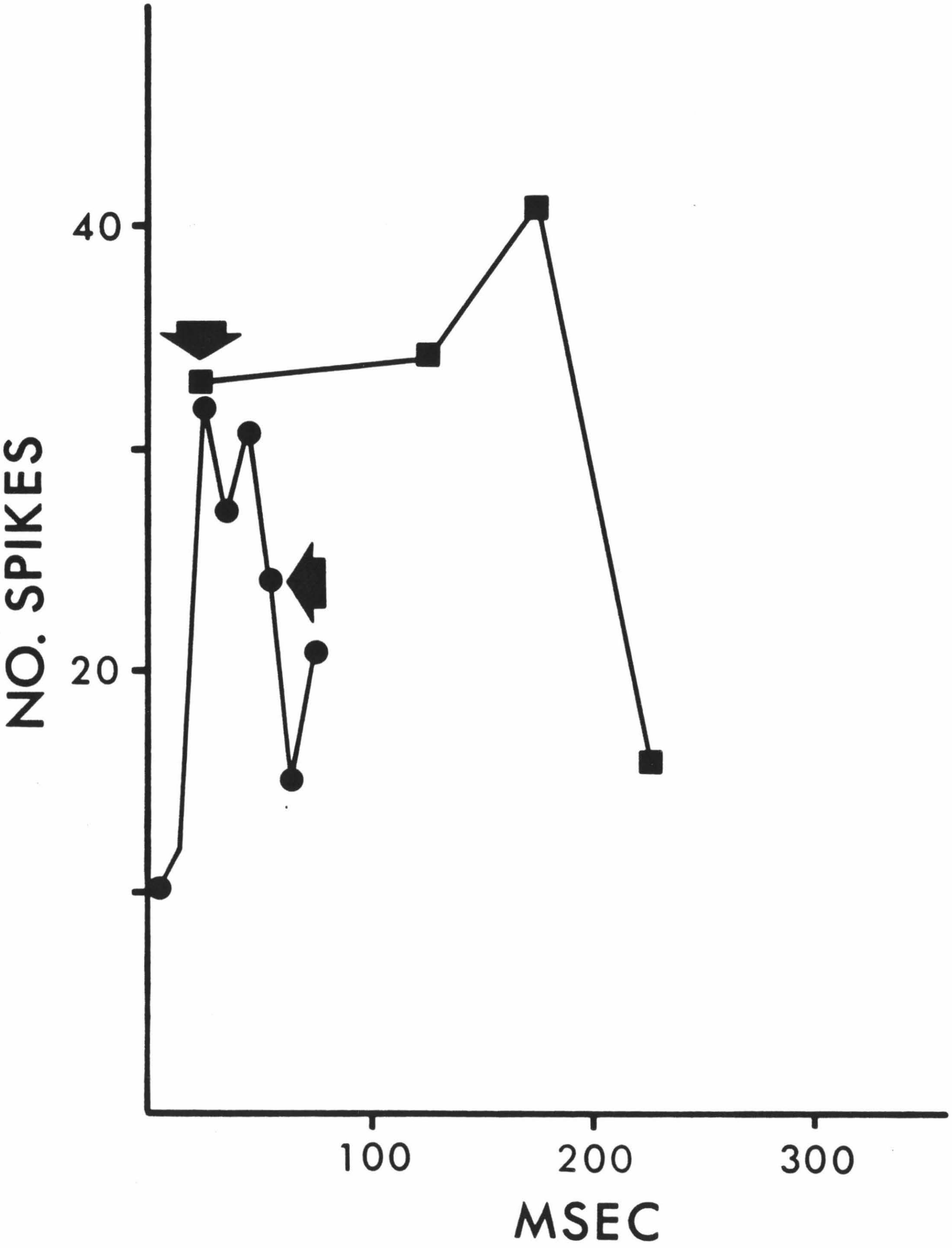

A
No. Spikes
25
15
5
KHz
6.0
2.0
DB
60
20
7000
Msec
B
C
D
2000
E
4000
F
G
H
I

Figure 6. Inter-tone interval tuning. Each point represents the total number of spikes at the second of two tones as a function of the varying interval between the tones. 10 repetitions per point. *Arrows* point to the natural interval between corresponding whistles in the birds' own songs. The *leftmost* curve (*circles*) is the tightest inter-tone interval tuning seen to date. The *broader* curve (*squares*), for another unit in another bird, is a more typical example.


40
20
NO. SPIKES
100
200
300
MSEC

Song specific units recorded in birds that sang abnormal, isolate-like songs exhibited considerable selectivity within a repertoire of naturally occuring songs. The most selective of these units responded only to the bird's own song, and not to any of the six tutored songs ("choice" paradigm: five sympatric species and one white crown song) to which it was exposed early in life, nor to any of nine other wild-type white crown songs tested. Song specific whistle-whistle units were also recorded in birds that sang wild-type songs. The response of one such unit recorded in a bird singing a song of the Bodega Bay dialect is shown in Figure 7. The initial phrase, comprising two whistles, elicited a robust response at the second whistle of the pair, followed by approximately two seconds of inhibition (Figure 7*A*). In isolation, the second whistle of the phrase elicited no response (Figure 7*B*), whereas the first whistle elicited only weak excitation at the offset (Figure 7*C*). The unit was inhibited by single tone bursts (Figure 7*D*), but was effectively stimulated by consecutive tone bursts of frequencies matching the center frequencies of the two whistles in the phrase (Figure 7*E*). As predicted by the single tone burst tuning, individual elements of the model were ineffective (Figure 7*F*).

The response decreased monotonically with variation in the frequency of either of the tone bursts of the model. For instance, the response was abolished by lowering the frequency of the second tone by 1 kHz while leaving all other parameters of the model unmodified (Figure 7*G*). In this way, it was possible to quantify the response of whistle-whistle units as a function of the frequency of either tone burst (Figure 8). This will be termed "leading", "following", or "dual consecutive tone tuning," depending on which tone burst is being referred to. No systematic differences in tuning were observed for leading versus following consecutive tone tuning within the small sample size recorded (n=6). Consecutive tone tuning for song specific whistle-whistle units confirmed that the optimal frequencies for either of the tone bursts of the model typically

Figure 7. Response of whistle-whistle unit 810102. *A*, Stimulus is the bird's own song. Note the response at the second whistle of the first phrase, followed by approximately two seconds of inhibition. *B*, Second whistle of first phrase. *C*, First whistle of first phrase. Note weak response at offset. *D*, Response to single tone bursts. ***Top raster*** is response to individual tone bursts 3.2 kHz to 4.6 kHz in 100 Hz steps, covering range of frequencies in the first phrase of the song. At 3.5 and 3.6 kHz unit seems to exhibit a slight offset excitation. ***Bottom raster*** is 10 repetitions of 3.6 kHz tone burst demonstrating no offset excitation. Frequency range of 500 Hz to 10 kHz in 500 Hz steps was also tested - no response (not shown). *E*, Consecutive tone model. Note response at second tone. *F*, Second tone of model in isolation. First tone was also tested - no response (not shown). *G*, Consecutive tone model as in *E* but second tone is decreased in frequency 1 kHz. *H*, Actual recording trace of unit's resonse (***top trace***, arrow) to the isolated first phrase (***bottom trace***). Good single unit isolation. *A* to *G*: time axes 7000 msec; all histograms 10 repetitions, 30 msec binwidth.

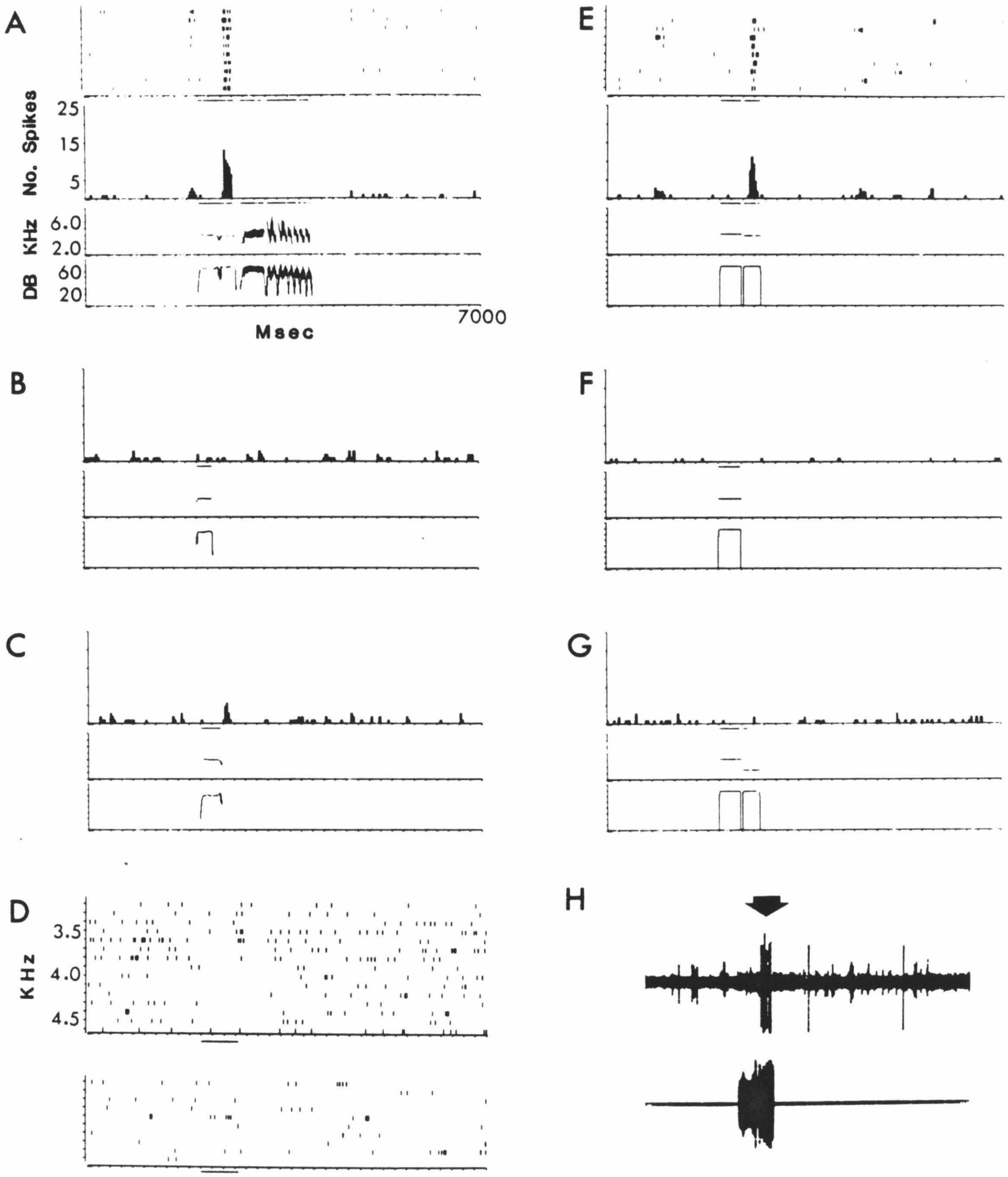

A
No. Spikes
25
15
5
KHz
6.0
2.0
DB
60
20
Msec
7000
B
C
D
3.5
KHz
4.0
4.5
E
F
G
H

Figure 8. Consecutive tone tuning curves. Leading consecutive tone tuning for a song specific whistle-whistle unit (*circles*). Following consecutive tone tuning for same unit (*triangles*) and for another unit (*squares*) - see text. The leading consecutive tone tuning curve represents an example of broad tuning, while the following consecutive tone tuning curves are examples of sharp tuning. The ***dashed portion*** of each curve spans the range of frequencies present in the corresponding whistles from the original songs.

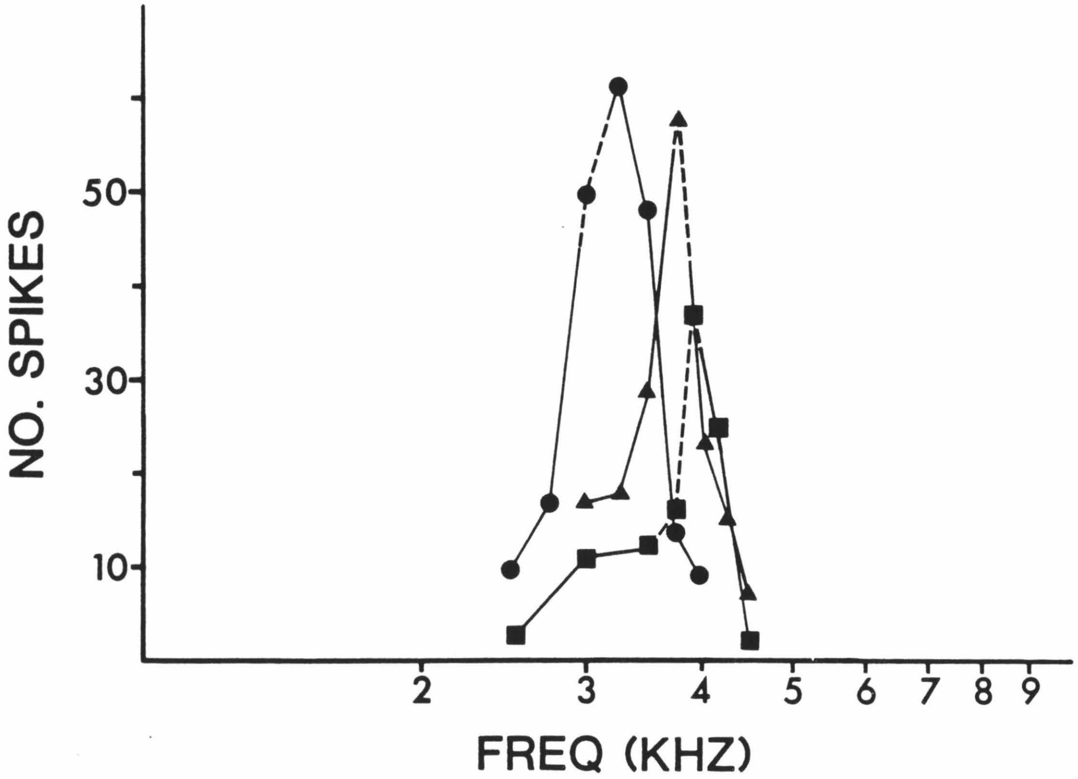

NO. SPIKES
50
30
10
2
3
4
5
6
7
8
9
FREQ (KHZ)

Figure 9. Song selectivity of unit 810102. Sonagrams of ten white- crowned sparrow songs with response histograms superimposed. The five songs in the left column are all of the same Bodega Bay dialect; top song is bird's own song. Note only one other song elicits strong excitation. Five songs in right column are sample songs from five different dialects of the same race of white crowns (*top to bottom*: Berkeley, Tilden Park, Pacific Grove, Lompoc, Lake Merced, California dialects). Note no song elicits strong excitation. All histograms 20 repetitions, 30 msec binwidth. The calibration mark in *lower right* represents 10 spikes.

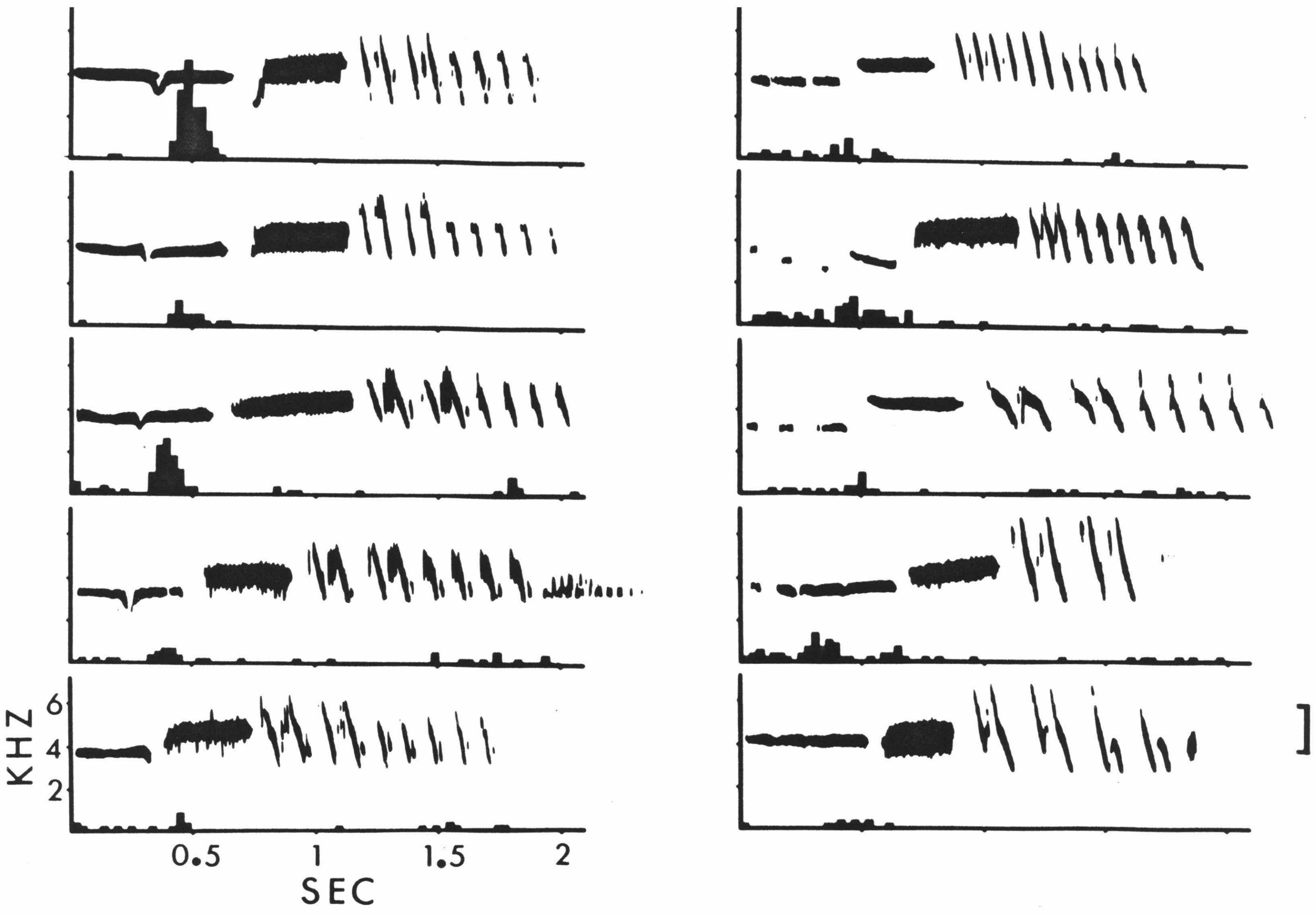
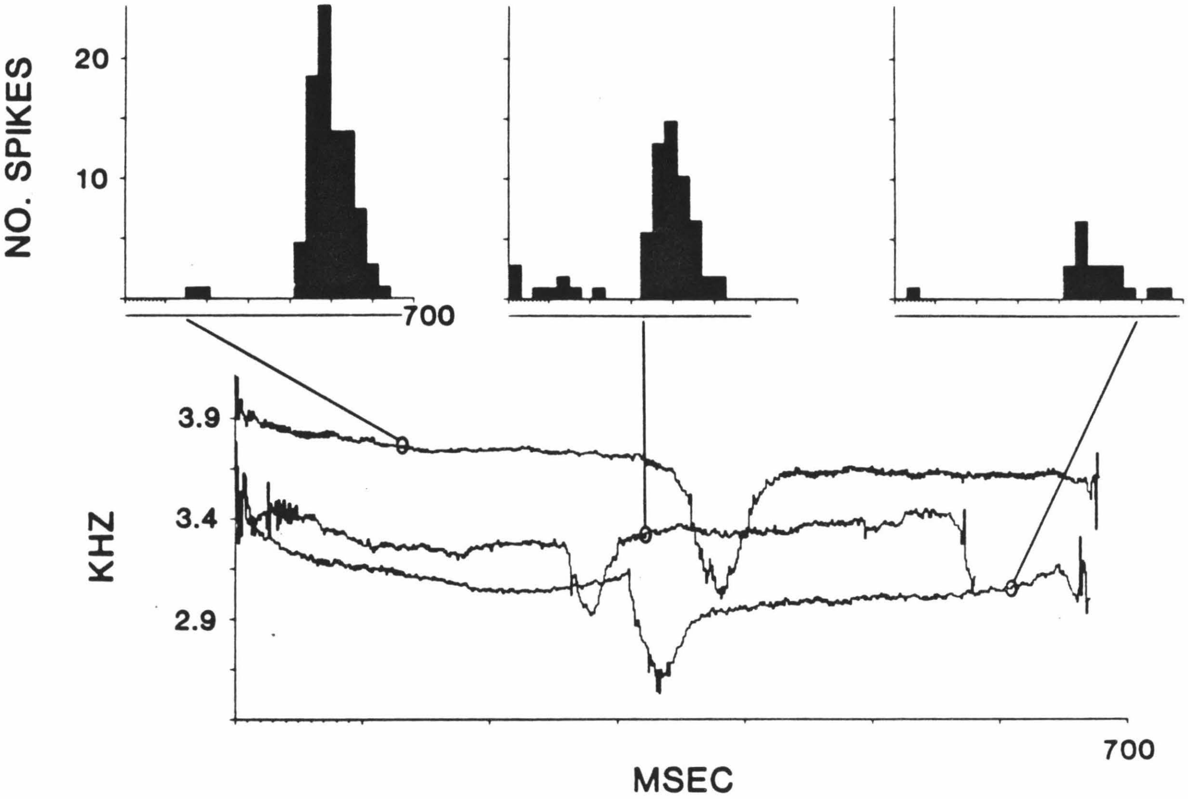

KHZ
6
4
2
0.5
1
1.5
2
SEC

Figure 10. Parameters underlying song selectivity of unit 810102. Response histograms (20 repetitions, 30 msec binwidth) and corresponding frequency vs. time plots for three Bodega Bay dialect phrases. *Leftmost histogram* is response to phrase in bird's own song. The dip in frequency at the middle of each phrase corrseponds to the dip in amplitude. Note that the one other phrase that elicited a strong response has similar frequency, but not timing, parameters.



NO. SPIKES
20
10
700
KHZ
3.9
3.4
2.9
700
MSEC

fell within the range of frequencies present in the corresponding whistles of the natural (song) stimulus (Figure 8).

The selectivity of the unit of Figure 7 within a group of wild-type songs is shown in Figure 9. Out of five Bodega Bay dialect songs, the bird's own song was most effective, and only one other song proved strongly excitatory. Eight other songs tested, including five samples of other Nuttalli dialects, elicited weak or essentially no excitation. These observations are consistent with the stimulus requirements as outlined above. That is, the two songs that evoked strong responses contained initial phrases comprising two whistles of appropriate frequencies (Figure 10). The strength of response varied systematically with the frequencies of the two whistles, as predicted by the consecutive tone tuning for this unit. Differences in duration of the various whistles and intervals did not account for the response selectivity (Figure 10). Five songs had whistles of approximately the same intensity as the whistle of the bird's own song; intensity did not account for the response selectivity. Those songs that contained single whistle phrases, even of the appropriate center frequency, did not elicit a response. An example is the song of Figure 9, ***fifth row, first column***, center frequency = 3.8 kHz. For song specific whistle-whistle units, all ineffective songs lacked one or more essential feature, namely two whistles, or the appropriate frequencies. Significant variation in timing information did not affect the response.

Those whistle-whistle units that responded optimally to consecutive tones of different frequencies exhibited an interesting form of two tone inhibition. For example, the unit of Figure 11*A* responded at the second phrase of the isolate-like song of the bird but did not respond to either of the two whistle phrases in isolation (Figure 11, *B* and *C*). While consecutive tones evoked a response at the second tone (Figure 11*D*), either tone in isolation was ineffective (Figure 11, *E* and *F*). When, in addition to the normal consecutive tone model, the second tone was played with the first tone (Figure 12*A*), or the first tone was played with the second tone (Figure 12*C*), the response was abolished. Since two tone bursts of equal intensity played simultaneously produce a louder overall sound, control stimuli with the first (Figure 12*B*) or second (Figure 12*D*) tone burst 6 dB louder were tested. These control stimuli were effective, and thus the two tone inhibition is not a result of an increase in stimulus intensity. The various observations outlined above are summarized in the discussion, and a model is proposed.

Response properties of other song specific units. Song specific units were not restricted to a single response category. Combinations of phrases – whistle-buzz, buzz-trill, and buzz-buzz – as well as combinations of elements within single trill phrases, elicited responses from units exhibiting considerable selectivity and specificity. The complexity of stimulus requirements for these units complicated their detailed analysis. Within the small sample size (20 units for all classes) the response properties of few units were fully explored. However, to the extent that these units do not respond to simple consecutive tone models, they must represent one or more different classes of response.

An example of a song specific buzz-trill unit is shown in Figure 13. The Bodega Bay dialect song of this subject evoked a strong response at the onset of the trill, the third phrase in the song, followed by approximately two seconds of inhibition (Figure 13*A*). Eliminating the whistle did not impair the response (Figure 13*B*). Although the unit exhibited inhibition for the duration and/or weak excitation at the offset of tone bursts (Figure 13*C*), it clearly responded to the isolated buzz at the offset (Figure 13*D*). The isolated trill was only weakly excitatory (Figure 13*E*). Attempts to stimulate this unit with consecutive tone models were unsuccessful. This particular unit showed the strongest response to an isolated phrase of any song specific unit tested. Nevertheless, the

Figure 11. Response of whistle-whistle unit 840601. *A*, Strong resonse to bird's own song. *B*, Weak offset resonse or no response to first phrase in isolation. *C*, No response to second phrase in isolation. *D*, Strong response to consecutive tone model. Note similarity to *A*. *E*, Weak offset response or no response to first tone of model in isolation. Note similarity to *B*. *F*, No response to second tone of model in isolation. Note similarity to *C*. *A* to *F*: time axes 7000 msec; all histograms 20 repetitions, 30 msec binwidth. *G*, Actual recording trace of unit's response (***top trace***, arrow) to bird's own song (***bottom trace***).

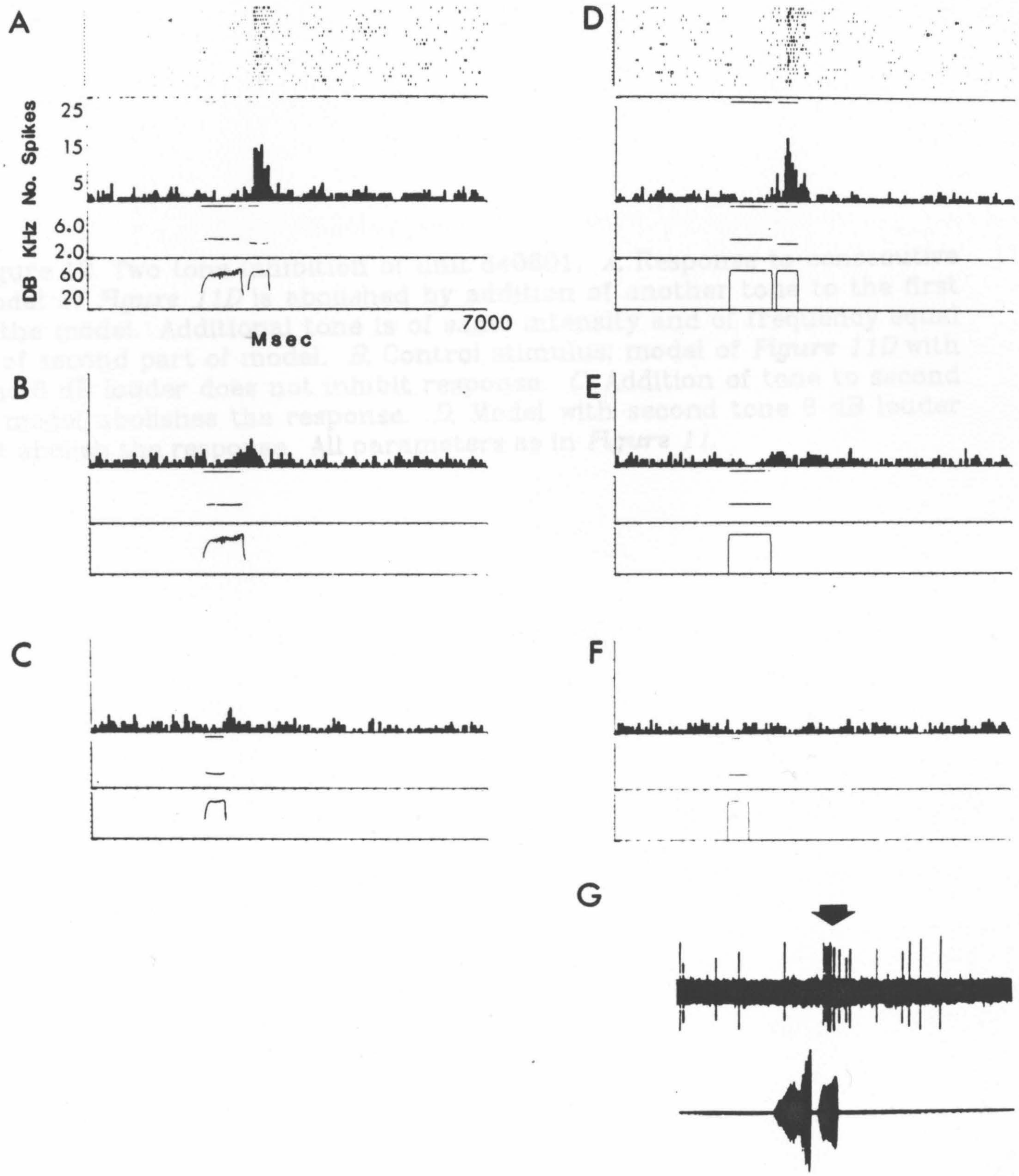
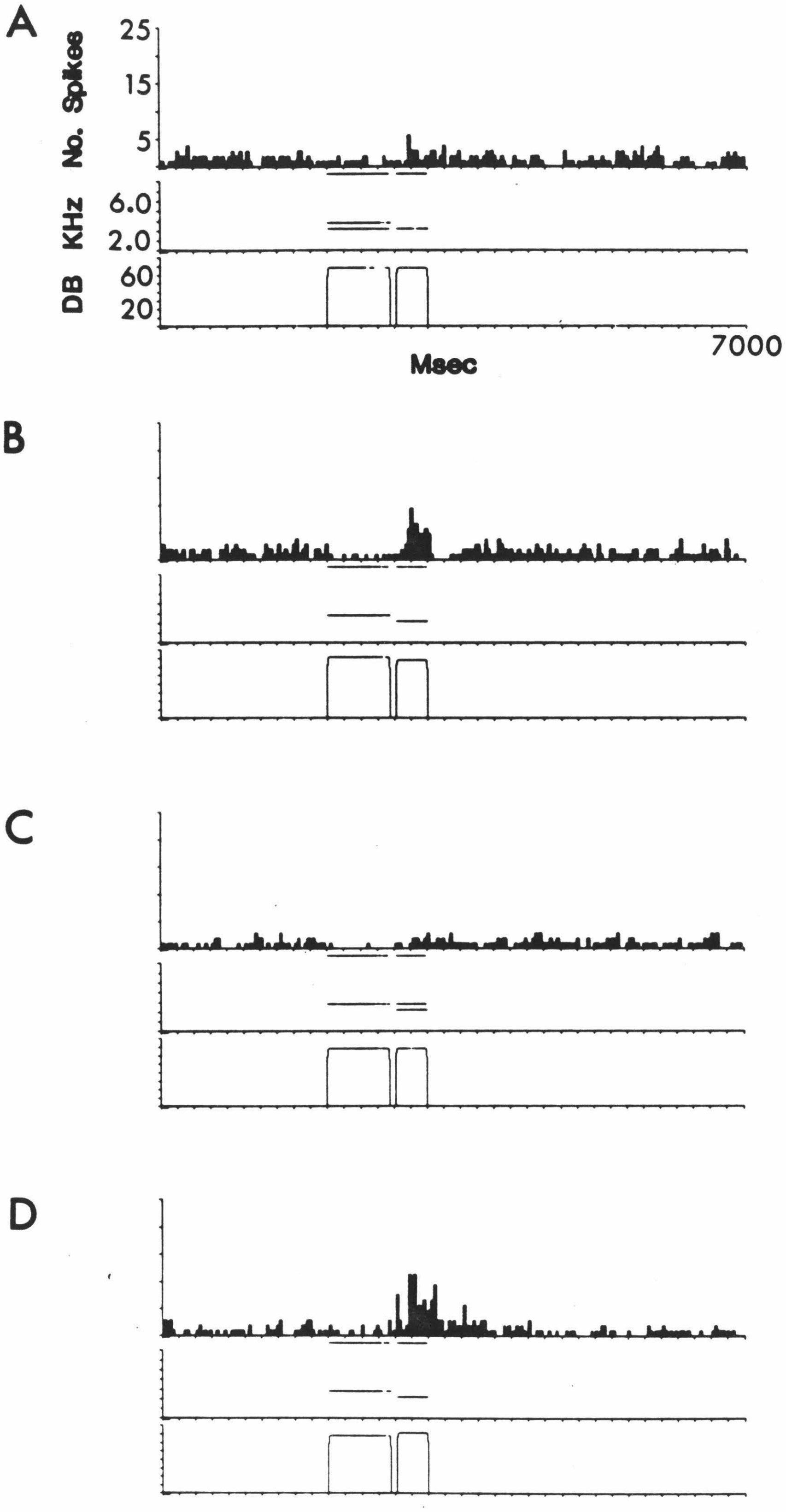

A
No. Spikes
25
15
5
KHz
6.0
2.0
DB
60
20
7000
Msec
B
C
D
E
F
G

Figure 12. Two tone inhibition of unit 840601. *A*, Response to consecutive tone model of *Figure 11D* is abolished by addition of another tone to the first part of the model. Additional tone is of same intensity and of frequency equal to tone of second part of model. *B*, Control stimulus; model of *Figure 11D* with first tone 6 dB louder does not inhibit response. *C*, Addition of tone to second part of model abolishes the response. *D*, Model with second tone 6 dB louder does not abolish the response. All parameters as in *Figure 11*.


A
25
15
5
No. Spikes
6.0
2.0
KHz
DB
60
20
7000
Msec
B
C
D

response to the buzz in isolation showed some diminution in response strength and considerably greater temporal variation than the response to the buzz-trill combination. No song specific unit responded as vigorously to an isolated phrase as it did to that phrase within its appropriate natural sequence.

Several song specific units responded to the trill of the bird's own song (n=9). The most extensively documented example is shown in Figure 14. The unit responded during the trill of the bird's own song (Figure 14*A*), and equally well to the isolated trill (Figure 14*B*), comprising three elements. Reversed trill (Figure 14*C*), and tone or noise bursts (Figure 14*D*), were ineffective stimuli. Slight filtering of the trill frequencies while leaving the amplitude function unaltered caused a slight decrease in the response (Figure 14*E*), and extensive frequency filtering abolished the response (Figure 14*F*). The stimulus requirements of this unit were quite closely matched to the trill parameters. Of six artificial models tested, only two (e.g. Figure 14*G*) produced a moderately strong response. These two models were derived as follows. The pictorial representation of the natural trill, comprising three elements, was magnified to fill the entire graphics screen. The experimenter could then precisely choose fiduciary points with a cursor to achieve an accurate piecewise linear approximation of the trill frequencies. Two other models that were ineffective in eliciting an excitatory response were of sufficiently similar acoustic parameters as to make the sounds hard to distinguish by the experimenter, a trained listener. These two models were not directly derived from the natural trill. Rather, they were versions of the two effective models simplified by reducing the number of segments in the piecewise linear approximation of the frequency modulations. Finally, two other models were also ineffective. They were identical to the effective models except in that they lacked the initial element. Elimination of the initial element of an effective model totally abolished the response (Figure 14*H*). Clearly this unit exhibited exquisite response specificity. While numerous tests

Figure 13. Response of buzz-trill unit 810402. *A*, Response to bird's own song. *B*, Response to buzz-trill combination. *C*, Response to single tone bursts. ***Top raster*** 2.0 kHz to 7.8 kHz in 200 steps, covering the range of frequencies of buzz and trill. Note weak response at offset of 4.8 and 5.0 kHz tone bursts. Frequency range of 500 Hz to 10 kHz in 500 Hz steps was also tested; no other response (not shown). ***Bottom historgram*** 20 repetitions of a 4.8 kHz tone burst. Note weak response at offset. *D*, Response to isolated buzz. Note response at offset is weaker and exhibits greater temporal variation than buzz-trill combination (*A* or *B*). *E*, Very weak response to trill. *A* to *E*: time axes 7000 msec; all histograms 20 repetitions, 30 msec binwidth. *F*, Actual recording showing isolation; response (***upper trace***) is to buzz (***lower trace***).

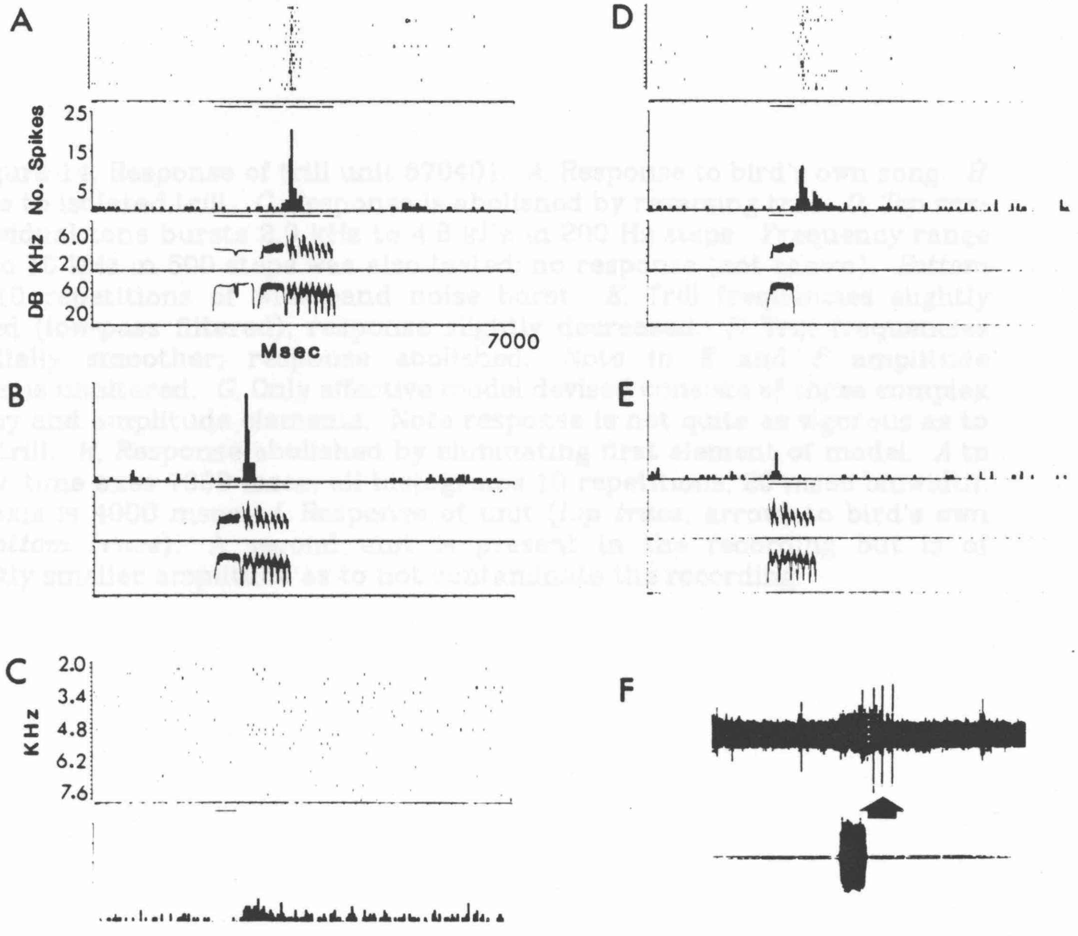
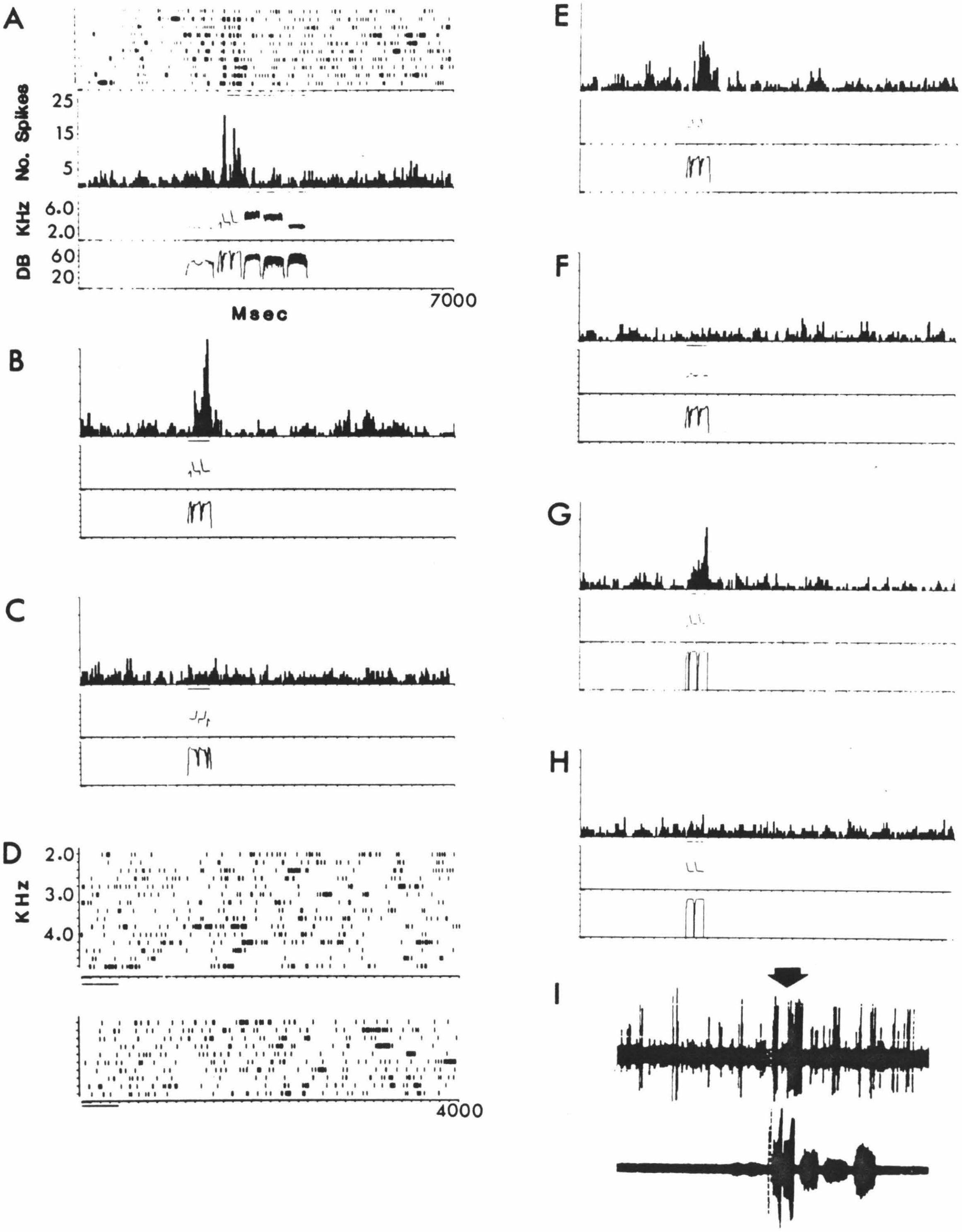

A
D
No. Spikes
25
15
5
KHz
6.0
2.0
DB
60
20
Msec
7000
B
E
C
2.0
3.4
4.8
6.2
7.6
KHz
F

Figure 14. Response of trill unit 570401. *A*, Response to bird's own song. *B*, Response to isolated trill. *C*, Response is abolished by reversing trill. *D*, *Top raster*: individual tone bursts 2.0 kHz to 4.8 kHz in 200 Hz steps. Frequency range 500 Hz to 10 kHz in 500 steps was also tested; no response (not shown). *Bottom raster*: 10 repetitions of wide-band noise burst. *E*, Trill frequencies slightly smoothed (low-pass filtered); response slightly decreased. *F*, Trill frequencies substantially smoother; response abolished. Note in *E* and *F* amplitude waveform is unaltered. *G*, Only effective model devised consists of three complex frequency and amplitude elements. Note response is not quite as vigorous as to natural trill. *H*, Response abolished by eliminating first element of model. *A* to *C*, *E* to *H*: time axes 7000 msec; all histograms 10 repetitions, 20 msec binwidth. *D* time axis is 4000 msec. *I*, Response of unit (*top trace*, arrow) to bird's own song (*bottom trace*). A second unit is present in the recording but is of sufficiently smaller amplitude as to not contaminate the recording.


A
No. Spikes
25
15
5
KHz
6.0
2.0
DB
60
20
7000
Msec
B
C
D
KHz
2.0
3.0
4.0
4000
E
F
G
H
I

suggested themselves, the unit was lost shortly after the control series of Figure 14*A*, and thus the essential acoustic parameters underlying the response were not fully determined. However, the data demonstrate that complex frequency modulations, as well as sequences of elements, were required to elicit excitation.

This and other song specific trill units exhibited significant selectivity for song. In a remarkable example, a bird exposed to the ``choice'' paradigm during the critical period eventually sang a song very similar to the white crown song of the tutor repertoire. The bird's own song differed from the original tutored song only in slight detail in the trill section. Nevertheless, a song specific trill unit isolated in that bird gave robust responses to the bird's own song, while remaining virtually unresponsive to playback of the original tutor song.

The song specific buzz-buzz unit of Figure 15 responded at the terminal buzz of the bird's own song, a Gambelii or Alaskan dialect song (Figure 15*A*). Reversing the song abolished the response (Figure 15*B*). Since individual buzzes are represented as rather symmetrical functions in time, this suggested that the unit was responding to a sequence of buzzes, which is not symmetrical. Indeed, although tone bursts and narrow-band and wide-band noise bursts did not drive the unit (Figure 15*C*), the final two buzzes in isolation elicited a strong response (Figure 15*D*). As could be predicted, neither of those two buzzes in isolation elicited a response (Figure 15, *E* and *F*). Because realistic but artificial buzzes were not available at the time of this experiment, attention was focused on this unit's response selectivity. Out of seven other songs, three Gambelii samples and four Oriantha or montane race dialects, only one elicited a response. This song was the only other song tested of the same dialect as the bird's own song (Figure 16). Thus, this unit exhibited clear interdialect selectivity. Because the response properties of this unit were not fully delineated

Figure 15. Response of buzz-buzz unit 600301. *A*, Response to bird's own song. *B*, Reversed song abolished the response. *C*, *Top*, *middle*, *bottom rasters*: tone bursts, narrow-band (600 Hz) noise bursts, wide-band noise bursts, respectively - no response. 2.0 kHz to 5.75 kHz in 250 Hz steps, covering frequency range of last two buzzes of the song. Frequency range of 500 Hz to 10 kHz in 500 Hz steps for tone bursts was also tested - no response (not shown). *D*, Last two phrases in isolation. *E*, Forth phrase in isolation. *F*, Fifth phrase in isolation. *A* to *F*: time axes 7000 msec; all histograms 10 repetitions, 30 msec binwidth. *G*, Good single unit response (arrow) to bird's own song. This trace was chosen to emphasize the strong multiunit activity song elicits.

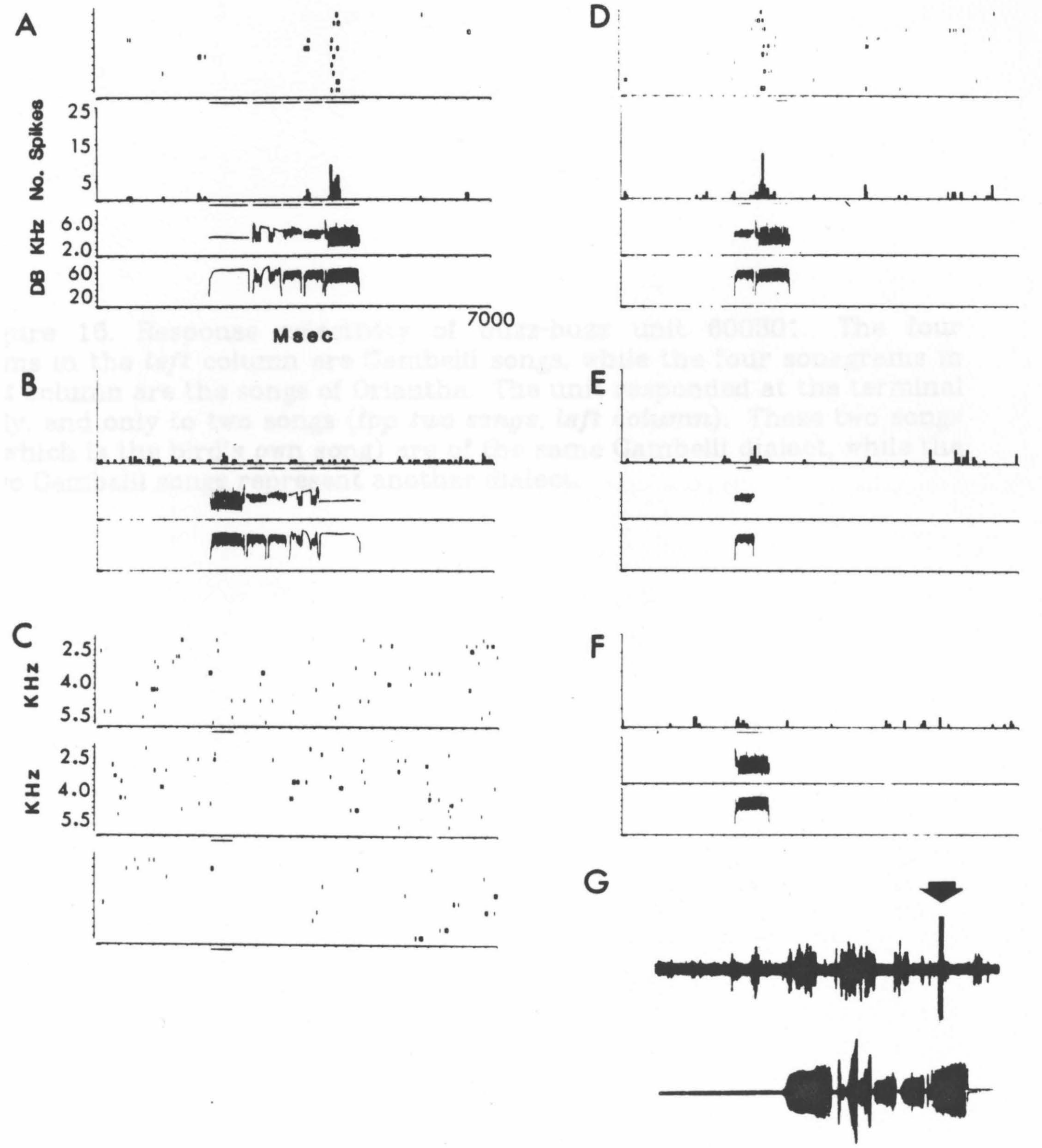
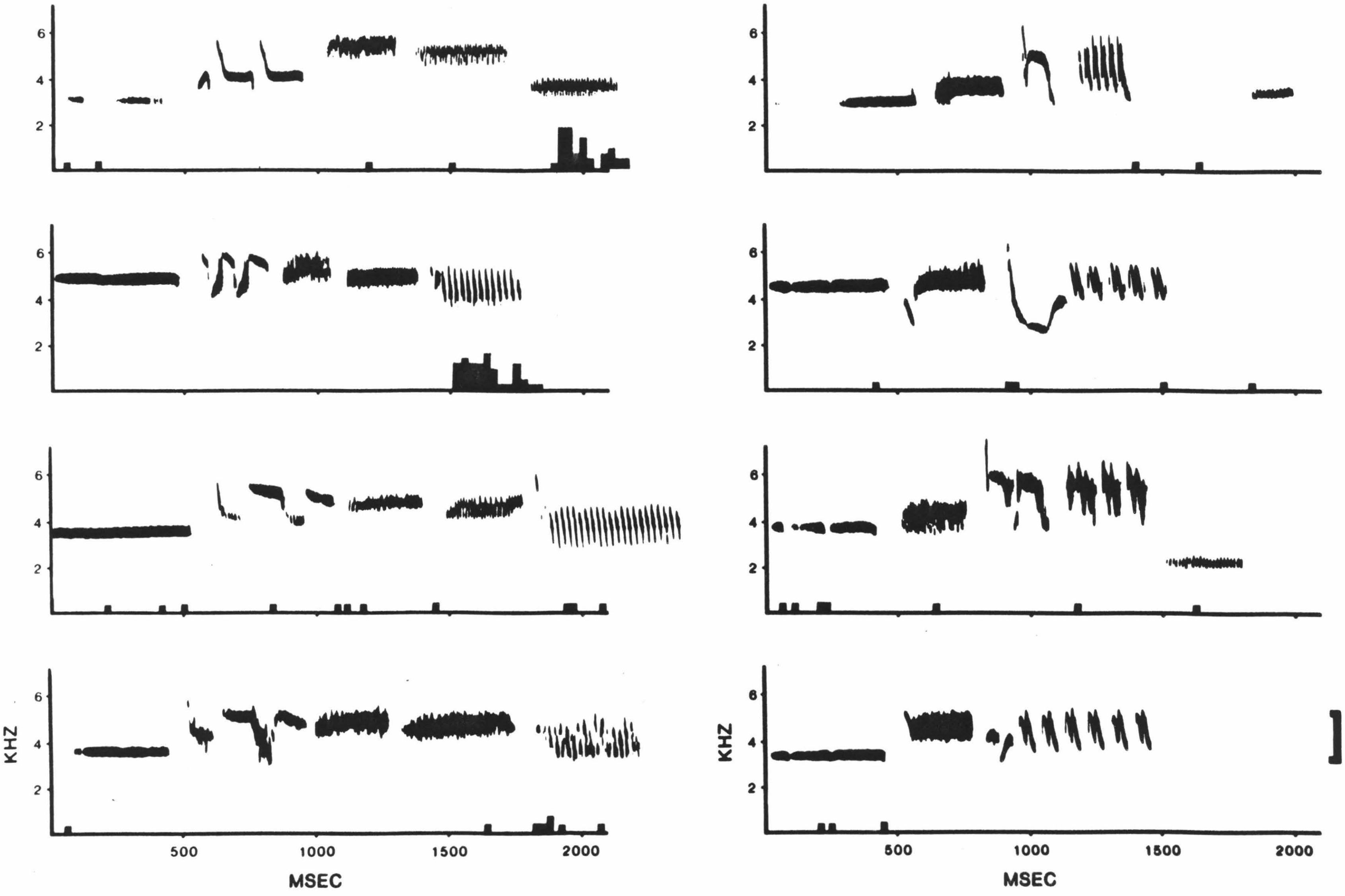

A
No. Spikes
25
15
5
KHz
6.0
2.0
DB
60
20
7000
Msec
B
C
KHz
2.5
4.0
5.5
KHz
2.5
4.0
5.5
D
E
F
G

Figure 16. Response selectivity of buzz-buzz unit 600301. The four sonagrams in the *left* column are Gambelii songs, while the four sonagrams in the *right* column are the songs of Oriantha. The unit responded at the terminal buzz only, and only to two songs (*top two songs, left column*). These two songs (one of which is the bird's own song) are of the same Gambelli dialect, while the other two Gambelii songs represent another dialect.

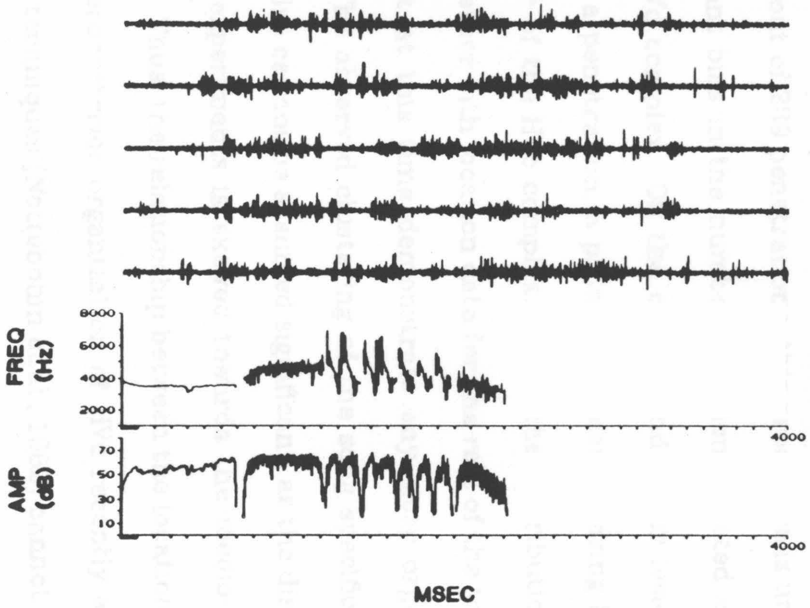

6
4
2
KHZ
500
1000
1500
2000
MSEC

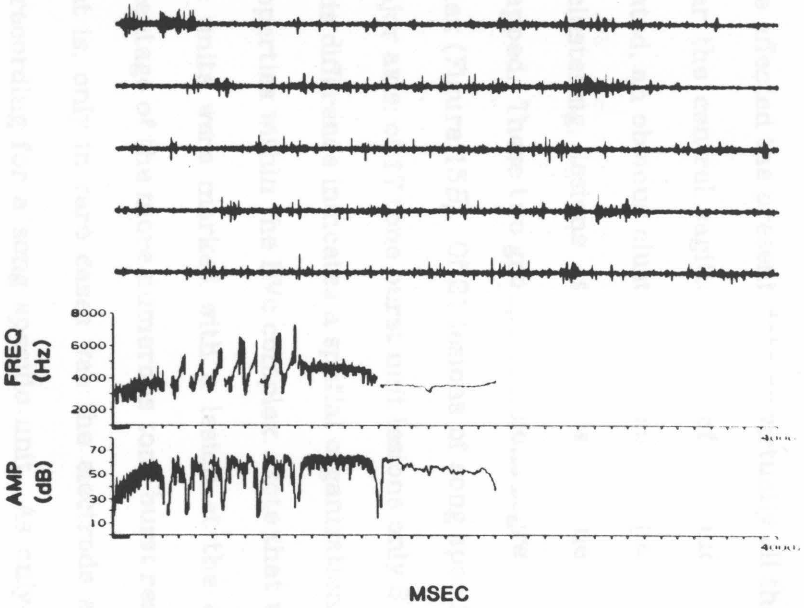
with artificial stimuli, it is only possible to speculate on the parameters underlying the unit's song selectivity. However, one significant similarity that only the two effective songs share is the frequency range of the terminal buzzes in those two songs. Taken as a whole, these observations suggest the existence of several classes of song specific responses, however, a firm conclusion awaits the complete elucidation of the response properties of such units.

Multiunit responses. Electrode penetrations through the HVc complex (see Figure 18) often encountered regions with strong background multiunit activity, often with physiologically distinct borders. At these recording sites the considerable spontaneous activity was quite variable and of a bursting nature. Within these regions acoustic stimuli often evoked strong responses. For some recording sites, tone bursts were effective stimuli, in others song was far more effective. Multiunits responding to song often exhibited surprising specificity. Reversed song was less effective and on occasion appeared to inhibit the spontaneous activity (e.g. Figure 17). Of various songs drawn from the repertoire, only those similar to the bird's own song elicited a strong response. Thus, song selectivity was also evident in multiunit recordings. Unfortunately, the inherent inability to clearly quantify such recordings limited any interpretations (see Discussion).

Anatomical distribution of functional properties. Of 27 song specific units, 22 were isolated near multiunit sites that yielded responses clearly selective for the bird's own song within the song repertoire. In an attempt to map the associated lesion data, it was observed that HVc is shaped as an ellipsoid, with the major axis parallel to the mid-sagittal plane (Figure 18*A*). Sagittal sections at any medio-lateral position generate an elliptical shape for HVc. Lesion locations thus can be mapped onto an ellipse while preserving relative position. While this mapping technique could in principle distort the distribution of locations near

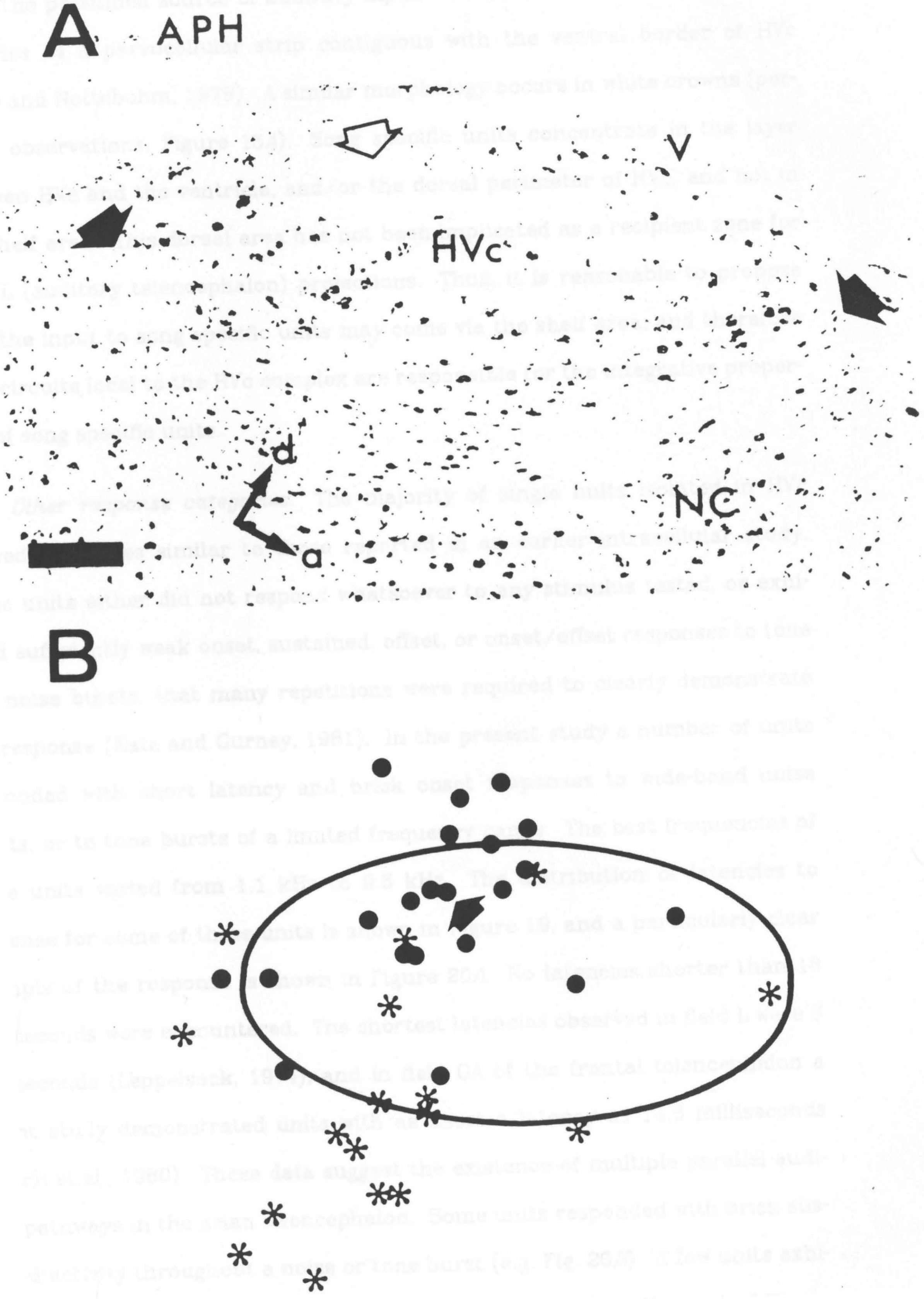
Figure 17. Specificity of multiunit responses. The five ***upper*** traces in each column represent multiunit responses to the stimulus in the ***lower*** part: ***left*** column is bird's own song; ***right*** column is same song played backwards. Note strong, time-locked activity to forwards song, while reversed song elicits inhibition during the song and some excitation at the offset of the song. Responses are from a continuous recording with an intersong period of 7 seconds. Only 4 seconds of data per repetition are shown so as to maintain clarity.


FREQ (Hz)
8000
6000
4000
2000
4000
AMP (dB)
70
50
30
10
4000
MSEC


FREQ (Hz)
8000
6000
4000
2000
AMP (dB)
70
50
30
10
MSEC

the lateral borders of HVc, where the shape and extent of the nucleus are not clear, this is unlikely to have affected the present data as virtually all the lesions were confined to areas near the central sagittal plane of the nucleus. When lesions of units were so treated, an obvious clustering became apparent (Figure 18*B*). To emphasize this clustering, lesions associated with tone burst unit recording sites were also mapped. These two groups of lesions segregated themselves within the HVc complex (Figure 15*B*). Of 21 lesions of song specific units, 19 occur at or above the major axis; of 17 tone burst unit lesions only 3 occur at or above the major axis. This difference indicates a spatial organization of units with differing functional properties within the HVc complex. Note that while the sites of most song specific units were marked with a lesion at the end of a recording, only a small percentage of the more numerous tone burst responding units were so marked. That is, only in rare cases was the electrode advanced after lesioning the site of recording for a song specific unit. As only 27 song specific units were isolated out of 239 penetrations, this procedure is unlikely to have introduced a significant bias in the number of lesions located along the dorsal perimeter of the HVc complex. On the other hand, other lesions were often placed at the end of a penetration, a procedure clearly biasing locations towards the ventral border of the HVc complex. Thus, the distribution of the tone burst unit lesions, together with location data for the rest of the tone burst responding units, does not at this time demonstrate any clear organization within the HVc complex. The observed clustering of the song specific units in the other two dimensions also cannot be assumed significant, as the distribution of penetrations for these experiments is skewed towards the medio-posterior aspect of the HVc complex. Thus, the relationship between the locations of song specific units and the anteroposterior organization of HVc recently elucidated on the basis of anatomical techniques (Nottebohm et.al., 1982) cannot presently be assessed.

Figure 18. Distribution of song specific and tone burst units within the HVc complex. *A*, A crestyl violet stained, 30 μm thick section showing the HVc complex. HVc is the magnocellular elliptically shaped nucleus in the center. Filled arrows delimit the antero-posterior extent of HVc and point to the "shelf", a thin parvocellular strip contiguous with the ventral border of HVc. A relatively cell-free layer interposes between the dorsal surface of HVc and the lateral ventricle. The open arrow points to an electrolytic lesion of the buzz-buzz unit 600301 of Figure 15. a - anterior, APH - Area Parahippocampus, d - dorsal, HVc - Hyperstriatum Ventrale, pars caudale; NC - Neostriatum Caudale; V - Lateral Ventricle. Calibration bar = 100um. *B*, Ellipse onto which lesion data have been mapped (see text). Orientation as in *A*. Circles are sites of song specific units, stars are sites of tone burst responding units. No lesion site was dorsal to the lateral ventricle. Arrow indicates location on map of lesion shown in *A*.


A
APH
V
HVc
d
a
NC
B

The presumed source of auditory input to HVc, the "shelf" area, exists in canaries as a parvocellular strip contiguous with the ventral border of HVc (Kelly and Nottebohm, 1979). A similar morphology occurs in white crowns (personal observations, Figure 15*A*). Song specific units concentrate in the layer between HVc and the ventricle, and/or the dorsal perimeter of HVc, and not in the shelf area. This dorsal area has not been implicated as a recipient zone for field L (auditory telencephalon) projections. Thus, it is reasonable to propose that the input to song specific units may come via the shelf area, and therefore that circuits local to the HVc complex are responsible for the integrative properties of song specific units.

Other response categories. The majority of single units isolated in HVc showed responses similar to those reported in an earlier intracellular study. These units either did not respond whatsoever to any stimulus tested, or exhibited sufficiently weak onset, sustained, offset, or onset/offset responses to tone and noise bursts, that many repetitions were required to clearly demonstrate the response (Katz and Gurney, 1981). In the present study a number of units responded with short latency and brisk onset responses to wide-band noise bursts, or to tone bursts of a limited frequency range. The best frequencies of these units varied from 1.1 kHz to 9.5 kHz. The distribution of latencies to response for some of these units is shown in Figure 19, and a particularly clear example of the response is shown in Figure 20*A*. No latencies shorter than 18 milliseconds were encountered. The shortest latencies observed in field L were 8 milliseconds (Leppelsack, 1974), and in field GA of the frontal telencephalon a recent study demonstrated units with as short a latency as 14.5 milliseconds (Kirsch et.al., 1980). These data suggest the existence of multiple parallel auditory pathways in the avian telencephalon. Some units responded with brisk sustained activity throughout a noise or tone burst (e.g. Fig. 20*B*). A few units exhibited rather complex responses to tone bursts. For example, the unit of Figure

Figure 19. Latency of HVc units. The minimal time to response for short latency HVc units (n=51). Stimulus was optimal one, either wide-band noise burst or tone burst, at 65 dB SPL. Shortest latency observed was 18 msec.

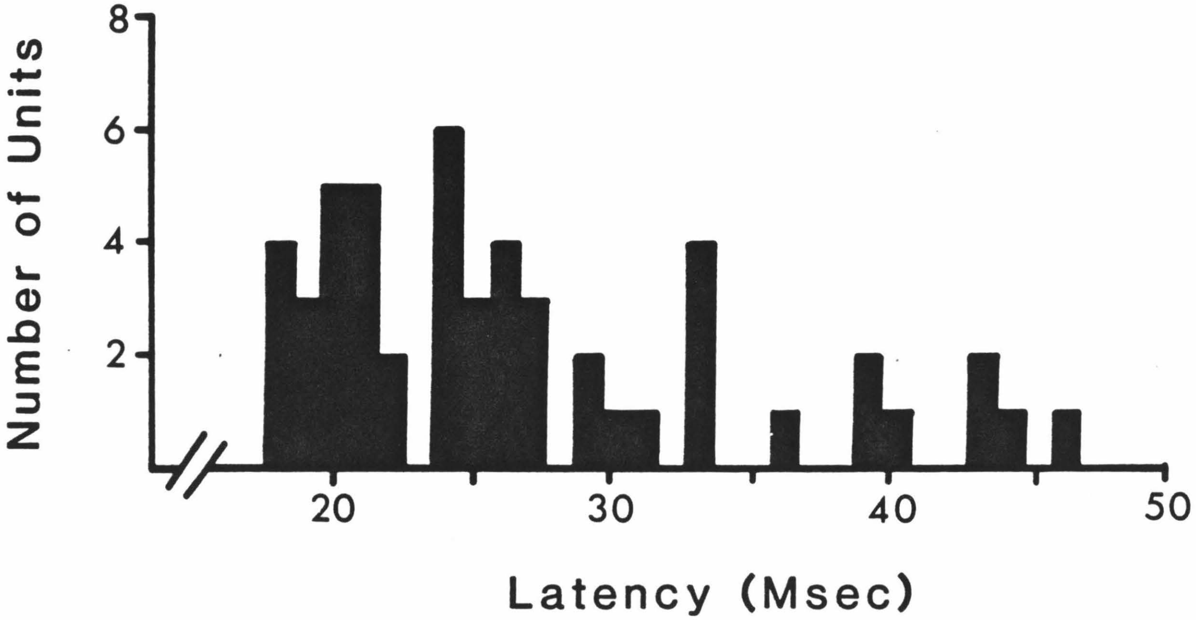
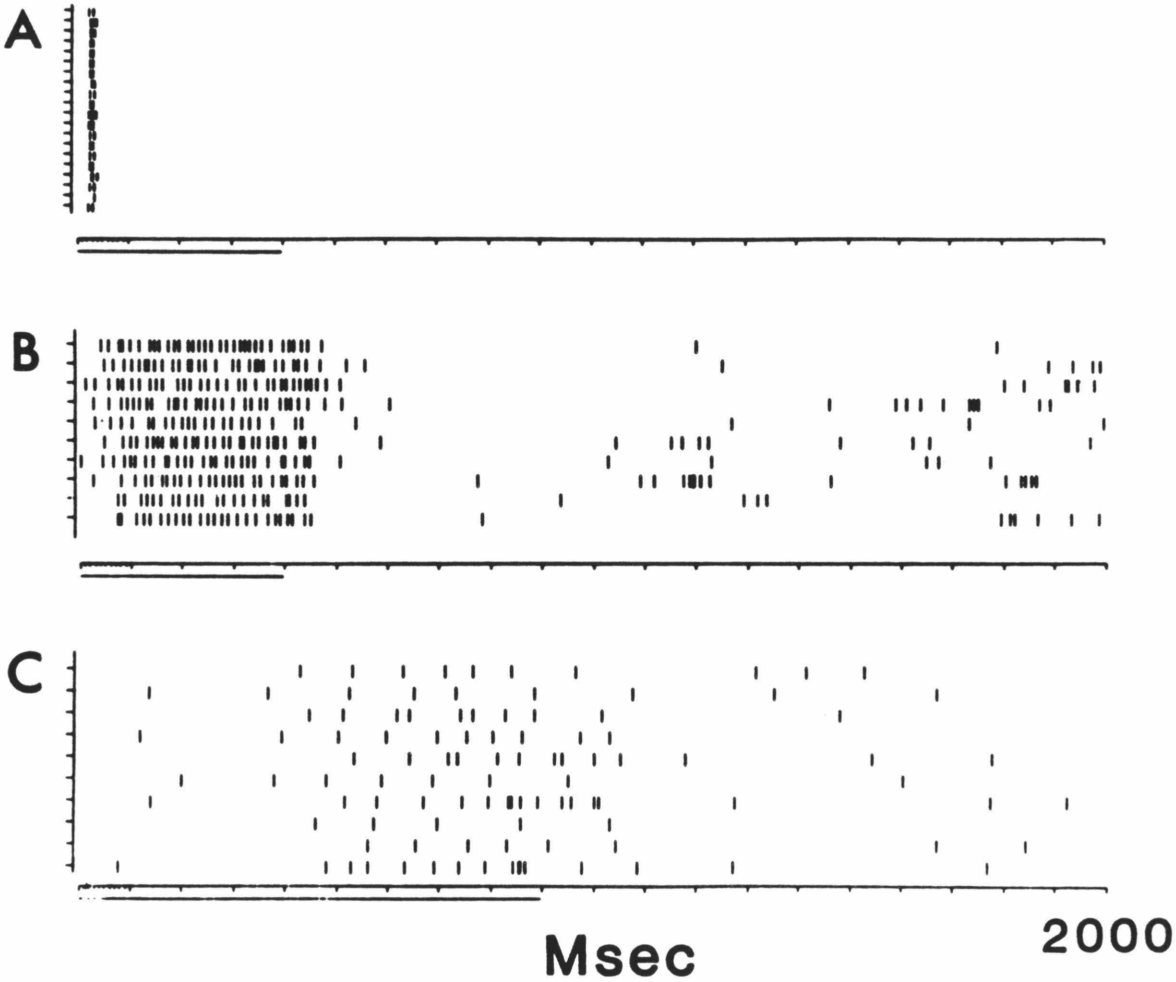

Number of Units
8
6
4
2
20
30
40
50
Latency (Msec)

Figure 20. Tone burst units in HVc. *A*, Particularly good example of short latency onset responding unit. Stimulus is wide-band noise, 65 dB SPL. Unit also responded well to tone bursts of 7 to 9 kHz. *B*, Particularly good example of HVc unit sustained response. Stimulus as in *A*. *C*, Responds to 65 dB, 900 msec tone bursts of 3.8 kHz, the center frequency of the whistle in the bird's own song.


A
B
C
Msec
2000

20*C* responded to tone bursts only of long duration, greater than 500 milliseconds. The response latency was similar at threshold (30 dB SPL) and at 35 dB above threshold. The unit was tuned rather sharply to 3.8 kHz, the center frequency of the 700 millisecond duration whistle in the bird's own song. On the basis of these response characteristics, this unit responded vigorously to the bird's own song, and responded poorly or not at all to ten other white crown songs, half of which were of the same dialect. The significance of the aforementioned response properties, seemingly intermediate in complexity between simple tone burst responding units and song specific units, remains unclear.

DISCUSSION

Criteria for specificity. A number of studies employing complex natural sounds as stimuli have relied on a variety of criteria in examining whether unit responses exhibited "specific", "selective", "sensitive", "feature detector", or "focal" properties. The potential significance of such a discovery is in its relationship to the presumed existence of functional sensory hierarchies that analyze behaviorally relevant stimuli, and the assumption that activation of elements at the top of such a hierarchy is an integral part of the process of perception for that specific stimulus (see Introduction). Thus, the response properties of a single unit must meet strict criteria for functional integration and stimulus selectivity if it is to be considered stimulus specific. In those systems such as echolocation in bats, where natural vocalizations are sufficiently simple that the attendant response requirements for acoustic specificity are clear, convincing stimulus specific complex units have been reported (O'Neill and Suga, 1982). In other systems, the results to date have not been entirely satisfactory.

In the current study, the following criteria were used as the basis for deciding that a unit's response was song specific.

1. Simple stimuli such as single tone bursts, and narrow-band and wide-band noise bursts, did not elicit responses that could be used to explain the responses to complex stimuli. This included inhibitory as well as excitatory effects.

2. The unit only responded to a subset within a repertoire of stimuli chosen on the basis of observations of naturally occuring behaviors. In every case, this subset of songs included the bird's own song. It was observed that this limited set of stimuli shared parameters that were missing in the ineffective stimuli of the repertoire.

3. Artificial stimuli synthesized on the basis of these shared parameters elicited the response. The systematic modification of the parameters of the artificial stimuli delineated the response properties and verified the acoustic parameters responsible for the efficacy of the various natural stimuli.

The first criterion attempts to distinguish between response properties that reflect peripheral frequency analysis and those that are dependent on central integrative processes. Even simple peripheral response properties can generate seemingly complex responses to complex stimuli. With the addition of inhibitory side-bands, complex responses of primary-like units can easily be misinterpreted as reflecting integrative mechanisms (e.g. see discussion of forwards/backwards paradigm). Furthermore, perhaps the only units with responses to complex sounds that can be fully predicted on the basis of simple tone burst tuning are those with primary-like response properties. Two tone excitatory/inhibitory interactions, and excitation to narrow-band or broad-band noise, are relatively simple mechanisms that nevertheless can generate significant response complexity not predictable simply on the basis of single tone burst tuning. Thus, in the absence of compelling observations to the contrary, it is essential to carefully search with tone bursts, including two tone stimulation, as well as with narrow-band and wide-band noise bursts, for simpler explanations for seemingly complex responses.

The second criterion is commonly accepted and applied by most researchers. Knowledge of the behavioral context of various vocalizations affords powerful insights into the potential significance of a unit's stimulus selectivity. For example, in this study the abundance of passerine neighbor-stranger discrimination studies (Brooks and Falls, 1975; Emlen, 1972; Verner and Milligan, 1971; Baker et.al.1981), and the well defined existence, if not role, of song dialects, has focused attention on the intra-dialect selectivity exhibited by some units. However, since the observed selectivity of a neuron reflects in part a repertoire biased by the experimenter, this criterion applied in isolation results in an inadequate description of neuronal response specificity.

Few studies have employed realistic but artificial sounds that mimic complex natural stimuli (criterion 3). This may, in part, reflect a technical limitation; nevertheless, the consequences are serious. For instance, it has been reported that in the starling field L units exhibited different levels of selectivity within a large repertoire of 80 naturally occuring call and song elements (Leppelsack and Vogt, 1976). Some units responded to many vocalizations, one extreme example responded to only one sound out of the entire repertoire and did not respond whatsoever to simple stimuli – tone and wide-band noise bursts. A very large repertoire of sounds with overlapping acoustic parameters limits the possibility that any observed response selectivity is capriciously generated by relatively simple parameters. Nevertheless, the use of synthetic yet effective sounds, modified parameter by parameter, could establish this conclusion convincingly by delineating precisely the parameters underlying the response selectivity.

In the present study, the aforementioned response property criteria were met by song specific units. When less rigorous requirements are applied, results can be ambiguous. For example, Scheich et.al. (1979) described 58 units that selected within a group and between groups of species-specific bird calls. Although approximately half of those units responded to nine or more out of 16 guinea fowl Iambus calls, few exhibited any response to the other guinea fowl calls comprising the repertoire. Forty-eight of these units responded to tone bursts, a high proportion responding to 1 kHz and to 1.8 to 1.9 kHz, which are prominent bands in the Iambi spectra. Response versus tone burst frequency histograms typically contained one or two major peaks, several or many minor peaks, and intervening inhibitory regions. Thus, a simple relationship between the spectral energy distribution of Iambi and the neuronal frequency tuning may acount for many of the observed responses. Furthermore, it was concluded that for many units the observed selectivity was based on spectral sensitivity to "frequency complexes" that define "focal properties" of Iambi. Whether this sensitivity was actually based on the entirety of the spectra, or on simple combinations of spectral components, cannot be ascertained because two tone stimulation studies were not undertaken.

The great majority of units in the HVc complex responded to complex stimuli such as song in a manner predictable on the basis of the excitatory and inhibitory responses to tones and narrow-band and wide-band noise bursts. Similar results have been reported in most studies of telencephalic and midbrain auditory nuclei in birds and mammals (Suga, 1965; Whitfield and Evans, 1965; Abeles and Goldstein, 1972; Watanabe and Katsuki, 1974; Leppelsack and Vogt, 1976; Scheich et.al., 1979; Bonke et.al., 1979; Symmes et.al., 1980; Langner et.al., 1981). In contrast to the conclusions of these authors stand the conclusions of Newman and Wollberg (1973), who investigated the functional properties of the auditory cortex of the squirrel monkey. They reported that most of the units responded to tone bursts and 89% responded to over half of 12 different vocalizations that constituted the chosen repertoire, They concluded that neither the tone burst tuning nor the parameters in the various call groups could explain the majority of unit responses. One explanation for this discrepancy is that the functional properties of the squirrel monkey auditory cortex are indeed different from what has been delineated for all other animals tested to date. However, the lack of artificial and modifiable versions of each call, the incomplete testing with tone and narrow-band and wide-band noise bursts of various amplitudes, as well as the lack of two tone stimulation studies, severely limits the possible interpretation of those results.

For many multiunit and single unit recordings, and for most song specific units, forward but not reversed song was an effective stimulus. The limitations in the interpretation of this manipulation must be stressed. Such discriminations can potentially be accounted for by the almost ubiquitous occurrence in sensory systems of neurons with antagonistic center-surround receptive fields. A complex stimulus moving in time from an excitatory center into an inhibitory surround of a unit could elicit an entirely different response if reversed (Suga, 1965). Furthermore, for multiunit recordings (McCasland and Konishi, 1981), it is difficult to distinguish between increased activation of the same population of units and a change in the composition of the population of units responding to different stimuli. Nevertheless, these and other data (McCasland and Konishi, 1981) suggest that in the HVc complex, the forward/reversed song paradigm is suggestive, but not demonstrative, of complex stimulus processing.

Possible physiological mechanisms of the responses. The present data afford some insight into the neural circuit generating the selectivity of song specific units. This analysis concentrates on the simplest of these units, the whistle-whistle units. It is assumed that all song specific whistle-whistle units derive their response properties from the same underlying neural circuit. The observations pertaining to the responses of song specific whistle-whistle units, that is, units that respond to a tone burst of frequency F2 if and only if it is preceded by another tone burst of frequency F1, can be summarized as follows.

1. Maximal response is elicited for frequencies F1 and F2. The consecutive tone tuning for these frequencies is reasonably sharp and peaked, typically with inhibitory side bands.

2. F2 can equal F1.

3. Simultaneous two tone stimulation inhibits the response to consecutive tones, even if the additional tone burst is of frequency F1 or F2.

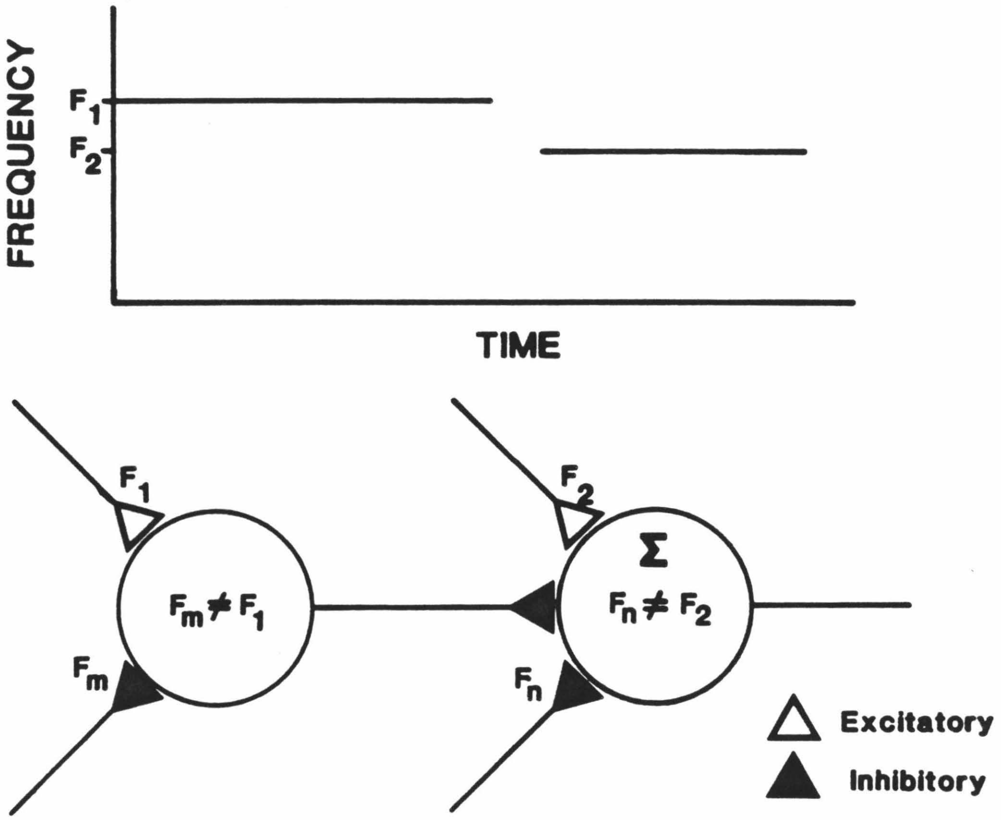
4. The timing information, the length of the tone bursts or the interval duration, can vary over hundreds of milliseconds with only slight variation in response strength.

5. The first tone often is inhibitory, and elicits weak excitation at the offset.

Numerous two or multi cell models can account for the responses of song specific whistle-whistle units. Single cell models fail either in that they cannot account for the timing, or if constrained to incorporate the timing requirements, do not have the necessary frequency requirements. A simple circuit that meets all the criteria is the two cell model of Figure 18, the second cell representing a song specific whistle-whistle unit. The first cell is tuned to F1, and is inhibited by tone bursts of frequencies other than F1. The output of the first cell inhibits the second cell for the duration of the first tone. At the offset of the first tone, a rebound from release from inhibition in the second cell produces a perithreshold depolarization. The second tone elicits a subthreshold depolarization in the second cell. The summed currents elicit a suprathreshold response.

All of the response properties of song specific whistle-whistle units can be accounted for by the proposed model. The leading consecutive tone tuning reflects the frequency tuning of the first cell in the model, and the following consecutive tone tuning reflects the tuning of the second cell. Even when F1=F2, a

Figure 21. Simple neural circuit to account for the response properties of song specific whistle-whistle units. The stimulus model is represented by upper graph. Lower schematic is of the excitatory and inhibitory input to two cells (circles). The second cell responds if rebound excitation from release from tonic inhibition by first cell is synchronized with activation of excitatory input to the second cell (see text).


FREQUENCY
F_1
F_2
TIME
F_1
$F_m \neq F_1$
F_m
F_2
Σ
$F_n \neq F_2$
F_n
△ Excitatory
▲ Inhibitory

single tone burst elicits only weak excitation at the onset (see Figure 4) because excitation at the second cell is rapidly followed by inhibition from the first cell. When F1≠F2, two tone stimulation, even if of frequencies F1 and F2, elicits inhibition in either the first cell or the second cell, and thus abolishes the response. The first tone in isolation activates an inhibitory input to the second cell. Perhaps the weak offset excitation to initial phrases or tone bursts seen in several recordings reflects the second cell's rebound from inhibition at the offset of the sound. If each cell in the circuit has elaborate frequency requirements, the more complicated responses seen in such units as the whistle-buzz and buzz-buzz units may possibly be modeled in the same general way.

Clear and convincing examples of temporal facilitation have already been reported at the level of single units – delay sensitive neurons of the bat auditory cortex (O'Neill and Suga, 1979,1982; Sullivan, 1982). Those units also respond strongly to specific temporal combinations of sound elements that in isolation elicit weak or no response. Other aspects of their response, however, such as sharp temporal tuning, lack of two tone inhibition, and response to high repetition rates, are different from song specific units, potentially reflecting the differing behavioral requirements of bats and songbirds. It will be interesting to see if the current model has general applicability.

The proposed model has several other implications. Intracellular recordings from the second cell proposed by the model would establish whether the functional integration seen in these recordings resides at the level of the HVc complex, or indeed reflects complex processing of sounds that occurs in field L (Leppelsack and Vogt, 1976; Scheich et.al., 1979; Langner et.al., 1981). The model makes specific predictions about the synaptic events expected in an intracellular recording of the second cell -- inhibition to frequencies other than F2, e.g. F1; excitation to F2; and excitation at the offset of the first tone. Finally, if these cells are indeed modified by the song learning and memory experience, then song learning should involve not only changes at synapses (Hebb, 1949), but also the establishment of multineuronal circuit properties.

Behavioral significance of the responses. It is tempting to conclude that the song specific responses reported here reflect some aspect of song learning. In the current study, however, a significant sampling bias was introduced by searching with one particular song, the bird's own song. In the extreme case, it is possible that the HVc complex contains a galaxy of highly complex and specific unit responses, and that the search paradigm merely selected for those few units with parameter requirements fortuitously satisfied by a specific song. A number of observations argue against this interpretation, and indicate that at least song specific HVc auditory responses are modified during song ontogeny.

As auditory feedback is known to be necessary for normal motor (song) development (Konishi, 1965), it is significant that song specific auditory units are found in an area with motor activity specifically related to song production (Nottebohm, 1976; McCasland and Konishi, 1981). In this study, song specific units were recorded from white-crowned sparrows of the following groups: wild caught Gambelii and Nuttalli, laboratory-reared birds that successfully copied computer tutor songs, and birds that sang isolate-like songs. The degree of specificity exhibited by song specific units would require a large parameter space to be embodied at the single unit level by the HVc complex of each bird if the system were innately specified. The two response classes accounting for the majority of the recordings in the HVc complex responded weakly to most sounds or did not respond at all. Song specific units, however, were typically inhibited by stimuli other than song and thus represented a minority population. Furthermore, some laboratory reared birds were successfully tutored with songs containing parameters never found in the wild (e.g. reversed trill - see Konishi, 1978). Isolate-like songs also did not exhibit wild-type morphologies. Nevertheless, song specific units in these birds reflected the abnormal songs (Figures 4, 10, and 11). When multiunit recordings responded vigorously to song, the bird's own song and allied songs always were the effective stimuli. Distantly related songs were never as effective. Finally, though many units were tested with reversed song, a complex stimulus, only a few units responded preferentially to reversed as opposed to forward song. If the response properties of complex units in HVc were specified independently of song, one would predict reversed song to be as effective a stimulus as forward song.

These observations cannot be construed as definitive proof, but are highly suggestive of a role for song ontogeny in modifying the responses of units of the HVc complex. If this contention is accepted, it is interesting that for those birds that failed to copy their tutor song(s) and sang either variations of the tutor song(s) or isolate-like songs, no units specific for tutor song were found. For these experiments tutor song(s) were also used as a search stimulus, although the relative number of presentations of tutor song(s) versus the bird's own song was not controlled. If this initial observation is further validated, it suggests experiments that could delineate the ontogenetic processes responsible for the song of an isolate bird, a behavioral phenomena for which no current explanation exists. Either the failure to accept a tutor model during the critical period produces isolate song specific units in HVc that then direct the song production process, or song specific units undergo a round of plasticity at or after the time of song crystallization. That is, song specific units may be the neural substrate of the template for song learning, or they may be involved in some other aspect of song production, such as the matching of auditory feedback to the template.

CONCLUDING REMARKS

The existence of neurons with highly selective response properties that clearly reflect a significant degree of central integration strongly supports suggestions (e.g. Barlow, 1972; Bullock, 1983) that at least some sensory pathways are organized in a clear functional hierarchy, with elements at the top of the hierarchy specialized for specific tasks. Consistent with the predictions of those authors, there is a relative paucity of song specific units, with perhaps only several hundred units in each hemisphere's HVc per each of four to five functional classes. Even the most complex responses observed in this study required appropriate parameters of only some parts of the song – two phrases, or several elements within a single phrase; and the songs of various birds that contained the appropriate phrase(s) elicited a response. No unit in the HVc complex responded at the end of the song, and only if all the song phrases were the appropriate ones. Thus, whichever behavioral task these units are involved in, the present data require the sequential and possibly parallel activation of different populations of neurons to encode the entire song. It is possible that the apex of sensory processing for birdsong at the level of single neurons would be only one stage of integration beyond that described in this study. There is no logic available to predict whether further levels of integration do exist, but the well characterized projections of HVc neurons suggests area X as a candidate site.

Many investigations of complex stimulus discrimination in sensory systems have been greatly aided by the insight provided from the neuroethological approach. In this study, the use of song as a search stimulus has uncovered responses that otherwise might have escaped notice. The significance of these responses has been assessed in relation to the natural history, critical period, crystallization and other behavioral observations pertaining to bird song. One of the serious limitations of "pure" neurophysiology is in the interpretation of results in terms of their function in the context of animal behavior. However, the suggestion that the song specific units described here reflect some aspect of the song learning experience can be directly tested by recording from young birds that have yet to sing but are past the critical period for acquisition of an appropriate model. It may be possible to directly assess the relationship between song specific units and the hypothesized neural template. The neural correlates of acoustic imprinting seen in birds may reside at the level of single cells. Perhaps some of the neural mechanisms underlying human vocal imprinting are similar.

Appendix

ANALOG SYNTHESIS OF ZERO-CROSSING ANALYZED BIRDSONGS

The song synthesis procedure described in Chapter 1 requires an analog device interfaced with the computer for the final song waveform synthesis. The control diagram of such a circuit, which converts voltage steps into piecewise linear frequency and amplitude modulations, is shown in Figure 22. The input to the circuit consists of two 1 kHz analog voltage step functions, and a 1 kHz TTL pulse. All three inputs are phase synchronized. The analog inputs represent frequency and amplitude, and the input voltage range of 0 to 5 volts controls an amplitude range of over 80 dB, and frequencies up to approximately 16 kHz.

Two distinct sections process the frequency and amplitude voltage functions independently. The input amplitude step function is exponentiated, which allows the limited dynamic range of the computer's digital to analog converter (in the current implementation 12 bits or $1{:}2^{12}$) to control the full dynamic range of the analog hardware while maintaining equivelent resolution (quantitization error) over the entire range. Each step function is converted into a piecewise linear ramp function by integrating between the past and current values of the steps. The voltage ramps are then converted into current ramps. The frequency current ramp serves as the input to a current controlled oscillator (CCO). The sinusoidal output waveform of the CCO may then be attenuated and summed with other frequency channels (not shown - three other channels in the current implementation) before serving as the input to an operational transconductance amplifier (OTA). The amplitude current ramp controls the OTA, and thus amplitude modulates the input signal. The output of the OTA is converted into a voltage and buffered.

The circuit diagram of the song synthesizer is shown in Figure 23. Exponentiation of the frequency input is achieved by signal conditioning the input (voltage offset and sign reversal) prior to input to a logarithmic element setup in antilogorithmic configuration. For both frequency and amplitude inputs, voltage steps functions are converted to voltage ramp functions as follows. At each timing pulse, the output of an integrator is sampled and held by momentarily closing an FET switch and charging a capacitor. (With the low output impedance of the integrator and high input impedance of the sample and hold op amp, the charging time and droop is negligible). This stored voltage, the final voltage for the preceeding interval, is differenced with the current input voltage. This difference voltage is the voltage gain or loss the ramp must achieve over a one millisecond interval. The difference voltage is applied to the input of the integrator and initiates a linear voltage ramp. Because the time constant of the integrator is adjusted to one millisecond, the integrator acheives the new input value at the end of the time interval. The voltage output of the integrator is converted into a current by a low input current offset op amp with a transistor in feedback. Use of a β optimized transistor ensures linearity for the conversion. For the frequency signal, the current ramp serves as the input to a current controlled oscillator. It was discovered essential to use an NPO (negative, positive, zero) capacitor for the timing capacitor to minimize temperature drift. After user-selected attenuation and summation with three parallel frequency channels, the composite signal is connected to the input of the first of two ganged OTAs. The amplitude signal serves as the control current for the OTAs. It was observed that the OTAs were nonlinear at low control currents. To curcumvent this problem, two OTAs were placed in series, thus achieving the desired dynamic range while keeping the controlling currents in the linear region. Furthermore, to achieve maximum pinch-off, the OTAs required lower supply voltages than the input devices to the OTAs. Thus, two analog power supply levels are provided. The output of the final OTA is buffered, converted into a voltage, and is the output of the device.

Figure 22. Control flow-chart of the song synthesizer. Inputs: VCA - voltage controlled amplifier; VCO - voltage controlled oscillator; TTL (digital). Other symbols: e^x - exponentiation; V/I - voltage to current conversion; I/V - current to voltage conversion; CCO - current controlled oscillator; OTA - operational transconductance amplifier. For the definite integral, $s(t)$ is a voltage step function, and τ represents the (fixed) time interval for the step function.

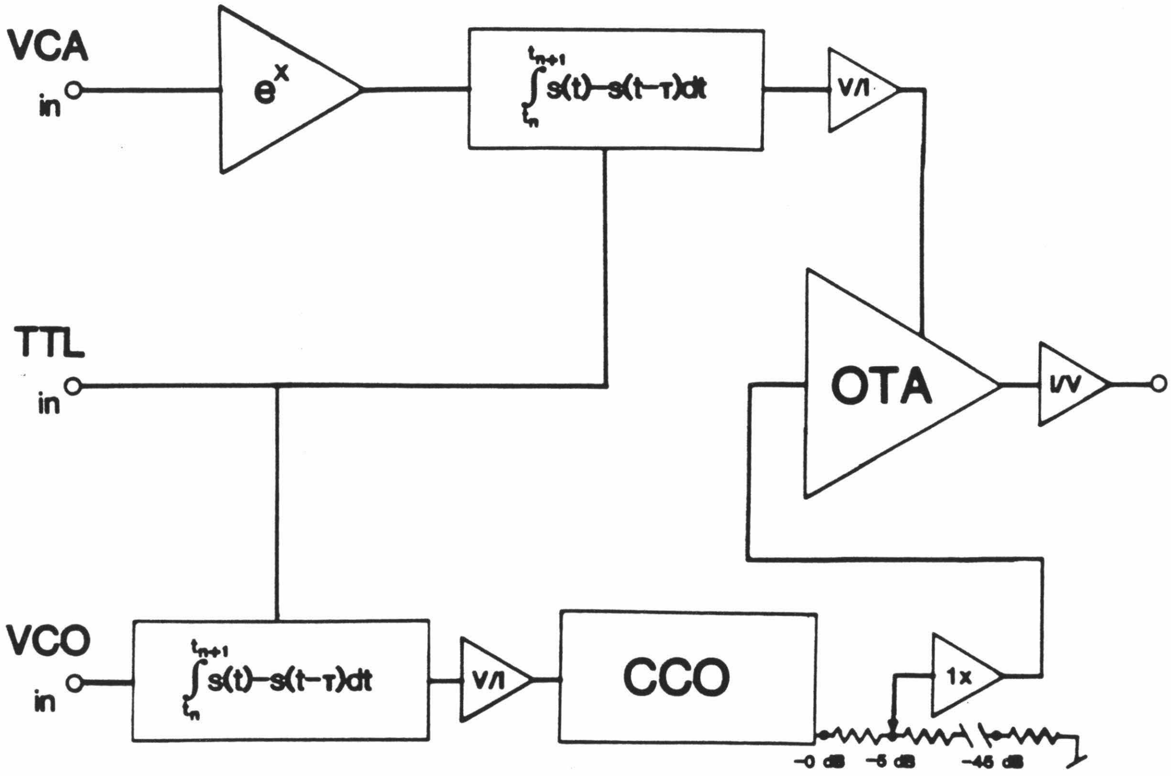
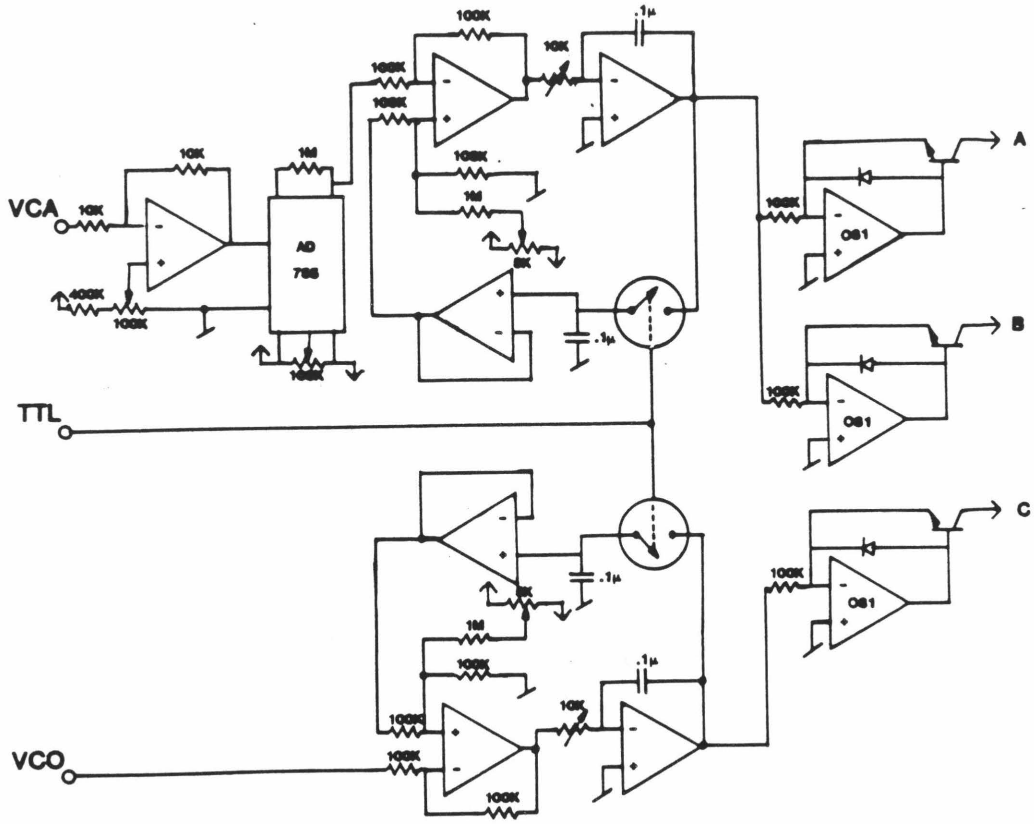
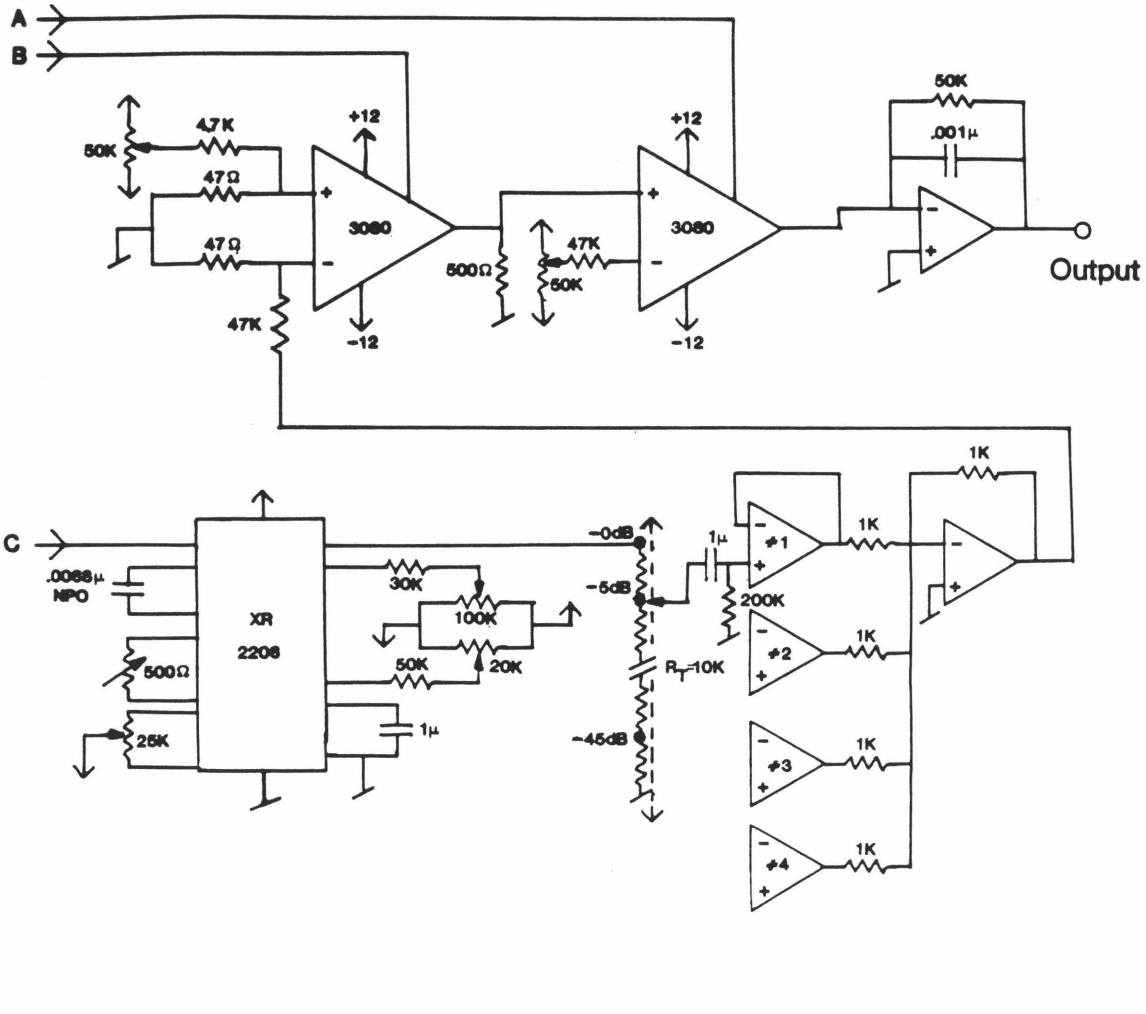

VCA
in
e^x
$\int_{t_n}^{t_{n+1}} s(t)-s(t-\tau)dt$
V/I
TTL
in
OTA
I/V
VCO
in
$\int_{t_n}^{t_{n+1}} s(t)-s(t-\tau)dt$
V/I
CCO
1x
−0 dB
−5 dB
−45 dB

Figure 23. Circuit diagram of song synthesizer. AD 755 is an Analog Devices logarithmic element. The analog switches are RCA CD4066 FET switches. TI 081 is a Texas Instruments low input offset current FET op amp. The transistors are Texas Instruments TIS 97 β optimized. CA 3080 is a RCA operational transconductance amplifier. XR 2206 is an EXAR voltage controlled oscillator. All op amps are TI 071 or TI 072. Diodes are 1N914. The 0.1 μf capacitors on the output of the FET switches should be low leakage capacitors, and low temperature drift coefficient capacitors such as NPO types should be used for the oscillator timing capacitors. All other components are non-critical. The power supply is ±12V for the CA 3080, and is ±15V for all other active devices.



VCA
10K
10K
400K
100K
AD
755
1M
100K
100K
100K
100K
10K
.1µ
100K
1M
5K
.1µ
TTL
A
B
C
100K
061
100K
061
100K
061
5K
.1µ
1M
100K
100K
100K
10K
.1µ
VCO
100K



A
B
50K
4.7K
47Ω
47Ω
+12
3080
-12
47K
500Ω
47K
50K
+12
3080
-12
50K
.001µ
Output
1K
C
.0068µ
NPO
500Ω
25K
XR
2206
30K
100K
50K
20K
1µ
-0dB
-5dB
R_T=10K
-45dB
1µ
200K
÷1
1K
÷2
1K
÷3
1K
÷4
1K

REFERENCES

Abeles, M. and M. H. Goldstein (1972) Responses of single units in the primary auditory cortex of the cat to tones and tone pairs. Brain Res. *42*: 337-352.

Allman, J. (1977) Evolution of the visual system in the early primates. Prog. Psychobiol. Physiol. Psychol. *7*:1-53.

Arbib, M.A. (1981) Perceptual structures and distributed motor control. In *Handbook of Physiology, The Nervous System, Vol. 2*, V. B. Brooks, J. M. Brookhart, and V. B. Mountcastle, eds., pp. 1449-1480, Williams and Wilkins, Baltimore.

Baker, M. C., D. B. Thompson, and G. L. Sherman (1981) Neighbor/stranger song discrimination in white-crowned sparrows. Condor *83*: 265-267.

Barlow, H. B. (1972) Single units and sensation: a neuron doctrine for perceptual psychology? Perception *1*: 371-394.

Bentley, D., and M. Konishi (1978) Neural control of behavior. Ann. Rev. Neurosci. *1*:35-59.

Bonke, D., H. Scheich, and G. Langner (1979) Responsiveness of units in the auditory neostriatum of the guinea fowl (*Numida meleagris*) to species-specific calls and synthetic stimuli. I. Tonotopy and functional zones of field L. J. Comp. Physiol. *132*: 243-255.

Brooks, R. J. and J. B. Falls (1975) Individual recognition by song in white-throated sparrows. III. Song features used in individual recognition. Can. J. Zool. *53*: 1749-1761.

Bruns, V. (1976) Peripheral auditory tuning for fine frequency analysis by the CF-FM bat Rhinolophus ferrumequinum. J. Comp. Physiol. *106*:77-86.

Bullock, T. H. (1961) The problem of Recognition in an analyzer made of neurons. In *Sensory Communication*, W. Rosenblith, ed., pp. 717-724, MIT Press, Cambridge.

Bullock, T. H. (1977) Ed., *Dahlen Konferenzenan on Recognition of Complex Acoustic Signals*. Abakon Verlagsgesellschaft, Berlin.

Bullock, T. H. (1983a) Perspectives on the neuroethology of sensory convergences. In *Multimodal Convergences in Sensory Systems*, E. Horn, ed. Academy of Science and Literature of Mainz. Mainz, Germany. in press.

Bullock, T. H. (1983b) Where do we go from here: need for cross-taxa comparison. In *NRP Research Bulletin*, NRP Press, Brookline, Mass., in press.

Chomsky, N. (1959) Review of *Verbal Behavior*, by B. F. Skinner. Language *35*:26-58.

Dawkins, R. (1976) Hierarchical organization: a candidate principle for ethology. In *Growing Points in Ethology*, P. P. G. Bateson and R. A. Hinde, eds., 1-54, Cambridge University Press, Cambridge.

Dickinson, R. (1980) *Animal Learning Theory*. Oxford U. Press, Oxford.

Dooling, R. J., C. Clark, R. Miller, and T. Bunnell (1982) Program package for the analysis and synthesis of animal vocalizations. Behav. Res. Meth. Instr. *14(5)*: 487.

Dorrscheidt, G. J. (1978) A reversible method for sound analysis adapted to bioacoustical experiments. Int. J. Bio-Med. Comp. *9*: 127-145.

Emlen, S. T. (1972) An experimental analysis of the parameters of bird song eliciting species recognition. Behavior *41*: 130-171.

Ewert, J. P., R. R. Capranica, and D. J. Ingle, eds. (1983) *Advances in Vertebrate Neuroethology*, Plenum Press, N.Y.

Frishkopf, L. S., and M. H. Goldstein Jr. (1963) Responses to acoustic stimuli from single units in the eighth nerve of the bullfrog. J. Acoust. Soc. Amer. *35*:1219-1228.

Green, S. (1975) Variation of vocal pattern with social situation in the Japanese monkey (Macaca fuscata): A field study. In *Primate Behavior* L. A. Rosenblum, ed. *4*:2-103, Academic Press, N.Y.

Greenewalt, C. H. (1968) *Bird song: acoustics and physiology.* Smithsonian Institute Press, Washington,D.C.

Gross, C. G., C. E. Rocha-Miranda, and D. B. Bender (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J. Neurophysiol. *35*: 96-111.

Hebb, D. O. (1949) *The organization of behavior, a neuropsychological theory*, Wiley, New York.

Hopfield, J. J. (1981) Human memory, error correction codes and spin glasses. CIT internal circular.

Hoyle, G. (1976) Approaches to understanding the neurophysiological bases of behavior. In *Simpler Networks and Behavior*, J. C. Frentress, ed., pp. 21-38, Sinauer, Sundeland, Mass.

Hubel, D. H. (1981) Nobel Lecture. In Nature (1982) *299*: 515-524.

Hubel, D. H., and T. N. Wiesel (1962) Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. J. Physiol. *160*: 106-154.

Hubel, D. H., and T. N. Wiesel (1974) Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. J. Comp. Neurol. *158*:295-305.

Julez, B. (1981) Textons, the elements of texture perception, and their interactions. Nature *290*:91-97.

Katz, L. C., and M. E. Gurney (1981) Auditory responses in the zebra finch's motor system for song. Brain Res. *211*: 192-197.

Kelly, D. B. and F. E. Nottebohm (1979) Projections of a telencephalic auditory nucleus - field L - in the canary. J. Comp. Neurol. *183*: 455-470.

Kirsch, M., R. B. Coles, and H.-J. Leppelsack (1980) Unit recordings from a new auditory area in the frontal neostriatum of the awake starling (*Sturnus vulgaris*). Exp. Brain Res. *38*: 375-380.

Knudsen, E. I., and M. Konishi (1978) A neural map of auditory space in the owl. Science *200*: 795-797.

Konishi, M. (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. Z. Teirpsychol. *22*: 770-783.

Konishi, M. (1971) Ethology and neurobiology. Amer. Sci. *59*:56-63.

Konishi, M. (1973) How the owl tracks its prey. Amer. Sci. *61*:414-424.

Konishi, M. (1978) Auditory environment and vocal development in birds. In *Perception and Experience*, R. D. Walk and H. L. Pick, Jr., eds., pp. 105-118, Plenum Publishing Corp., N. Y.

Konish, M., and E. I. Knudsen (1982) A theory of neural auditory space. In *Cortical Sensory Organization, Vol. 3, Multiple Auditory Areas*, C. N. Woolsey, ed., pp. 219-229, Humana Press, Clifton, N.J.

Kroodsma, D. E. (1977) Re-evaluation of song development in the song sparrow. Animal Beh. *25*: 390-399.

Kupfermann, I., and K. R. Weiss (1978) The command neuron concept. Behav. Brain Sci. *1*:3-39.

Langner, G., and H. Scheich (1978) Active phase coupling in electric fish: behavioral control with microsecond precision. J. Comp. Physiol. *128*:235-240.

Langner, G., D. Bonke, and H. Scheich (1981) Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. Exp. Brain Res. *43*: 11-24

Lee, Y. W. and M. Schetzen (1965) Measurement of the Wiener kernels of a non-linear system by cross-correlation. Int. J. Control *2*:237-254.

Lehrman, D. S. (1953) A critique of Konrad Lorenz' theory of instinctive behavior. Quart. Rev. Biol. *28*:337-363.

Leppelsack, H.-J. (1974) Funktionelle eigenshaften der horbahnim Feld L des neostriatum caudale des staren (*Sturnus vulgaris* L.,Aves). J. Comp. Physiol. *88*: 271-320.

Leppelsack, H.-J. and M. Vogt (1976) Responses of auditory neurons in the forebrain of a songbird during stimulation with species-specific sounds. J. Comp. Physiol. *107*: 263-274.

Leppelsack, H.-J. (1978) Unit responses to species-specific sounds in the auditory forebrain center of birds. Fed. Proc. *37*: 2236-2241.

Leppelsack, H.-J. (1981) Einfluβ von Gesangsleren auf das Autwortverhalten auditorischer Vorderhirnneuronen eines Singvogels. *Habilitationsschrift* Ruhr-Universitat, Bochum.

Licklider, J. C. R. and G. A. Miller (1951) The perception of speech. In: *Handbook of Experimental Psychology*, S. S. Stevens, ed., pp. 1040-1074, Wiley, N.Y.

Marler, P. (1969) Tonal quality of bird sounds. In: *Bird Vocalizations*, R. A. Hinde, ed., pp. 5-18, Cambridge University Press, Cambridge.

Marler, P. (1970) A comparative approach to vocal learning: song development in white-crowned sparrows. J. Comp. Physiol. Psychol. *71(2), Part 2*: 1-25.

Marler, P. and M. Tamura (1962) Song dialects in three populations of white-crowned sparrows. Condor *64*: 368-377.

Marler, P. and P. Mundinger (1971) Vocal learning in birds. In: *Ontogeny of vocal behavior*, H. Moltz, ed., pp. 389-450, Academic Press, N.Y.

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

Marr, D. (1982) *Vision*. Freeman, San Francisco.

Maynard, D. M., and A. I. Selverston (1975) Organization of the stomatogastric ganglion of the spiny lobster IV. The pyloric system. J. Comp. Physiol. *100*: 161-182.

McCasland, J. and M. Konishi (1981) Interaction between auditory and motor activities in an avian song control nucleus. Proc. Natl. Acad. Sci. USA *78*: 7815-7819.

McCulloch, W. S. (1945) A heterarchy of values determined by the topography of nervous nets. Bull. Math. Biophys. *7*: 89-93.

McCulloch, W. S., and W. Pitts (1943) A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. *5*: 115-133.

Mesarovic, M. D. (1968) Systems theory and biology - view of a theoretician. In ***Systems Theory and Biology***, M. D. Mesarovic, ed., pp. 59-87 Springer-Verlag, N.Y.

Miles, F. A., and S. G. Lissgerger (1981) Plasticity in the vestibulo-ocular reflex: a new hypothesis. Ann. Rev. Neurosci. *4*:273-299.

Moiseff, A. and M. Konishi (1981) Neuronal and behavioral sensitivity to binaural time differences in the owl. J. Neurosci. *1*:40-48.

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

Morse, P. A. (1979) The infancy of infant speech perception: the first decade of research. Brain Behav. Evol. *16*:351-373.

Mountcastle, V. B. (1976) The world around us: neural command functions for selective attention. Neurosci. Res. Program Bull. *14*:1-47.

Mountcastle, V. B. (1978) Brain mechanisms for directed attention. J. R. Soc. Med. *71*:14-28.

Mulligan, J. A. (1966) Singing behavior and its development in the song sparrow, ***Melospiza melodia***. University of California Publications in Zoology *81*: 1-73.

Newman, J. D. and Z. Wollberg (1973) Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. Brain Res. *54*: 287-304.

Nottebohm, F. E., T. M. Stokes, and C. M. Leonard (1976) Central control of song in the canary, ***Serinus canarius***. J. Comp. Neurol. *165*: 457-486.

Nottebohm, F. E., D. B. Kelly, and J. A. Paton (1982) Connections of vocal control nuclei in the canary telencephalon. J. Comp. Neurol. *207*: 344-357.

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 and its representation in the auditory cortex of the mustached bat. J. Neurosci. *2*: 17-31.

Partridge, B. L., W. Heiligenberg, and J. Marsubara (1981) The neural basis of a sensory filter in the jamming avoidance response: no grandmother cells in sight. J. Comp. Physiol. *145*: 153-168.

Poggio, T. (1981) Wiener-like identification techniques. In *Theoretical Approaches in Neurobiology*, T. Poggio and W. E. Reichardt, ed. pp. 60-63, MIT Press, Cambridge.

Reichardt, W. E., and T. Poggio (1981) Visual control of flight in flies. In *Theoretical Approaches in Neurobiology*, T. Poggio and W. E. Reichardt, ed., pp. 135-150, MIT Press, Cambridge.

Richards, D. G. (1979) Recognition of neighbors by associative learning in rufous-sided towhees. Auk *96*: 688-693.

Roeder, K. P., and A. E. Treat (1961) The reception of bat cries by the tympanic organ of noctuid moths. In *Sensory Communication*, W. Rosenblith, ed., pp. 545-560, MIT Press, Cambridge.

Scheich, H., G. Langner, and R. Koch (1977) Coding of narrow-band and wide-band vocalizations in the auditory midbrain nucleus (MLD) of the guinea fowl (*Numida meleagris*). J. Comp. Physiol. *117*: 245-265

Scheich, H., G. Langner, and D. Bonke (1979) Responsiveness of units in the auditory neostriatum of the guinea fowl (*Numida meleagris*) to species-specific calls and synthetic stimuli. II. Discrimination of Iambus-like calls. J. Comp. Physiol. *132*: 257-276.

Selverston, A. I., and J. P. Miller (1982) Application of a cell inactivation technique to the study of a small neural network. TINS *5*: 120-123.

Simmons, J. A. (1979) Perception of echo phase information in bat sonar. Science *204*:1336-1338.

Skinner, B. F. (1938) *The Behavior of Organisms*. Appelton, N.Y.

Staddon, J. E. R., L. W. McGeorge, R. A. Bruce, and F. F. Klein (1978) A simple method for rapid analysis of animal sounds. Z. Tierpsychol. *48*: 306-330.

Stone, J., B. Dreher, and A. Leventhal (1979) Hierarchical and parallel mechanisms in the organization of visual cortex. Brain Res. Rev. *1*: 345-394.

Stone, J., and B. Dreher (1982) Parallel processing of information in the visual pathways. TINS *5*: 441-446.

Suga, N. (1965) Analysis of frequency-modulated sounds by auditory neurons of echo-locating bats. J. Physiol. *179*: 26-53.

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.

Sullivan, W. E. III (1982) Neural representation of target distance in auditory cortex of the echolocating bat *Myotis lucifugus*. J. Neurophysiol. *48*: 1011-1032.

Symmes, D., G. E. Alexander, and J. D. Newman (1980) Neural processing of vocalizations and artificial stimuli in the medial geniculate body of squirrel monkey. Hearing Res. *3*: 133-146.

Szentagothai, J., and M. A. Arbib (1975) *Conceptual models of neuronal organization*. MIT Press, Cambridge.

Thorpe, W. H. (1961) *Bird song: the biology of vocal communication and expression in birds*. Cambridge University Press. Cambridge, England.

Uttal, W. R. (1981) *A taxonomy of visual processes*. Lawrence Erlbaum, Hillsdale, N.J.

Van Essen, D. C., and J. H. R. Maunsell (1983) Hierarchical organization and functional streams in the visual cortex. TINS, in press.

Verner, J., and M. M. Milligan (1971) Inter-population song dialect discrimination in the white-crowned sparrow. Condor *73*: 208-213.

von Uexkull, J. (1909) *Umwelt und Innenleben der Tiere*. Berlin.

Watanabe, T. and Y. Katsuki (1974) Response patterns of single auditory neurons of the cat to species-specific vocalization. Jap. J. Physiol. *24*: 135-155.

Whitfield, I. C., and F. E. Evans (1965) Responses of auditory cortical neurons to stimuli of changing frequency. J. Neurophysiol. *28*: 655-672.

Wiener, N. (1958) *Nonlinear Problems in Random Theory*. John Wiley, N.Y.

Wiersma, C. A. G. (1952) Neurons of arthropods. Cold Spring Harbor Symp. Quant. Biol. *17*:155-163.

Wiersma, C. A. G., and K. Ikeda (1964) Interneurons commanding swimmeret movements in the crayfish Procambarus clarkii (Girard). Comp. Biochem. Physiol. *12*:509-525.

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

Worden, F. G., and R. Galambos, eds. (1972) *Auditory Processing of Biologically Significant Sounds*. MIT Press, Cambridge.

Zoloth, S. R., M. R. Petersen, S. Green, P. Marler, D. B. Moody, and W. Stebbens (1979) Species-specific perceptual processing of vocal sounds by monkeys. Science *204*:870-873.

Zoloth, S. R., R. J. Dooling, R. Miller, and S. Peters (1980) A minicomputer system-for the synthesis of animal vocalizations. Z. Tierpsychol. *54*: 151-162.