7

The Neuroethology of Vocal Communication: Perception and Cognition

Timothy Q. Gentner and Daniel Margoliash

1. Introduction

In its most common sense, acoustic communication occurs between animals, but in special cases it may occur within individual animals in the context of autocommunication. Communication can be described as an information exchange that alters the behavior of the communicating animals. Acoustic communication signals (typically vocalizations) are shaped by the physics of the sound-producing organs, the physical media they traverse, and the physics of the receptor organs (Bass and Clark, Chapter 2; Fitch and Hauser, Chapter 3; Ryan and Kime, Chapter 5). Vocal communication signals are also shaped by the perceptual mechanisms of the receiver, by the proximate behavioral states of the senders and receivers (Boughman and Moss, Chapter 4; Yamaguchi and Kelley, Chapter 6), and by the evolutionary history of the senders and receivers, most often in the context of sexual selection. The information in vocal signals is represented by non-random acoustic variation that may either form discrete categories or fall along graded continua. Likewise, vocal signals may be perceived as members of discrete categories or along graded continua. Whether graded or discrete, animals must account for the statistical variation in vocal signals as they are produced, transmitted, and perceived. Ultimately, this means that acoustic behaviors are constrained by the natural variation in communication signals.

The wide diversity of behavioral constraints makes exploration of the neural mechanisms of acoustic communication both exciting and daunting. On one hand, one of the grand challenges in neurobiology has been the application of mechanistic analysis to perceptual and cognitive components of brain function, and the neuroethology of acoustic communication provides a logical and rigorous approach to such problems. At the same time, however, proper treatment of the mechanisms of acoustic communication requires a neural analysis that is sensitive to multiple levels of biological organization along with information-theoretic analyses that capture variations in behavior and signal acoustics. Despite good progress on many fronts, and a bright future, such analysis has yet to be achieved in any single system.

This chapter attempts to place the neuroethology of acoustic communication in the context of its multiple levels of analysis. Rather than striving to review the entire field, our approach is to focus on the neural and behavioral mechanisms that underlie what we consider to be a fundamental class of behaviors present in most vocal communication systems, namely vocal recognition. We have chosen specific examples to elaborate the points we make, and to some extent these choices reflect our own backgrounds. The choice of examples, however, is fundamentally immaterial because the evolutionary backdrop of neuroethology provides a logical framework for inferences drawn across phylogeny. In addition, although we are ultimately interested in questions of auditory perception and cognition, we do not incorporate the various literatures pertaining to the psychophysical basis of hearing and synthetic sound perception. Again, rather than any basic incompatibility, this reflects the practical constraints of the present format. A mature neuroethology of acoustic communication will complement the knowledge gained from research in other areas of comparative hearing and audition, which are reviewed elsewhere (e.g., Dooling et al. 2000; Feng and Ratnam 2000; see Simmons, Chapter 1).

1.1. The Neuroethological Approach

In one of his most important theoretical contributions to biology, Tinbergen (1963) pointed out that any given behavior could be understood in four primary ways. One can consider the evolutionary history of a behavior, its adaptive significance (i.e., its relation to survival or reproduction, past and present), its developmental history within an organism, and its physiological bases. The first two of these are often referred to as "ultimate" causes because they concern behavioral function on an evolutionary time scale. The latter two are often referred to as "proximate" causes of behavior because they are concerned, on a much finer time scale, with ontogenetic and physiological mechanisms within an organism. Although it is possible, and in fact common, to study behavior at only one or the other level, Tinbergen's heuristic is not meant to imply independence across levels. On the contrary, because evolutionary selection mechanisms operate on variation at the level of the individual, the natural history of a behavior, and its adaptive significance, are likely to have profound effects on underlying physiological and developmental processes. Likewise, the proximate mechanisms provide physical constraints for the evolutionary trajectory of a particular behavioral trait. The bidirectional interactions between ultimate and proximate mechanisms of behavior form the basis of neuroethology.

Neuroethology focuses on the mechanisms of ethologically analyzed behaviors by bringing a top-down approach to questions of brain function. By this, we mean that neuroethology takes a behavior whose adaptive significance has been studied in natural conditions (e.g., prey capture, food storage, vocal recognition) as the functional output of the central nervous system and then attempts to determine the underlying neural mechanisms. The neurobiological analysis may follow either a top-down approach (e.g., Moiseff and Konishi 1981) or a bottom-up approach (e.g., Heiligenberg 1991), but in either case the general methodology is predicated on prior knowledge about adaptive behavior.

Traditionally, neuroethologists interested in acoustic communication have investigated species with well-developed vocal systems and have focused on neuronal selectivity and specificity within the vocal repertoire. Given this focus, the extent to which the neural mechanisms of vocal communication in specialized animals can yield insight into more general mechanisms of audition should be considered. Historically, such consideration has led to criticisms that mimic those leveled against early ethology in that generalization is hampered by concentration on a limited component of the acoustic biotope (i.e., vocalizations) and by the choice of animals highly specialized for acoustic communication.

These criticisms fail at both the theoretical and empirical levels. Foremost, such criticisms are inconsistent with an evolutionary perspective maintaining that the perceptual world of each species is a unique consequence of its evolutionary history. As with morphological traits, similarities in behavioral phenotypes across species may result from either common origins or convergent evolution and so do not ensure a corresponding similarity in the underlying neural mechanisms. Comprehensive theories of behavior must therefore embrace, not ignore, evolution (e.g., Gallistel 1990; cf. Dickinson 1980) and must be allowed to emerge from comparative studies of many different species engaged in natural behaviors. A similar reasoning applies to the derivation of general neural mechanisms and their emergence through comparative studies. Moreover, as a matter of practical experience, the foregoing criticisms of the neuroethological approach to acoustic communication are not substantiated by the experimental literature. For example, among the most extreme cases of acoustic specialization is autocommunication in the contexts of echolocation and vocal learning. Yet many organizational features common to the vertebrate auditory (or octavolateralis) system have been usefully described—in many cases first elucidated—in relation to processing of vocalizations in echolocating bats, weakly electric fish, other fishes specialized for vocal communication, frogs, songbirds, and related systems. These principles include the forms and actions of parallel and hierarchical systems (including distorted tonotopic maps, feedforward, feedback, and lateral dynamic connections), distributed representations, single neurons with complex receptive field properties, temporal coding in single neurons and populations of neurons, sensorimotor interactions, state-dependent and dynamic receptive field properties, and sexual dimorphisms. Current studies of these systems continue to be highly productive and, if anything, reflect an expanding scope of issues successfully being addressed.

Finally, an additional question concerns the extent to which the neuroethological approach to acoustic communication can give insight into human acoustic perception. These concerns rest largely on the assumption that human perceptual processes, particularly those involved in speech and language, are unique. Nevertheless, determining which phenotypes of speech are unique and which are conserved requires comparative analysis. For example, although early work on speech processing emphasized what appeared to be unique properties of the human system, many of these initial claims evaporated in the face of later animal research (Kuhl and Miller 1975; Kluender et al. 1987). More recent theoretical perspectives emphasize that features of the speech signal, and associated neural processing mechanisms, are elaborations of general features of the mammalian and even vertebrate auditory system (see Fowler 1996; Lotto et al. 1997). Comparative studies have also demonstrated specialized sensory representations for autocommunication signals (see Popper and Fay 1995; Brenowitz et al. 1997). Such representations are likely to exist in humans because humans are well-known to be sensitive to the acoustic structure of their own vocalizations, but such representations have hardly been studied at all in humans. Furthermore, where there are specific theories of vocal learning in animals, they can lead to specific predictions regarding speech learning in humans (Margoliash 2001). These examples give confidence that insight into human speech perception can be gained by studying the perception of acoustically complex signals from a comparative perspective (Doupe and Kuhl 1999; Fitch 2000).

1.2. Structure of Animal Communication Systems

Arriving at explicit definitions for animal communication is a notoriously difficult problem. However, most will agree on a minimum description of communication as a process involving the transmission of information, via a signal, from a sender to a receiver. Where debate arises is in the extent to which various researchers attribute intent to the sender and in the degree to which the various fitness benefits for either the sender or the receiver are emphasized. Although such considerations remain a topic of continued debate among theoreticians (Dawkins and Krebs 1978; Beer 1982; Smith 1997; Bradbury and Vehrencamp 1998, 2000), from a proximate standpoint, it is the transmission of information that is of interest.

For information to be transmitted via some signal, there must be parity between the sender of that signal and the receiver. That is, the receiver must interpret at least some of the variability (i.e., information) in the signal in a predictable manner. Thus, the structure of a communication system can be considered as a behavioral feedback loop in which information flows

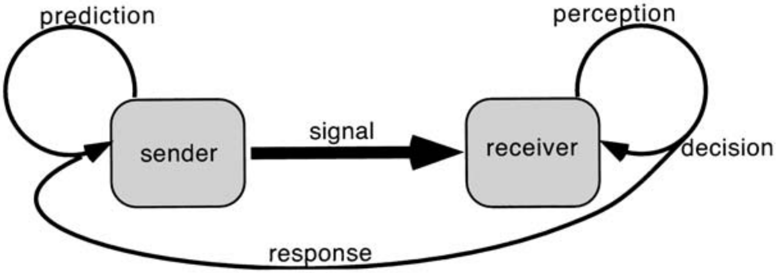

FIGURE 7.1. Diagram showing the theoretical cognitive relationship between sender and receiver in a communication system. Primary information is transmitted from the sender to the receiver in the form of a signal and from the receiver to the sender in the form of a behavioral response. The receiver must decide whether to engage in specific behaviors following the detection of some signal, and the sender must decide whether the receiver's behavior coincides with the predicted response for a given signal.

from the sender to the receiver and then back to the sender again, where the receiver's response is compared to some predicted behavior (Fig. 7.1). At the ultimate level, physiological mechanisms in the sender are under selective pressures to produce a specific variability in the signal, and those in the receiver are under pressure to perceive this variability in specific ways. Although particular circumstances may not always ensure mutual benefits to the sender and receiver, and thus that there is convergence upon specific signal characteristics, the general requirement for signal parity will hold as long as the signal remains functional. Conversely, given a functional signal, one can assume parity and thus that the production and perception mechanisms in communication systems are closely tuned to specific aspects of the acoustic variation in that signal. At the behavioral level, this is evidenced by that fact that specific functional behaviors come under the direct and strong control of acoustic variation in the signal.

One excellent example of the relationship between signal variation and behavior can be seen in female choice among songbirds. In many species of songbirds, males sing elaborate vocalizations (songs) that are often directed at females, and there is an extensive body of literature indicating that male song serves as the basis for female behavioral preference and choice in many species (reviewed in Searcy and Yasukawa 1996). Depending on the species, females choosing among individual males may attend to variation in song duration, dialect, output rate, "quality" (i.e., presence or absence of specific features), repertoire size, or measures of the acoustic complexity in a song. (Curiously, amplitude, which is so important in anuran and insect studies, has rarely been tested as a parameter for song preference and choice in female songbirds; see Dabelsteen and Pederson 1993; Searcy 1996.) Some of the behavioral variables describing male song may map simply onto neuronal mechanisms in females, whereas others may not. For example, in the case of repertoire–size preferences, it has been suggested that the acoustic entropy in the signal may drive neuronal habituation mechanisms (see Ryan 1998; Gentner et al. 2001) as opposed to more direct measures of repertoire sizes employed by researchers, such as simple song counts.

1.3. Technical Challenges

Given the attractiveness of vocal communication systems as research objects, it is somewhat surprising that the underlying neural mechanisms are relatively poorly understood. More than anything else, this lack of understanding reflects the tremendous technical challenges involved in studying these complex biological systems. Assessing perceptual responses to conspecific stimuli often requires elaborate conditions difficult to stage in the laboratory, and ethologists have only cautiously embraced conditioning procedures otherwise common in the study of perceptual phenomena (e.g., Dooling and Searcy 1980; Weary and Krebs 1992; Adret 1993). In part, such caution has been in an effort to avoid potential confusions that can arise when interpreting arbitrary conditioning paradigms in the context of natural behavior.

A related technical limitation is that the proper analysis of vocal communication will often require one or more animals to be actively engaged in communication as the experimenter records physiological activity. Ideally, this should include both field and laboratory settings. For the most part, this vision is more fanciful than real because the ability to reliably conduct single-neuron recordings in awake, behaving vertebrate animals of the small size commonly employed in neuroethological studies is only beginning to emerge (Dave et al. 1998b; Nieder and Klump 1999; Venkatachalam et al. 1999), so far mainly under relatively constrained laboratory conditions (Yu and Margoliash 1996; Dave et al. 1998a; Nieder and Klump 1999).

Yet another reason for the relative paucity of analysis of vocal communication is in some sense accidental and historical. Three of the most compelling vertebrate neuroethological systems studied from a vocal communication perspective—bats, weakly electric fish, and songbirds—involve autocommunication in the form of echolocation or feedback-mediated learning. Although some lessons and principles may emerge from the study of autocommunication that are common to sender/receiver systems, some clearly cannot. In autocommunication, the animal as the receiver has knowledge of the timing and structure of its own motor behavior (although an efference copy of the final motor output may not be directly available to the CNS—e.g., Heiligenberg 1977). An independent receiver cannot have such detailed knowledge. Thus, the signal coding and subsequent processing mechanisms, along with the role of attentional systems, are likely to vary between receivers engaging in autocommunication and those that are not.

In addition, in echolocation systems and passive localization systems, the structure of the vocal signal, or the parameters of the signal to be analyzed, are strongly and directly constrained by physical characteristics of the signal-generating mechanism, the medium for signal transmission, and the physical characteristics of the target. This means that sensory mechanisms for localization are influenced principally by natural selection. In all localization systems, there is compelling evidence for behavioral sensitivity to microsecond timing and considerable evidence for specialized neural systems to accurately represent information on such time scales. To date, there is no behavioral or physiological evidence to suggest that receiver sensitivity to microsecond timing is important outside the context of autocommunication. Vocal communication, on the other hand, generally arises in the context of reproductive behaviors such as courtship and territorial defense. Thus, vocal communication is influenced principally by sexual selection, although physical constraints can also influence the design of communication signals (Konishi 1970; Ryan and Kime, Chapter 5). The differences in selection pressures on different types of signaling may also have profound influences on CNS organization.

2. Organizational Principles for Neural Representations of Conspecific Vocalizations

Internal representations of information in the world are a function of both the form of the input and state of the neural system. Neuroethological studies using conspecific vocalizations bridge these "input" and "state" components in a biologically plausible way and have made important contributions to theories of representation and brain organization at multiple levels of analysis. In this section, we focus on the organizational principles of representation that underlie the discrimination and classification of vocal communication signals. These neuroethological analyses provide insight into behavioral mechanisms at both proximate and ultimate levels. At the proximate level, such studies relate to questions of neural coding, such as the forms of hierarchical and parallel representation, the representation of spectrally and temporally complex acoustic signals, and representational plasticity. Insights into ultimate processes that derive from neuroethological studies of conspecific vocalizations include coupling mechanisms between motor and sensory systems and the relationship between vocal recognition at inter- and intra-specific levels.

2.1. *The Feature Detector Hypothesis*

An understanding of the hierarchical and parallel organization of neural activity patterns is fundamental to the study of integrative brain mechanisms. Behavioral phenomena such as sensory perceptions may manifest themselves as particular states of activity in large aggregates or "assemblies" of neurons, with each neuron providing only a coarse coding of the stimulus in terms of a graded response (Hebb 1949). This is called the population hypothesis and has received particular attention in recent years with the emergence of technologies to record and analyze data from multiple electrodes in behaving animals (e.g., Eichenbaum and Davis 1998; Nicolelis 2001). Alternatively, 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 and is closely tied to the idea of "feature detectors" (Barlow 1972). The two hypotheses are not mutually exclusive, and although the early theoretical literature rejected the single-cell hypothesis (e.g., Marr 1982), in fact both theories may be considered substantially established. The analysis of conspecific vocalizations has helped in synthesizing a unified perspective.

2.1.1. Single Cells and Distributed Representations

In the extreme case, the single-cell doctrine has been characterized, or perhaps caricatured, as the "grandmother cell" hypothesis (see Martin 1994; Barlow 1995). Imagine a cell that responds always and only whenever you perceive the face of your grandmother and is a requisite for that perception. The grandmother cell concept is defined by necessity and sufficiency arguments in relation to the percept. However, where these constraints have been proposed in other contexts of neural coding, such as motor control, they have generally not proven to be satisfactory criteria (Kupfermann and Weiss 1978; Eaton 1983). The grandmother cell hypothesis has difficulty addressing the issues of combinatorial explosion, redundancy, and coverage. For instance, consider the case for a theoretical system containing N neurons. If the response of one (and only one) neuron codes for one stimulus, the system can only represent N stimuli. If precepts are represented at high resolution, or in combination, the number of percepts can easily overwhelm the number of neurons available to code those percepts. Loss of a single neuron would represent loss of the percept. These facts combined with the uncertain support in the experimental data suggest that the limits of a single-cell coding scheme are not realized, at least in vertebrates.

The extreme case of the distributed population-coding hypothesis is equally implausible. For the theoretical system containing N neurons, a fully distributed system could represent at least 2^N stimuli (given a binary response for each cell). Many more percepts could be represented than there were neurons, and loss of any one neuron or small set of neurons would result only in gradual degradation of the information represented in the network. These are desirable properties that superficially mimic biological neural networks, and highly interconnected networks are attractive to theoreticians because they are amenable to quantitative analysis.

Nevertheless, such coding schemes are biologically unrealistic. In the distributed model, neurons would be fully interconnected (i.e., there would be no anatomical specificity of feedforward, feedback, or lateral connections). There would be no topographic representation of information. Neurons would fail to show stimulus specificity. Such anatomical and physiological patterns of organization, however, are not observed in nervous systems, which exhibit precision of connections between and within different classes of neurons, with the spatial location of a neuron typically representing aspects of the neuron's specificity for parameters that are encoded. Thus, the extreme forms of the single cell and distributed-representation hypotheses are both rejected because they fail to reflect known biological reality.

The synthesis of the single-cell and distributed-coding hypotheses arises from an appreciation of the different end states of a behavioral continuum. Perceptions of arbitrary objects are unlikely to be processed by highly object-specific neurons. This conclusion derives from constraints on proximate mechanisms as described above and constraints on ultimate mechanisms, especially the unpredictable behavioral significance of an arbitrary object. For arbitrary objects, a distributed representation is likely, where the activation of ensembles of less specialized neurons is ultimately related directly to the perceptual event. In contrast, a predictable environment may allow for either genetic fixation or learning during ontogeny (or both) to establish specialized processing for reliable objects. There is considerable evidence for the existence of such specialized hierarchical streams of processing (see below) with mnemonic cells that exist at higher levels of these hierarchies. Rather than yielding a combinatorial explosion, extension of the hierarchical organization within the constraints provided by a predictable environment may increase coding efficiency. Perception is still ultimately related to the activation of ensembles of neurons, but the size of the ensemble may be different from that in the case of an arbitrary stimulus, and the neurons that encode the predictable stimulus may be more highly specialized. Thus, there is a balance among the various computational requirements associated with different behavioral requirements (Rolls 1992). Assessing neuronal variation across such behavioral variation is at the heart of the neuroethological approach.

In the context of the proposed framework, neurons act not as single-cell indicators of complete percepts representing components of signals that in combination result in perception of objects (cf. Barlow 1972), nor as featureless cogs in a vast distributed machine, but as localized feature detectors. In this proposal, the mechanisms for combining features may be single cells, population dynamics, or both. For more arbitrary objects, the features might be described by the statistics of natural scenes (Simoncelli and Olshausen 2001). Each species also lives in its own unique perceptual environment, and this serves to delimit the more complex features and combination of features that are represented at the level of single cells.

The evidence for feature detectors of specific signals is extensive, especially in acoustical-vocal systems (see below), which have been extensively studied from the perspective of natural stimuli. The evidence in other systems, especially vision, is not as complete, in spite of pioneering work on the frog visual system (Lettvin et al. 1959; Ewert et al. 1983a). Recent work suggests that combinations of visual feature detectors (columns) are a part of the process of object recognition (Tsunoda et al. 2001). Resolution of the issue of visual feature detectors may await more consistent application of neuroethological principles to vision research. There is strong suggestive evidence and considerable dispute, for example, regarding the existence of feature detectors in the context of visual face recognition in monkeys (e.g., Fujita et al. 1992; Gross 1992; Wang et al. 2000). Yet, far more of visual processing than face and hand recognition—in the case of monkeys, for instance, for fruits and other food sources, classes of predators, and classes of habitat—may be dominated by specialized mechanisms than is generally appreciated. Visually guided behaviors for these classes of stimuli provide logic for searching for corresponding biases in the visual system. The well-documented referential call system based on classes of predators in vervet monkeys (Seyfarth et al. 1980a) suggests an interesting parallel that could guide vision research.

Finally, if neurons in some pathways act as local feature detectors for certain classes of stimuli, then it becomes important to specify how subsets of those neurons interact to produce the percept. There is considerable theoretical work on population coding and the role of synchronous neuronal activity to dynamically "bind" simpler response profiles into global percepts (Singer and Gray 1995; see deCharms and Zador 2000) and some computational work that might suggest how simple subfeatures are combined to detect objects (Lee and Seung 1999). This is an area where a neuroethological approach could be quite advantageous but so far has only begun to be applied.

2.1.2. The Organization of Feature Detectors

In early single-neuron studies of sensory systems, neuroethologists observed specializations of the auditory system related to vocalizations of interspecific predators and courtship signals (Capranica 1965; Roeder 1966). These observations were interpreted in terms of a hierarchical organizational scheme of sensory systems, with neurons higher in the hierarchy having more selective responses related directly to behavioral output. Extensive data, collected from many systems following these early studies, now support the existence of both hierarchical organization schemes and feature-detector cells. The salient observations included tuning of peripheral responses to a behaviorally relevant range of parameters and cells at higher levels of a sensory hierarchy sensitive to particular combinations of spectral and temporal components of behaviorally relevant sounds. The extent of peripheral and central specialization is species- and behavior-dependent and, at least in some cases, can also be influenced by sex (e.g., Narins and Capranica 1976). It is noteworthy that peripheral and central specializations were predicted in the vigorous debates of the early ethologists (Lehrman 1953). Obviously, each species must vocalize in a range that it can hear, but this does not predict the observed peripheral specializations. For example, in several species of anurans, different frequency ranges are represented in different peripheral auditory organs (Capranica 1978). In some species of bats, the important second-harmonic region of the echolocation signal is highly overrepresented in the cochlea (Bruns and Schmieszek 1980). Peripheral specializations have also been observed in reptiles and birds (Manley 1990).

It follows that if the evolution of vocal signals can shape the action and distribution of peripheral receptors to match the vocal signal (or vice versa), then there are likely to be central effects as well. Indeed, in numerous systems, central neurons have been observed that respond selectively to specific spectral and temporal features of conspecific or autogenous (self-produced) vocalizations or are selective for conspecific vocalizations within a repertoire of sounds (Leppelsack and Vogt 1976; Mudry et al. 1977; Suga et al. 1978, 1979; Scheich et al. 1979b; Margoliash 1983; Rose and Capranica 1983; Rose et al. 1988; Rauschecker et al. 1995; Bodnar and Bass 1997, 1999; Crawford 1997; see Ewert et al. 1983b; Feng and Schellart 1999). In many cases, the specializations involve processing for spectral combinations and temporal sequences of sounds. These are reviewed separately in the next sections.

The feature-detector concept does not require stimulus and response invariance. Feature-detector cells may exhibit dynamic modulation of response properties (Zhang et al. 1997), in some cases in response to changes in behavioral state (Dave et al. 1998a), and more permanent plastic changes in response profiles (Knudsen 1985; Doupe 1997). These dynamic properties may complicate the analysis of the feature representations, but if the dynamic process is lawful, then it does not invalidate the concept of feature detection (cf. Manley and Müller-Preuss 1978). In the best-studied cases, the dynamic changes observed have been closely correlated with changes in behavior.

2.1.2.1. Parallel Hierarchical Pathways

Feature detectors imply hierarchical organization, but there is also extensive evidence that within a sensory system several parallel pathways may exist, each hierarchically processing different sets of sensory cues associated with different behavioral tasks (e.g., different aspects of recognition or localization). Such data have provided further support of the feature-detector hypothesis. The highly specific behavioral deficits in human patients with certain brain lesions are particularly compelling in this regard because they link regional localization with profound yet restricted perceptual specializations (Damasio et al. 1990). Animal studies complement this analysis by permitting controlled experimental manipulations. In electric fish, the magnitude and phase of the electric organ discharge are processed separately by different “P” (probability) and “T” (timing) classes of receptors that encode amplitude and phase, respectively, and distinct pathways arising from these receptor classes ascend the CNS until they converge at higher levels (Heiligenberg 1991). The neurons in the highest levels can be described as “recognition” neurons, whose activity is directly predictive of behavior (Rose et al. 1988). In bats, separate ascending systems of projections and multiple cortical areas appear to mediate the differential processing of constant-frequency and frequency-modulated components of the echolocation calls (Olsen and Suga 1991a, 1991b; see Casseday and Covey 1995), and lesion studies support this idea (Riquimaroux et al. 1991). Similarly, in barn owls, sound amplitude and timing information appear to be encoded in separate ascending pathways prior to convergence at the level of the midbrain (Moiseff and Konishi 1981; Sullivan and Konishi 1984). Again, lesions of each pathway produce specific behavioral deficits associated with loss of discrimination of one dimension in the parameter space but not the other (Takahashi et al. 1984). In frogs, the different frequency pathways that arise in the periphery are eventually combined at the level of single neurons (Fuzessery and Feng 1982, 1983). A similar conclusion appears to obtain for the primate auditory system, where separate ascending pathways exhibit differential sensitivity for stimulus morphology and location (Romanski et al. 1999; Rauschecker and Tian 2000). Collectively, these and other data provide strong support for a hierarchical-, modality-, and parameter-specific organization of vertebrate sensory systems (Ulinski 1984).

2.1.3. Combination Sensitivity

Neurons can exhibit processing for complex sounds by virtue of elaborations in spectral or temporal components of their receptive fields. In audition, the sequence of discrete events is critical for perception, and neuronal specializations for detecting sequences of events can be expected. One form of neuronal sequence sensitivity is called combination sensitivity. A neuron that is combination-sensitive responds with nonlinear summation (also called facilitation) to a combination of sounds (either discrete spectral components of the sound or two or more temporally discrete elements of a sequence) as compared with the response of the neuron to subsets (typically individual components) of the sound. There is no standard for what a “sufficient” nonlinear response is for a neuron to be considered combination-sensitive as long as the additional nonlinear component reaches statistical significance. Neuronal combination sensitivity may be far more common than has been appreciated (e.g., Brosch and Schreiner 2000;

Gehr et al. 2000). Different magnitudes of the nonlinear components of the combination response, and different temporal windows of integration, may reflect different mechanisms that give rise to combination sensitivity.

2.1.3.1. Spectral Combination Sensitivity

Neurons with spectral combination sensitivity are reported in a broader range of species than are neurons with temporal combination sensitivity. This may reflect, at a neurobiological level, the fundamental nature of the process of integration of spatial information arranged across a sensory epithelium and, at a behavioral level, that sensitivity to temporal structure is probably a secondary adaptation. Neurons that require two or more spectral lines before exhibiting a facilitated, typically excitatory response have been well-characterized in bats (Suga et al. 1979), birds (Langner et al. 1981; Margoliash and Fortune 1992), cats (Sutter and Schreiner 1991; Nelken et al. 1994a, 1994b), and frogs (Mudry et al. 1977; Fuzessery and Feng 1982, 1983), and there is weaker evidence in other systems as well. The responses have almost always been linked to specific spectral components of vocalizations, whereas spectral combination sensitivity was not observed in the auditory cortex of monkeys trained on a missing-fundamental task (Schwarz and Tomlinson 1990). Spectral combination-sensitive neurons can be common throughout the auditory system in some animals (bats), or locally common within a specific neural structure (frogs, cats) or within a specific neural pathway (songbirds). Because these response properties can be difficult to identify without the adequate stimulus, the failure to find spectral (or temporal) combination-sensitive neurons cannot be taken as proof that such response properties are absent.

The mechanisms for producing spectral combination-sensitive neurons are probably best described from extracellular recordings conducted in the echolocating CF-FM mustached bat, *Pteronotus parnelli*. Members of this species produce a biosonar "pulse" for orientation composed of a constant-frequency (CF) and frequency-modulated (FM) component, each of which has four harmonics (Fig. 7.2). The returning echoes of these pulses are delayed temporally, reflecting target distance, and have a Doppler shift in frequency based on the relative velocity of the target, typically an insect. Periodic frequency (and amplitude) modulations in the echo may result from insect wing beats. Many types of combination-sensitive neurons have been described in this system, and the auditory cortex can be divided into multiple subregions based on the particular combinations of pulse and echo components that give rise to facilitated responses. For example, neurons in the CF_1/CF_3 subregion are tuned to the first harmonic of the pulse and the third harmonic of the CF component in the returning echo (see Fig. 7.2). Within this region, specific cells are tuned to specific frequency differences between the pulse and echo and/or phase-locked to frequency modulation of the harmonic echo, representing a specific relative target velocity or

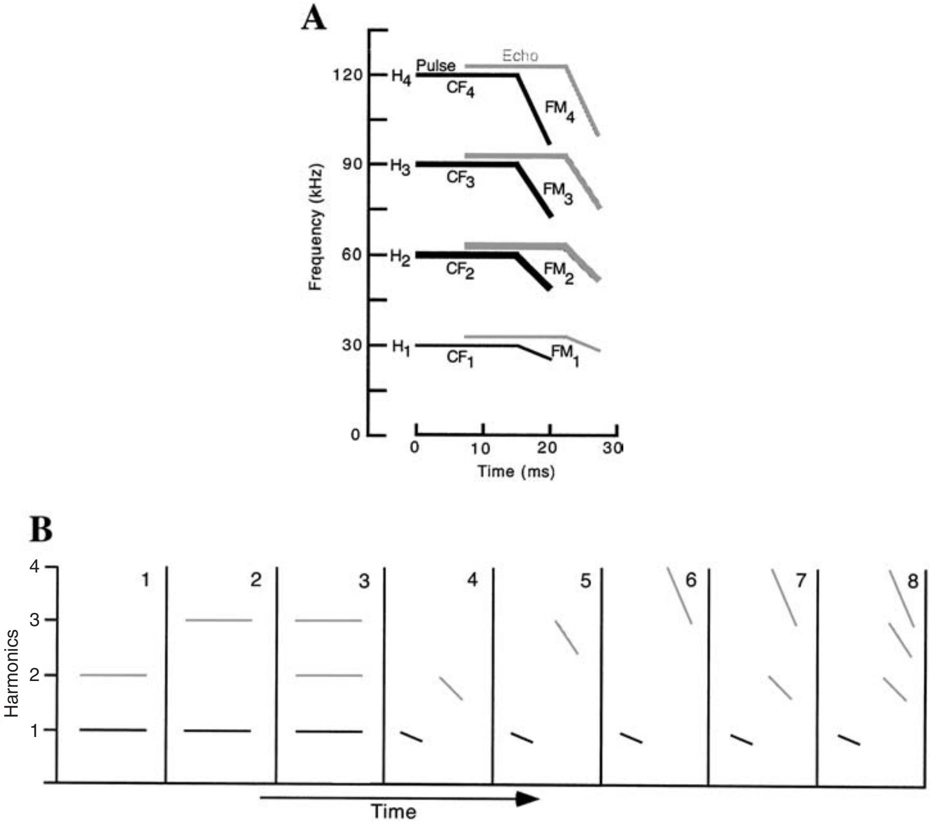

Figure 7.2. Schematic spectrograph (**A**) of the echolocation signal of the mustached bat, *Pteronutus parnellii*, showing the constant-frequency (CF_{1-4}) and frequency-modulated (FM_{1-4}) components for each of the four harmonics (H_{1-4}) of the pulse (black lines) and the echo (gray lines). Note that the echo is delayed in time and Doppler-shifted in frequency compared with the pulse. In (**B**), schematic representations of the signal components that facilitate common forms of spectral and temporal combination sensitivity are shown. Pulse harmonics are in black and echo harmonics in gray. (After Suga et al. 1998.)

beating wings, respectively. These multiple regions form a frequency versus frequency topographical representation that systematically maps relative target velocity. Another class of spectral combination-sensitive cells is found in the Doppler-shift constant-frequency (DSCF) region of the cortex. These neurons are tuned to very fine-grain frequency and amplitude modulations of the CF_2 component (see Fig. 7.2) and may contribute to determining target angle along with velocity (Riquimaroux et al. 1991; Kanwal et al. 1999). There are additional cortical areas in mustached bats that exhibit combination sensitivity, including some that respond to unusual combinations of stimulus components (e.g., Tsuzuki and Suga 1988;

Fitzpatrick et al. 1998). Different neural mechanisms may underlie processing of echolocation signals in different species of CF-FM bats and in different classes of bats (e.g., FM bats), but this is not reviewed here (see O'Neill 1995).

Spectral combination-sensitive neurons are also present in the ascending auditory system of mustached bats at the level of the medial geniculate body of the thalamus (MGB) (Olsen and Suga 1991a; Wenstrup 1999) but have not been found in the inferior colliculus (O'Neill 1985). This led to the hypothesis that the spectral combination sensitivity originated at the level of the MGB as a result of converging inputs from inferior collicular (IC) neurons tuned to specific frequencies in the sonar pulse and higher harmonics in the returning echoes. Recent findings, however, have challenged the idea that the IC is the primary source of converging input to the MGB (Wenstrup and Grose 1995), so additional study is clearly needed.

One potential source of combination-sensitive responses in the MGB is descending cortical projections. Corticofugal projections originating in layers V and VI of the auditory cortex project to both the IC and the MGB as well as subcollicular auditory nuclei. These projections appear to maintain tonotopic organizations such that high-frequency tuned regions of the cortex project to high-frequency tuned regions of the MGB and IC, for example. Although it is clear that these projections enhance the tuning of cortical cells through feedforward stimulation of frequency-matched subcortical units and the reduction of responses in unmatched subcortical units (Suga et al. 1997; Zhang and Suga 2000), their role in the generation of combination-sensitive responses has only recently come under investigation. The complexity of this corticofugal system is compounded by the presence of both direct excitatory connections and inhibitory connections via the thalamic reticular formation (see Wenstrup 1999).

2.1.3.2. Temporal Combination Sensitivity

Temporal combination-sensitive (TCS) neurons have been well-described in bats and birds, and there is some evidence for TCS neurons in other systems (primates: Olsen 1994; cats: Brosch and Schreiner 2000; Gehr et al. 2000). The behavioral requirements for temporal combination sensitivity differ dramatically for songbirds and bats, and the observed TCS response properties are tightly linked to the behaviors. This suggests that in systems where the animal's behavioral requirements are not sufficiently understood, it might be possible to overlook neurons with a TCS response for lack of use of the appropriate stimulus.

In mustached bats, TCS neurons have been described in the ascending auditory system, starting at the level of the inferior colliculus (Mittmann and Wenstrup 1995; Leroy and Wenstrup 2000) and continuing in the medial geniculate nucleus (Olsen and Suga 1991b; Wenstrup 1999) and in multiple cortical areas (e.g., Suga et al. 1978; Suga and Horikawa 1986; Taniguchi et al. 1986). Cortical neurons typically require FM_1–FM_N ($N = 2, 3, 4$) stimuli, that is, the fundamental FM component and one of the higher harmonics. The neurons typically are tuned to a fixed delay between the two FM components, in which case they are sensitive to a particular target distance (echo delay). In some cases, FM/FM neurons respond to systematic changes in delay between the two FM components, which could facilitate tracking an approaching target such as an insect (O'Neill and Suga 1979). As with the CF/CF neurons, FM/FM neurons are organized in a topographic manner. In the cortex, each pair of components (i.e., FM_1 paired with $FM_{\{2|3|4\}}$) is mapped in a separate subregion. The principal parameter mapped within each FM/FM subregion is delay. That is, time delays (which relate to target distance) are systematically mapped within the cortex (O'Neill 1995).

Recent data from the mustached bats suggest that TCS facilitation first arises at the level of the inferior colliculus (Portfors and Wenstrup 1999; Wenstrup et al. 1999) and is mediated via glycineric inhibitory mechanisms (Wenstrup and Leroy 2001). Neural tuning for sound duration also appears to arise via inhibitory processes in IC (Casseday et al. 1994), and neurons in IC are also tuned to periodic frequency modulations according to the rate and amount of the FM (Casseday et al. 1997). The data support a model in which duration tuning arises from the interaction among several different types of subthreshold inputs and suggest that similar mechanisms may give rise to FM tunings as well (Covey 2000). It may be the interaction of these processes at the population level that gives rise to TCS facilitation. Here again, corticofugal projections may play a role as their effect on the delay tunings for FM–FM combination-sensitive neurons has recently been shown (Yan and Suga 1999).

In songbirds, TCS neurons have been described in forebrain components of the song vocal-control system. TCS neurons in songbirds are sensitive to sequences of notes or syllables (Margoliash 1983; Margoliash and Fortune 1992; Doupe 1997; Dave and Margoliash 2000). Starting with what appears to be the first song system nucleus with auditory responses (Janata and Margoliash 1999), all forebrain areas of the song system are selective for acoustic features of the individual bird's own song. It is probable that each of these areas contains TCS neurons, although this has yet to be tested.

The extracellularly described properties of TCS neurons differ dramatically between birds and bats. Whereas in mustached bats the range of time delays represented in the neuronal population spans roughly 0.4–18 msec (corresponding to typical echo delays experienced during predation), the time course for TCS cells in birds is much longer (~80–350 msec) and can include many song syllables (Margoliash 1983; Margoliash and Fortune 1992; Dave and Margoliash 2000). Delay tuning of TCS neurons is assessed by systematically varying the interval between the first and second components that, in combination, result in the TCS response. When such tests were conducted in white-crowned sparrows (Margoliash 1983) or zebra finches

(Margoliash and Fortune 1992), the memory of TCS neurons was observed to extend up to approximately 300–500 msec. The long integration periods and high degree of stimulus specificity place the TCS cells described in songbirds among the most complex auditory neurons known.

Originally, a simple model of the TCS response in birds was proposed whereby release from inhibition resulted in the nonlinear component of the response (Margoliash 1983). Data from intracellular studies found evidence for that model but also for other interacting subthreshold mechanisms similar to those described in the bat IC, along with additional threshold nonlinearities (Lewicki and Konishi 1995). Whereas axonal delay lines may account for the shortest-time-scale TCS responses in bats (Kuwabara and Suga 1993), the much longer time scales of the TCS response in songbirds cannot be accounted for by axonal delays. In songbirds, recent data show that different subthreshold responses are associated with two distinct populations of projection neurons and a population of interneurons in the song-control nucleus HVc (Mooney 2000). TCS responses are thought to be present for at least some neurons in both classes of projection neurons and for the interneurons. Thus, as in bats, multiple mechanisms in birds may be responsible for the TCS response in different classes of neurons. One difference between TCS neurons in CF-FM bats and birds is that, in the latter, the neurons are predominately found in vocal-control areas—areas that also participate in generating motor output. This suggests the hypothesis that a relatively undifferentiated auditory input to song-system nuclei might be patterned in interaction with central pattern generators for song that are shaped by song learning (Margoliash et al. 1994). TCS responses are apparently rare in the ascending auditory system of birds, whereas they are common in the ascending auditory system of bats.

2.2. *Input Constraints on Representational Systems*

The search for specializations in the auditory system related to vocal behavior has often resulted in identification of complex, nonlinear neural-response properties such as the combination-sensitive neurons described above. However, practical limitations on stimulus power, stimulus and recording duration, and stationarity, coupled with the high-order nonlinearities central neurons typically exhibit, have handicapped quantitative approaches to nonlinear analysis (e.g., white-noise analysis). As a result, generalized procedures to characterize nonlinear neuronal responses have not been established, and linear techniques to analyze complex, especially natural, stimuli are being developed (e.g., spike-triggered receptive fields: Klein et al. 2000; Theunissen et al. 2000; spike-based stimulus reconstruction: Rieke et al. 1996). Information-theoretic approaches can describe nonlinear statistical properties of neurons, but these often require more data than is practical to collect. Furthermore, the ultimate utility of information-theoretic descriptions is not yet resolved. From this perspective, the use of natural stimuli in neurophysiological experiments is very valuable. Although the theory underlying their use is not yet well-developed in a formal sense, natural stimuli realistically constrain the search space to a range of input parameters that, because of their behavioral relevance, are likely to be represented by central neurons.

The use of natural stimuli to characterize neuronal responses is somewhat analogous to the application of psychophysical paradigms to characterize neuronal responses. Just as a psychophysical illusion may identify new neuronal response properties (von der Heydt et al. 1984), so may presentation of a new category of natural sounds identify new components of receptive fields, even in cases where neurons have previously been extensively studied. For example, although bat cortical neurons had been extensively studied in the context of echolocation, entirely new excitatory regions of their receptive fields that had previously been missed were uncovered when social communication calls were first presented to those neurons (Ohlemiller et al. 1996). The receptive field domain suitable for communication calls may represent the plesiomorphic condition. The existence of two separate receptive field domains highlights the degree of nonlinearity or high dimensionality (complexity) of the parameter space and emphasizes that the choice of stimulus repertoire can be a subtle decision that shapes the limitations of the experimental paradigm. Beyond this obvious caveat, procedures have been established for using natural stimuli in neurophysiological experiments.

2.2.1. Selectivity and Specificity for Conspecific Vocalizations

Two strategies have been employed to assess the responses of neurons to species-typical vocalizations. The first is to use natural vocalizations and derivative sounds to identify the potential behavioral relevance of neural responses. This is assessed in terms of a neuron's selective responses to a subset or category of natural vocalizations (i.e., its selectivity). The second strategy is to identify the specific acoustic features underlying a neuron's selective responses (i.e., its specificity) by using artificial stimuli to characterize its response properties. These two strategies are complementary, and both may be necessary to adequately describe a neuron's responses to natural stimuli. Testing for selectivity rarely unambiguously identifies the specific acoustic features that result in the selective response. Conversely, exploration of specific features of a neuron's response properties may give only limited biological insight into the significance of those features in the absence of a well-delineated behavioral context. Absent knowledge of the acoustic behaviors, neural analyses may ultimately emphasize response parameters that are not central to the behavioral decisions associated with real-world tasks such as vocal recognition.

In cases of complex vocal behavior (i.e., where multiple acoustic parameters affect the behaviors under study), an experimental approach is mandated that renders the distinction between selectivity and specificity ultimately arbitrary. Such cases require the combination of selectivity and specificity approaches. To accomplish this, the relevant natural vocalizations are systematically decomposed into simpler signals. At the same time, artificial sounds are used to synthesize increasingly closer approximations (models) of the natural vocalizations. Optimally, both decomposition and synthesis procedures are sensitive to behaviorally salient variation identified in the natural vocalizations. The goal is to bring the decomposed and synthetic stimuli to some common intersection and thus to establish a logical relationship between the variation in neuronal response and the variation in the stimuli.

Combining the approaches of modified natural and artificial stimuli need not result in convergence on the same solution—different parameter spaces may be identified using the two approaches. In addition, errors are possible, because there are practical limits to the size of a repertoire that depend on the stability of the recordings and the natural rate of stimulus repetition. In the absence of a specific model, these limits may prevent the choice of a sufficient repertoire in cases of large and complex vocal repertoires or highly selective neuronal responses. Errors in analysis may also result from experimental decisions about the appropriate stimuli that have to be made online without the benefit of retrospective analysis. These can be exceptionally challenging and exciting experiments! In the limited number of cases where this approach has been employed, data from natural and artificial stimuli have converged upon common solutions (e.g., Margoliash 1983). When the two approaches converge on a common set of parameters, this gives confidence that a uniform model of the neuron's response profile has been achieved. In this case, the derived model can account for the neuron's response in terms of a specific set of acoustic parameters in the natural vocalizations.

The success of a receptive-field model of neuronal response selectivity can be independently determined by quantitative predictions of response selectivity based on acoustic specificity that in turn can suggest specific predictions regarding cellular mechanisms that give rise to the selective responses. Achieving such quantitative predictions may itself require a significant modeling effort and is not frequently attempted. For example, the responses to songs of some avian auditory thalamic neurons can be quantitatively predicted from their responses to tone and noise bursts (Bankes and Margoliash 1993; Anderson et al. 1996). Achieving this result required extensive testing to find the appropriate model (network) architectures. Interestingly, the predictive power of the models was most sensitive to the dynamics of the neuronal responses to the artificial stimuli. Model output was less sensitive to manipulation of traditional static descriptions of neuronal response such as the frequency-amplitude response curves and rate intensity functions.

2.3. State Constraints on Representational Systems

Selection processes shape the acoustics of vocal communication signals. Constraints that result from these pressures are not restricted to the signal (i.e., input) acoustics but also the organization (i.e., state) of the representational system. In the following section, we consider a number of ways in which the organization of the representational system is constrained by both phylogenetic and ontogenetic processes. These constraints include potential links between motor and sensory representations, learning, and the relationship between species and subspecies-level recognition behaviors.

2.3.1. Motor-Sensory Linkages

2.3.1.1. Genetic Coupling and Coevolution

Signal production and perception are linked behaviorally (see Section 1; Fitch and Hauser, Chapter 3; Ryan and Kime, Chapter 5). One possible explanation for this parity is that the production of a class of sounds and perception of sounds of the same class are tightly linked within each individual. Such an arrangement could result from genetic coupling, wherein the neural mechanisms for production and perception of a signal would share common elements controlled by the same set of genes (Alexander 1962). Thus, modification of the genetic material would result in concomitant changes in production and perception. Although the simplicity of this idea is attractive, there is no guarantee that the changes will be coordinated or that the modified element will contribute the same change, in the same direction and magnitude, for both modalities. Rather than postulate a single genetic basis for separate production and perception systems, a more parsimonious solution is to postulate a single central pattern generator responsible for both production and perception (Bentley and Hoy 1972; Hoy et al. 1977). The predicted role of central pattern generators in perception represents an additional prediction of the genetic-coupling hypothesis.

Alternatively, the match between perception and production could result from coevolution (von Helversen and von Helversen 1994). By this account, the mechanisms of production and perception of communication signals are genetically independent but evolve under reciprocal selective pressure toward representation of a common set of communication features. Coevolution may arise through the common effects of a shared environment. For example, males in the gray tree frog species *Hyla versicolor* and *H. chrysoscelis* both produce pulsatile calls to which gravid females are attracted (Gerhardt 1982). Although the mean pulse rate varies between species, it is also temperature-dependent, such that a male *H. chrysoscelis* calling at 20°C could have a pulse rate similar to a male *H. versicolor* calling at 15°C. In areas where the species are sympatric, this temperature-dependent property of male calls creates the potential for mating errors.

However, female preferences for calls within the conspecific frequency range are also temperature-dependent (Gerhardt and Doherty 1988). Because these temperature dependencies scale differently under similar temperature ranges, they are likely to be controlled by different neural mechanisms (see van Dijk et al. 1997). Of course, genetic coupling and coevolution are not mutually exclusive.

2.3.2. Motor-Sensory Linkages in Learned Signals

Linkages between production and perception can also result from developmental changes that selectively modify phenotypic patterns. A compelling example is the development of human speech. It has long been known that human infants can distinguish many of the phonetic contrasts in speech, even those in languages to which they have not been exposed (e.g., Lasky et al. 1975). Adult speakers of different languages, on the other hand, not only differ from one another in the location of perceptual boundaries between phonetic contrasts but also lose altogether the ability to perceive some nonnative contrasts (e.g., Best et al. 1988; Logan et al. 1989). The loss of sensitivity to selected nonnative contrasts occurs somewhere near the end of the first year (Werker and Tees 1984).

The developmental changes in speech perception are roughly matched to developmental changes in vocal production, and such findings have been taken to support the theory (Liberman et al. 1967) that speech perception is guided by a process that matches speech sounds to the vocal gestures required to produce those sounds. This is the so-called "motor" theory of speech perception. As with the genetic-coupling hypothesis, in motor theory, the same pattern generators participate in perception and production. The modern version of the motor theory of perception (Liberman and Mattingly 1985) stresses representational modularity as conceived by Fodor (1983).

Psychologists have staged a series of strong assaults on the motor theory (e.g., Lindblom 1991; Fowler 1996; Nearey 1997). Because the motor theory of speech perception appears to make clear predictions regarding the recruitment of specific motor pathways in speech perception, it would appear possible to test predictions of motor theory with neurobiological techniques. Unfortunately, instantiations of the motor theory of speech perception do not specify the neural mechanisms implied by requisite processes such as vocal gesture, information encapsulation, and so forth. Absent these precise definitions, motor theory is not falsifiable at a neurobiological level. In addition, experimental neuronal data from humans with the spatial and temporal precision required to test theories of perception such as motor theory cannot yet be obtained.

The comparative approach can help resolve questions that might otherwise be difficult to resolve with the available human data. Humans are phylogenetically isolated with regard to the central feature of interest, vocal learning. This limits the effectiveness of a phylogenetic analysis (Fitch 2000). Nevertheless, computational problems are shared by all biological systems that exhibit vocal learning, which may result in similar mechanistic solutions (Margoliash 2001). The following section describes production and perception of song in relation to various pathways in the avian song system.

2.3.3. Production-Perception Linkages in Birdsong

The bird song system provides an experimental model system amenable to detailed neurobiological approaches with which to pursue questions of production/perception linkages in the context of a learned vocalization. Comparisons of neural responses to conspecific songs with responses to a bird's own song(s) explore the hypothesis that a bird's own song holds a position of particular perceptual significance within the representational system. This hypothesis is well-supported by a variety of data. For instance, several field playback studies demonstrate that aggressive responses from territorial males are often dependent on whether the particular playback song is similar to one that the subject sings (see Owings and Morton 1998). In most species, playback of the individual's own song elicits a level of aggressive response intermediate to the weakly aggressive responses to songs of established neighbors and strongly aggressive responses to unfamiliar (strangers') songs (Stoddard 1996). The salience of a bird's own song is also apparent under more controlled conditions, as for example when examining acquisition rates or category formation in operantly trained tasks (Cynx and Nottebohm 1992; Gentner and Hulse 1998) or behavioral preferences (Pytte and Suthers 1999; Okanoya et al. 2000). Together, such data suggest that some form of coupling exists between the song-production and perception systems.

2.3.3.1. The Vocal Motor Pathway

The extensive analysis of the avian song system provides a basis for understanding how different components of the system contribute to perceptive and productive components of behavior. Several forebrain pathways have been identified as components of the song system, of which two have been extensively analyzed (Fig. 7.3). A vocal motor pathway (VMP) is required for song production, participates in moment-to-moment control of singing, and represents the descending motor outflow of the forebrain song system. An anterior forebrain pathway (AFP) shares similarities with the mammalian corticothalamo-basal ganglia loop. The AFP is required for normal song development and may be necessary for adult song maintenance, although contributing little (depending on species) to moment-to-moment control of singing. The relative contributions, if any, of the VMP and AFP to components of perceptual processing of conspecific songs would represent different schemes for establishing perception/production linkages, a

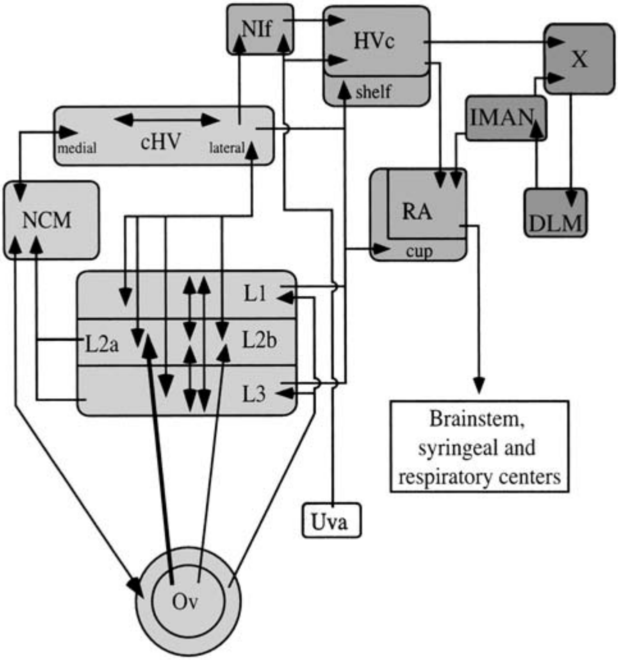

FIGURE 7.3. Diagram indicating the primary auditory pathways (light gray), vocal motor pathway (VMP, gray), and anterior forebrain pathway (AFP, dark gray) in songbirds. Ov: nucleus ovoidalis; L1–L3: field L complex; NCM: caudomedial neostriatum; cHV: caudoventral hyperstriatum; NIf: nucleus interfacialis; Uva: nucleus uvaeformis; RA: robust nucleus of the archistriatum; lMAN: lateral magnocellular nucleus of the anterior neostriatum; DLM: dorsolateral region of the medial thalamus; X: area X; HVc is used as the proper name. Ov comprises a core and a surround. HVc and RA have subjacent regions, the "shelf" and "cup," respectively.

direct connection with motor output in the case of the VMP, or a shared developmental trajectory in the case of the AFP.

In the VMP, the roles of the nuclei HVc and RA in song production have been well-established through electrolytic lesions (Nottebohm et al. 1976), electrophysiological recordings in singing birds (McCasland and Konishi 1981; Yu and Margoliash 1996), and perturbation experiments using electrical-stimulation (Vu et al. 1994). Zebra finch songs are composed of repeated motifs, which are sequences of different syllables. Each syllable typically comprises several notes. Recordings of single neurons in singing birds demonstrate that premotor activity, in what are probably HVc interneurons (see below), is organized by the syllable type. That is, repetitions of the same syllables are associated with similar activity patterns, independent of position within the song, whereas different syllables are associated with different patterns of activity (Yu and Margoliash 1996). This results even if the syllable types share some or many notes. RA neurons, which receive from HVc, exhibit patterns of physiological activity that contrast with those of HVc neurons. Premotor activity of RA neurons is organized by note type, independent of the syllable in which the note is embedded (Yu and Margoliash 1996).

The syllable is probably the largest unit of vocalization that represents a motor program. When singing birds are startled, they tend to complete the current syllable before ceasing to sing (Cynx 1990). Upon electrical stimulation of HVc during singing, birds tend to restart the song at the beginning of motifs (i.e., at large segmental boundaries representing repeated sequences of syllables). In contrast, stimulation of RA during singing disrupts the syllable morphology but not the sequence of syllables (Vu et al. 1994). Thus, information about the large-scale organization of song (sequences of motor programs) could be available to the VMP and the AFP through the two distinct classes of HVc projection neurons; HVc projection neurons target exclusively either RA or area X (Fig. 7.3).

The different classes of HVc projection neurons, and HVc interneurons, exhibit differences in their intrinsic properties and subthreshold auditory-response properties (Mooney 2000). Furthermore, recent data from electrical-stimulation studies designed to establish identities of extracellularly recorded classes of HVc neurons suggest that most of the chronic recordings were of HVc interneurons (Shea et al. 2001). Thus, it remains to be seen whether the mapping from syllable-level to note-level representations occurs within HVc (at the level of the RA-projecting HVc neurons) or in the projection of HVc to RA. The HVc to RA mapping potentially represents a transition from categories of behavior to motor output. It is likely, in any case, that the VMP and the AFP receive different mixes of premotor activity and sensory feedback during singing.

The physiological properties of other VMP nuclei contrast with those of HVc and extend the concept of a hierarchical arrangement of sensorimotor control. Neurons in NIf and Uva (Fig. 7.3), which project to HVc, are active during singing (McCasland 1987; Williams and Vicario 1993). Multiunit recordings from Uva suggest that some activity is more closely associated with the timing of entire motifs, and bilateral lesions of Uva disrupt the suprasyllabic organization of song but do not abolish singing (Williams and Vicario 1993). Recent data from zebra finches suggest that NIf is a major source of auditory input to HVc (Janata and Margoliash 1999; Boco and Margoliash 2001). Bilateral lesions of NIf have phrase-level effects on singing in Bengalese finches (Hosino and Okanoya 2000) and apparently only transient effects on singing, with full recovery, in zebra finches (Vu et al. 1995). Thus, lesions of VMP nuclei afferent to HVc affect but do not abolish singing whereas lesions of HVc or RA abolish all singing behavior.

2.3.3.2. Auditory Responses in HVc

Neurons in HVc show auditory responses; in particular, they exhibit selective responses to the bird's own song compared with conspecific or parametrically modified songs (Katz and Gurney 1981; McCasland and Konishi 1981; Margoliash 1983, 1986, 1987; Margoliash and Konishi 1985; Margoliash and Fortune 1992; Volman 1993; Lewicki 1996; Yu and Margoliash 1996; Theunissen and Doupe 1998; Mooney 2000). The selective responses of song-system neurons have motivated the hypothesis that HVc and the AFP may be involved in production-independent perceptual processing of songs that is influenced by representations of a bird's own song (Margoliash 1986, 1987). This hypothesis suggests developmental linkage of responses to a bird's own song with motor output, without invoking direct motor processing during perception.

Studies of HVc-lesioned female canaries are consistent with a role for HVc in song perception. Lesions to HVc abolish female behavioral preferences for conspecific over heterospecific song (Brenowitz 1991) and for sexually attractive song phrases over other phrases of conspecific song (Del Negro et al. 1998; and see Vallet and Kreutzer 1995). That is, although only conspecific song elicits copulation displays in prelesion females, all songs (conspecific and heterospecific) elicit solicitations following HVc lesions. Multiunit recordings from HVc in sexually receptive female canaries have identified inhibition that is restricted to presentations of sexually attractive conspecific song phrases (Del Negro et al. 2000). Thus, there may be a linkage between physiological inhibition and that observed in behavior. Because the projection of area X onto DLM (Fig. 7.3) is probably inhibitory (Luo and Perkel 1999), this suggests an overall release from inhibition in the rest of the AFP (DLM and lMAN) that is proportional to the response in HVc. In female canaries, the degree of release would be proportional to the sexual potency of the stimulus.

The role of HVc in song discrimination may be species-dependent. In both male and female starlings, HVc lesions do not affect the retention of learned conspecific song discriminations but do affect the ability to form new associations with familiar songs (Gentner et al. 2000). In contrast to the studies in canaries and starlings, in female zebra finches, *Taeniopygia guttata*, lesions to HVc have no effect on copulation solicitations (MacDougall-Shackleton et al. 1998a). Qualitative differences in sexual dimorphism between zebra finches and other species, especially the magnitude of HVc projections onto RA and area X, may help to explain these differences (Nottebohm and Arnold 1976; Gurney 1981). The data suggest that rather than providing a direct role in perceptual representation, HVc may be acting as a selective filter regulating motor control of subsequent behaviors, including copulation responses in females and song output in males. Under this hypothesis, HVc filtering, and hence the perceptual processing of song, varies across species depending on the degree of sexual dimorphism in the brain of the animals.

2.3.3.3. The Anterior Forebrain Pathway

A second pathway in the songbird forebrain involves an avian corticobasal ganglia-thalamocortical pathway (Bottjer and Johnson 1997; Farries and Perkel 2000; Perkel and Farries 2000). This three-nucleus pathway makes an obligatory contribution to song development (Bottjer et al. 1984; Sohrabji et al. 1990; Scharff and Nottebohm 1991). In zebra finches, the AFP has also been shown to contribute to song maintenance (Williams and Mehta 1999; Brainard and Doupe 2000). Although neurons in the AFP in zebra finches show premotor activity during singing (Jarvis and Nottebohm 1997; Hessler and Doupe 1999), lesions of AFP nuclei—in contrast to lesions of HVc or RA—have little disruptive effect on motor output (Nordeen and Nordeen 1993). In Bengalese finches, partial lesions of an AFP nucleus result in the transient disruption of song (Kobayashi et al. 2001). The premotor activity in the AFP has been interpreted as an efference copy signal that combines with auditory feedback to create an error signal that helps stabilize vocal output (Williams and Mehta 1999; Brainard and Doupe 2000, 2001; Solis et al. 2000). The differences between zebra and Bengalese finches may result from the apparent greater reliance on auditory feedback, and hence a greater effect of the AFP, in song maintenance in adult Bengalese finches (Okanoya and Yamaguchi 1997).

Different populations of HVc neurons project to RA in the VMP and area X in the AFP. Thus, the apparent role of HVc in some aspects of song perception does not distinguish between potential differential roles of the VMP and AFP in song perception. If the VMP is involved in song perception, this would provide strong evidence for a "motor" theory of birdsong perception (Williams and Nottebohm 1985) and its related theory for speech perception by reference to production (Liberman et al. 1967; Liberman and Mattingly 1985). Alternatively, the perceptual role of HVc could be mediated through its AFP projections. In this case, the observed linkage between production and perception would arise through a shared developmental history, not through direct moment-to-moment coupling.

The data appear to support the ontogenetic theory. In what is perhaps the most convincing evidence that the song system is involved in non-BOS acoustic stimulus recognition, lesions to lMAN in female canaries disrupt the retention of auditory discriminations (Burt et al. 2000). These lesions effect discrimination among pairs of conspecific songs, heterospecific songs, and synthetic sounds, but not among pairs of visual stimuli, and therefore suggest a general auditory processing role for the AFP rather than one specifically related to conspecific songs or a bird's own song. In contrast, the effects of lesions to the AFP in adult male zebra finches appear to be restricted to conspecific discriminations involving a bird's own song (Scharff et al. 1998). However, the later study only examined acquisition rates, not retention, and used a small stimulus set. Thus, it remains to be seen whether general auditory effects are observable in male songbirds, where the AFP shows clear selectivity to BOS (Doupe 1997).

Volumetric studies are consistent with the lesion work in suggesting a role for the AFP in song perception, particularly among females. In adult female brown-headed cowbirds, the volume and number of neurons in lMAN is positively correlated with the bird's ability to discriminate among male songs (Hamilton et al. 1997). In addition, preliminary evidence suggests that female lMAN volume may also correlate positively with male song repertoire size among several species of European warblers (DeVoogd et al. 1996).

2.3.3.4. Behavioral-State Modulation of Auditory Responses

A role for the AFP in song perception does not exclude a role for the VMP in song perception. Some controversy remains as to whether the VMP contributes to song perception. Throughout the song system, auditory responses of neurons are strongly modulated by behavioral state (Dave et al. 1998a; Schmidt and Konishi 1998). Surprisingly, the modulation of response is such that activity is suppressed in awake birds relative to sleeping birds. One report that systematically studied behavioral state changes found virtually no auditory activity in RA in awake birds but strong auditory responses in sleeping birds (Dave et al. 1998a). (Subsequent single-unit studies have suggested that the sleep state may be involved in song playback, perhaps associated with motor program consolidation; see Dave and Margoliash 2000.) Another study, which focused on the contribution of HVc and lMAN to RA auditory responses in anesthetized birds, reported (in a small part of the overall design) RA auditory responses in awake birds (Vicario and Yohay 1993). It is difficult to resolve these apparently conflicting observations. The Dave, Yu, and Margoliash (1998a) study used chronic recording techniques, and male birds were sufficiently unhindered that they freely sang in response to females. The auditory stimuli were presented after the females were removed from cages adjacent to those of the subjects. The Vicario and Yohay (1993) study used restrained animals in a sound-attenuation chamber and did not directly observe the animals during acoustic stimulation. It is difficult to assess whether these differences in design could result in sufficient differences in behavioral state to explain the variation in RA auditory responsiveness. A recent lesion study of RA also claimed to show an effect on conspecific song recognition (Vicario et al. 2001). The behavioral effects were small, however, and whether the locus of the effects of the lesions was within RA remains unresolved.

The failure to find auditory responses in the VMP in awake birds is a strong challenge to the theory that the VMP is involved in song perception (see Margoliash 1997). A related question is whether any song-system neurons exhibit auditory activity during the day. One study recording multi-unit activity identified an apparent loss of auditory activity in male zebra finch HVc during daytime recordings (Schmidt and Konishi 1998). In contrast, several other studies identified daytime auditory activity in HVc of canaries, white-crowned sparrows, and zebra finches (McCasland and Konishi 1981; Margoliash 1986; Yu and Margoliash 1996). A study of single HVc neurons concluded that many neurons retain auditory responsiveness during the day, albeit at consistently lower levels than observed at night. Consistent if weak auditory responses were also observed in area X in daytime recordings (Rauske and Margoliash 1999), and daytime auditory responses in nonsinging birds have also been reported for lMAN recordings (Hessler and Doupe 1999). This suggests that the AFP but not the VMP retains auditory activity during the day. If this conclusion is established, and these results are obtained in other species and under other behavioral measures of perceptual responses, they would provide strong evidence against the motor theory hypothesis of perception and strong evidence for the ontogenetic-coupling hypothesis.

2.3.4. Plasticity and Functional Constraints

The ethologically relevant functions of any representational system derive from a combination of an organism's evolutionary history and ontogenetic experience. Both of these processes can lead to plasticity in the neural system and thus can constrain the form of a representation at any given point in time. Plasticity can result from time-dependent components of ontogeny and from experience-dependent components (i.e., learning). The combination of these two components gives rise to the so-called "critical period" in development where certain experiences must occur at specific times in order for representations to develop normally (for discussions of critical periods, see Bateson 1979). Even after a normal sensory development, representational plasticity can be induced by altering experience in the periphery (e.g., by peripheral denervation, stimulus restriction, or illusion; Pons et al. 1991; Recanzone 1998) or by central lesions (Nudo et al. 1996). These effects have been observed in many sensory modalities, and although usually examined in relation to cortical representations (e.g., Jones 2000), they have also been observed in more peripheral structures (Gao and Suga 1998; Gold and Knudsen 2000). Moreover, representational plasticity is observed not only in response to abnormal experience. Simply training a monkey to perform a frequency-discrimination task increases the cortical representation (number of cells), sharpness of tuning, and response latency for the behaviorally relevant frequencies (Recanzone et al. 1993). Likewise, in humans and bats, tonotopic responses in the auditory cortex can be modulated by simple conditioning procedures (Morris et al. 1998; Gao and

Suga 2000). These plasticity effects appear to have important differences, depending on the specializations of particular regions (Sakai and Suga 2001). Based on such data, we hypothesize that the representations of complex acoustic events, such as vocal communication signals, are also mediated by experience. Thus, it may be that, at any one time, the representational system is "tuned" to a subset of conspecific vocalizations related to relevant tasks and contexts consistent with individual experience in a manner similar to that observed for synthetic stimuli (e.g., Kilgard and Merzenich 1998; Kilgard et al. 2001). Under such conditions, uninformed choices of natural vocalizations presented during an experiment may be incongruous with the representational state. One must be aware of the behavioral relevance associated with each vocalization and attempt to match the functional significance (i.e., the induced plasticity) of specific vocalizations to the experience of specific animals.

The behavioral relevance of different vocalizations is also shaped by evolutionary history, and this provides an additional source for constraints on the representational system. However, these effects are not always intuitively obvious. For example, vocal-communication based recognition can take many forms, depending on the species under consideration. One commonly observed form of recognition is between heterospecific and conspecific signals (i.e., species recognition). Among the midshipman, *Porichthys notatus*, a species of nocturnally active fish, for example, females appear to use male acoustic signals to localize prospective mates (McKibben and Bass 1998). The same is true for many anurans (Capranica 1965; see Feng and Ratnam 2000) as well as birds and mammals (Searcy and Yasukawa 1996). Based on such widespread observations, one general function of the perceptual system may be to differentiate between heterospecific and conspecific signals. However, in many cases, animals are able to make much finer discriminations between conspecific signals, and to the extent that behaviorally relevant classifications are made on the basis of intraspecific acoustic variation, species recognition can be expected as an indirect by-product of an auditory system tuned to other information. This leads to the hypothesis that *inter*species recognition is related to the structure of the ascending auditory system. It may not be necessary to postulate any special mechanisms or "templates" to account for innate predispositions for recognition of conspecific vocalizations (cf. Konishi 1978; Marler and Sherman 1983). Perhaps the most dramatic and compelling evidence for this comes from chick/quail chimera studies, where perceptual predispositions were associated with midbrain structures that could be transferred across species (Long et al. 2001). Because auditory feature detectors have mostly been described in forebrain pathways, for example in birds and primates, we propose the hypothesis that *intra*species recognition is related in these species to the structure of the forebrain auditory system. In Section 3, we develop intraspecies recognition as a model for perceptual and cognitive mechanisms of vocal communication.

2.3.3.1. Categorization and Classification of Conspecific Vocalizations

One way that experience-dependent plasticity can influence the structure and form of perceptual representations is through the emergence of categorical boundaries between various sets of stimuli. At the behavioral level, categorical perception is observed as a nonlinear relationship between stimulus variation along some dimension and a corresponding behavior. Changes along a stimulus dimension that span the boundaries between categories are easy to detect, whereas equal-magnitude changes within a category are more difficult to detect. Such behaviors suggest the obvious hypothesis that similarly categorical (i.e., nonlinear) neural responses should be observed when particular neurons are exposed to fixed levels of variation both within and across category boundaries. In general, the complexity of most natural stimuli suggests that categorical neural responses are likely to arise at higher levels within forebrain hierarchies. An example of this is the recent data showing differential responses of neurons in the prefrontal cortex of monkeys to different categories of visual objects (Freedman et al. 2001). However, in cases where categorical, perceptual behavior corresponds closely to relevant dimensions of peripheral tunings (e.g., frequency), categorical responses may first arise in peripheral structures.

Not all categories of stimuli will necessarily meet the behavioral definitions required for strict demonstrations of categorical perception. More general category formation can be described at varying levels of abstraction (Herrnstein 1990). For instance, some categories of stimuli may simply reflect rote collections of arbitrary objects, whereas exemplars in other categories may share common physical features or be the predicate of an abstract rule. For instance, classes of predators, food items, or potential mates may form natural categories. These general forms of categorization appear to be nearly ubiquitous among vertebrate species (Herrnstein 1990), and some forms may occur among invertebrates as well (e.g., Wyttenbach et al. 1996). Moreover, the fact that learning influences most categorical behavior suggests an amazingly high degree of plasticity in the neural processes underlying these capabilities.

Behavioral tests are required to determine the location of category boundaries within natural stimulus sets. Here, the use of vocalization repertoires holds a distinct advantage because much of the behavioral work on communication systems has been directed at the determination of behaviorally relevant classes of stimuli. The use of repertoires of natural stimuli has been developed most prominently in the analysis of bird and primate auditory systems, where neuronal response variation across different categories of vocalizations has been tested. For example, in European starlings, a species of songbird, neurons throughout the field L (analogous to the primary auditory cortex in mammals), the caudal medial neostriatum (NCM), and caudal ventral hyperstriatum (cHV) respond to conspecific vocalizations, and the specificity of the response properties shows increasing complexity between field L and NCM/cHV (Leppelsack and Vogt 1976; Müller and Leppelsack 1985; see Section 3). In general, neurons responsive to specific, behaviorally significant features of vocalizations form a plausible substrate for categorical representations (see Section 2.1). Very few studies, however, have explored the neuronal responses to natural variation within and between categories of stimuli (Freedman et al. 2001). This work is only now beginning in primates and birds (Gentner and Margoliash 2001; Tian et al. 2001; Wang and Kadia 2001). This is an area ripe for future research. In Section 3, we provide a logical basis for understanding these neural mechanisms in the context of vocal recognition.

3. Emerging Model Systems for the Neuroethology of Vocal Recognition

The preceding sections discuss a number of ways in which neuroethological approaches using vocal communication signals can constrain both the input to and state of the representational system as well as its output. The biological plausibility of these constraints derives directly from the functional (i.e., adaptive) role served by vocal communication signals. Thus, the extent to which these various constraints can be invoked for any one system depends on our knowledge of how that organism uses particular signals under natural conditions. Ideally, one wants to know how behaviors are driven by acoustic variation in a given communication signal and then adapt the critical components of such behaviors to laboratory procedures that are amenable to physiological preparations. This requires sophisticated knowledge of both behavior and basic auditory physiology, and to date only a few systems meet these requirements. In this section, we review the central auditory physiology and vocal-recognition behavior in two communication systems, those of songbirds and primates. We concentrate on vocal recognition because it captures many features of communication systems described above (e.g., categorization) and is widespread among many taxa, making the results amenable to comparative analyses.

3.1. Perception of Vocal Signals

Variation in communication signals can occur in the spectro-temporal properties of the signal itself and also in the spatial-temporal distribution of signal sources. Together, this variation leads to at least two general classes of receiver behavior. The first class derives from the fact that not all acoustic events are of equal interest, so animals must be able to dissociate appropriate target signals from irrelevant/background noise, including nontarget conspecific vocalizations. This has been studied in the context of the so-called cocktail party effect (Cherry and Taylor 1954) and more generally in terms of auditory stream segregation (Bregman 1990). Despite its obvious importance, only a relatively few studies have addressed this phenomenon in nonhumans (Fay 1998; MacDougall-Shackleton et al. 1998b). Only a single study has examined stream segregation using natural stimuli (Hulse et al. 1997). It may be that the acoustic parameters governing stream segregation of acoustic communication signals vary dramatically from those involved in the segregation of pure tone sequences. In any case, the basic ability is likely to be widespread, and initial data suggest that such processing occurs at or before the level of the primary auditory cortex (Fishman et al. 2001). Recent reviews cover both stream segregation and the closely related topic of auditory spatial localization in nonhuman animals (Feng and Ratnam 2000; Klump 2000), and we do not address them further. Most research using conspecific communication signals assumes that the test subject has successfully extracted the target signal by presenting stimuli in isolation. This also assumes independence of segregation and subsequent classification behaviors.

Once an auditory object is formed, a second general class of behavior emerges as these objects or events are organized into behaviorally relevant classes. For example, females might rely on male vocalizations to help choose a mate and therefore are likely, under appropriate conditions, to distinguish between heterospecific and conspecific male vocalizations. The presence of such distinctions, or class boundaries, implies discrimination among multiple auditory objects along with an associative link between the resulting internal representation and some behavioral response. At the behavioral level, these processes are collectively referred to as recognition.

According to this definition, recognition can take many forms, depending on the specific boundaries between classes of vocalizations. Often these acoustic boundaries correspond to other behaviorally relevant distinctions (e.g., species, sex, kin, and individual). That is, they are not arbitrary but rather reflect the ecology of the particular animal under consideration. Although not all forms of recognition behavior are likely to be mediated by the same neural mechanisms, there are likely to be shared features across species, particularly when relevant classification requires discrimination among subsets of conspecific vocalizations. Recognition based on intraspecific acoustic variation is widespread (Boughman and Moss, Chapter 4), and several of the most recent examples from different taxa are given below.

3.2. *Intraspecific Recognition Behavior*

One common distinction within species is in the degree of relatedness between individuals. Vocal recognition often follows these lines. For example, king penguin chicks, *Aptenodytes patagonicus*, and emperor penguin chicks, *A. forsteriis*, can each recognize the calls of their parents (Jouventin et al. 1999; Aubin et al. 2000), in the latter case by using harmonic interference patterns generated by the simultaneous production of two sounds in different halves of the syrinx (the vocal-production organ in birds). Adult king penguins can also recognize their mate's calls (Lengagne et al. 2000), as can spectacled parrotlets, *Forpus conspicillatus* (Wanker et al. 1998), and several species of songbirds (Lind et al. 1997; O'Loghlen and Beecher 1997, 1999; Beguin et al. 1998).

Vocalization-based kin recognition is also apparent among a variety of mammalian species, including some populations of gray seals, *Halichoerus grypus* (McCulloch and Boness 2000), and Northern fur seals, *Callorhinus ursinus*, where mother-offspring recognition is maintained for many years beyond the breeding season (Insley 2000). In addition, bottlenose dolphins, *Tursiops truncatus*, have individually distinctive "signature" whistles long thought to function in recognition. Recent data confirm this by showing that mothers can recognize the whistles of their independent offspring and that independent offspring can recognize the whistles of their mother (Sayigh et al. 1999).

Female African elephants, *Loxodonta africana*, appear to have extensive networks of vocal recognition, distinguishing the infrasonic calls of female family, bond group, and even more distant kin from those of females outside these categories (McComb et al. 2000). Female spotted hyenas, *Crocuta crocuta*, can also recognize specific vocalizations of their own pups (Holekamp et al. 1999). Female greater spear-nosed bats, *Phyllostomus hastatus*, give screech calls whose acoustic structure varies between groups from different caves, and individuals appear to discriminate among the calls from different caves, although the capacity for individual vocal recognition remains unknown (Boughman and Wilkinson 1998).

3.3. The Songbird Model

Various forms of intraspecific vocal recognition have been observed in nearly every species of songbird studied to date (see Stoddard 1996) and have been examined more extensively here than in any other group of animals. In general, vocal recognition in songbirds provides for the association of specific songs with specific singers or locations and thereby serves as a basis for decisions in more elaborate social behaviors such as female choice (Wiley et al. 1991; Lind et al. 1997), female preference (O'Loghlen and Beecher 1997), and kin recognition among communally breeding birds (reviewed by Beecher 1991). Another complex social behavior in which vocal recognition plays an important role is territoriality, where it functions in both the manipulation and maintenance of territorial boundaries (Peek 1972; Falls and Brooks 1975; Falls 1982; Godard 1991) and thus may have indirect effects on reproductive success (Hiebert et al. 1989).

3.3.1. Field Studies

The function of male song in maintaining and establishing songbird territories is well-known. For example, removing a male songbird's ability to sing has dramatic effects on his success at holding a territory (Peek 1972, Smith and Reid 1979; McDonald 1989), and simply broadcasting a conspecific song from an unoccupied territory leads to significantly lower rates of settlement in that territory than in others' territories where control sounds or no sounds are played (Krebs et al. 1978; Yasukawa 1981; Falls 1987). Moreover, territory residents often respond weakly, or not at all, to a neighbor singing from a familiar location but more strongly to a stranger singing from that same location. Using these facts along with a variety of clever song-playback techniques in the field, a very large number of studies have demonstrated that males in many (at least 23) songbird species are capable of discriminating among neighbors and strangers on the basis of song alone. Furthermore, for several (at least 8) species, listeners are capable of recognizing individual singers on the basis of their songs (see Stoddard 1996).

3.3.2. Signal Variation

There are several ways that singer identity could be represented in the acoustic variation of male song. In the simplest case, individual males might sing a unique song or sets of songs (i.e., repertoires), and recognition would follow by the association of specific songs with specific singers. This strategy appears to be used by song sparrows, *Melospiza melodia* (Beecher et al. 1994), and European starlings (Gentner et al. 2000). One obvious feature of this strategy, potentially worth exploring, is that it may be heavily constrained by the memory capacity of the recognition system. Although this question deserves further attention, initial results suggest that the capacity of these systems is, in fact, quite high (see Gentner et al. 2000) and, at least for song sparrows, exceeds the number of exemplars that an individual is likely to face at a single time in the wild (Stoddard et al. 1992).

A second recognition strategy employed relies on morphological differences in the acoustics of shared song types. In both white-throated sparrows and field sparrows, the songs of neighboring territorial males share several acoustic features but vary slightly in frequency. Neighbors rely on these subtle frequency differences to recognize one another (Brooks and Falls 1975; Nelson 1989). Related to this is a third possible strategy for recognition. If the morphology of the vocal-production apparatus varies slightly between individuals, then this variation might impart unique spectral features, or so-called "voice characteristics," to all of an individual's vocalizations. The use of voice characteristics has been suggested for great tits, *Parus major* (Weary and Krebs 1992), but does not appear to be a relevant cue for either song sparrows (Beecher et al. 1994) or starlings (Gentner et al. 2000). Finally, vocal recognition might also rely on the sequence in which multiple song types are sung. That is, different males may share song types but sing them in individually distinctive temporal patterns. The role of this final cue has not been extensively studied in songbirds, but there is some evidence to suggest that European starlings are sensitive to the sequence of motifs within familiar song bouts (Gentner and Hulse 1998).

For species in which males sing multiple songs, the four mechanisms outlined above may not be mutually exclusive. There is no *a priori* reason to believe that vocal recognition in a single species relies on individual variation coded in only a single dimension, nor is there any reason to suspect that all species of songbirds use the same recognition strategies. Given the approximately 4,500 different species of songbirds—each singing acoustically distinct songs and the occurrence of vocal recognition in a wide range of behavioral contexts, it is likely that vocal-recognition information is coded at multiple levels throughout a songbird's repertoire (Braaten 2000).

3.3.3. Laboratory Studies

Given the likely diversity of vocal-recognition behaviors across songbird species, it is reasonable to consider whether there are corresponding peripheral perceptual specializations among songbirds that in theory might provide an "open channel" of communication within a species while limiting confusion across species. For instance, different species might concentrate the spectral energy with their songs in defined spectral bands. This hypothesis is supported by several observations of species-specific advantages during operant discriminations of multiple conspecific and heterospecific songs in several different species (Sinnott 1980; Okanoya and Dooling 1990; Cynx and Nottebohm 1992; Dooling et al. 1992). However, the overwhelming data from psychophysical studies of hearing in birds indicate that most basic sensory processing capabilities (e.g., frequency sensitivity) are conserved across songbird species (Dooling et al. 2000). Thus, it appears that biases for the discrimination of species-specific vocalizations, and hence mechanisms for vocal recognition, result from evolutionary or ontogenic changes in central processing structures. This inference is consistent with the more general assumption that the cognitive processes underlying vocal recognition take the neural representation of acoustically complex signals (i.e., song) as their input. Recent laboratory studies of European starlings have addressed these questions by determining more precisely the form of the acoustic signal controlling recognition in this species.

Male starlings tend to present their songs in long episodes of continuous singing referred to as bouts. Song bouts, in turn, are composed of much smaller acoustic units referred to as motifs (Adret-Hausberger and Jenkins 1988; Eens et al. 1991), which in turn are composed of still smaller units called notes. (This usage of "motif" for starling songs is at slight variance with motifs as defined for zebra finch songs in Section 2.3.3.1.) Notes can be broadly classified by the presence of continuous energy in their spectrographic representations, and although several notes may occur in a given motif, their pattern is usually highly stereotyped between successive renditions of the same motif. One can thus consider starling song as a sequence of motifs, where each motif is an acoustically complex event. The number of unique motifs that a male starling can sing (i.e., his repertoire size) can be quite large, and consequently different song bouts from the same male are not necessarily composed of the same set of motifs. This broad acoustical variation in their song provides several potential cues that starlings might use when learning to recognize the songs of an individual conspecific and while maintaining that recognition over time. One straightforward recognition mechanism is the association of specific motifs with specific singers. Although some sharing of motifs does occur among captive males (Hausberger and Cousillas 1995; Hausberger 1997), the motif repertoires of different males living in the wild are generally unique (Adret-Hausberger and Jenkins 1988; Eens et al. 1989, 1991; Chaiken et al. 1993; Gentner and Hulse 1998). Thus, learning which males sing which motifs can provide a discriminative cue for song classification.

Data from operant studies in starlings support the idea that recognition is based at the level of the motif. Starlings trained to recognize individual conspecifics by one set of song bouts can readily generalize correct recognition to novel song bouts from the same singers (Gentner and Hulse 1998; Gentner et al. 2000; Fig. 7.4A, B). However, when these novel song bouts have no motifs in common with the training songs, and when song exposure outside of the operant apparatus is restricted, recognition falls to chance (Gentner et al. 2000; Fig. 7.4C). Likewise, starlings trained to discriminate among pairs of motifs will reverse the discrimination when transferred to the same motif sung by the opposite individual and perform at chance when transferred to novel motifs sung by the training singers (Gentner 1999). This failure to generalize correct recognition to songs composed of novel motifs, or to single novel motifs, is inconsistent with the use of individually invariant source and/or filter properties (voice characteristics) for vocal recognition.

The data suggest that starlings learn to recognize the songs of individual conspecifics by attending to information contained at (or below) the level of the motif and by then associating distinct sets of motifs with individual singers. If this is true, then once recognition is learned, it should be possible to control it systematically by varying the proportions of motifs in a given bout that come from two "vocally familiar" males. That is, recognition behavior ought to track the statistical distribution of motifs from two vocally familiar males rather than the presence or absence of single motifs from either male. Recent data confirm this prediction (Gentner and Hulse 2000; Fig. 7.4) and thereby suggest that when starlings are compelled to

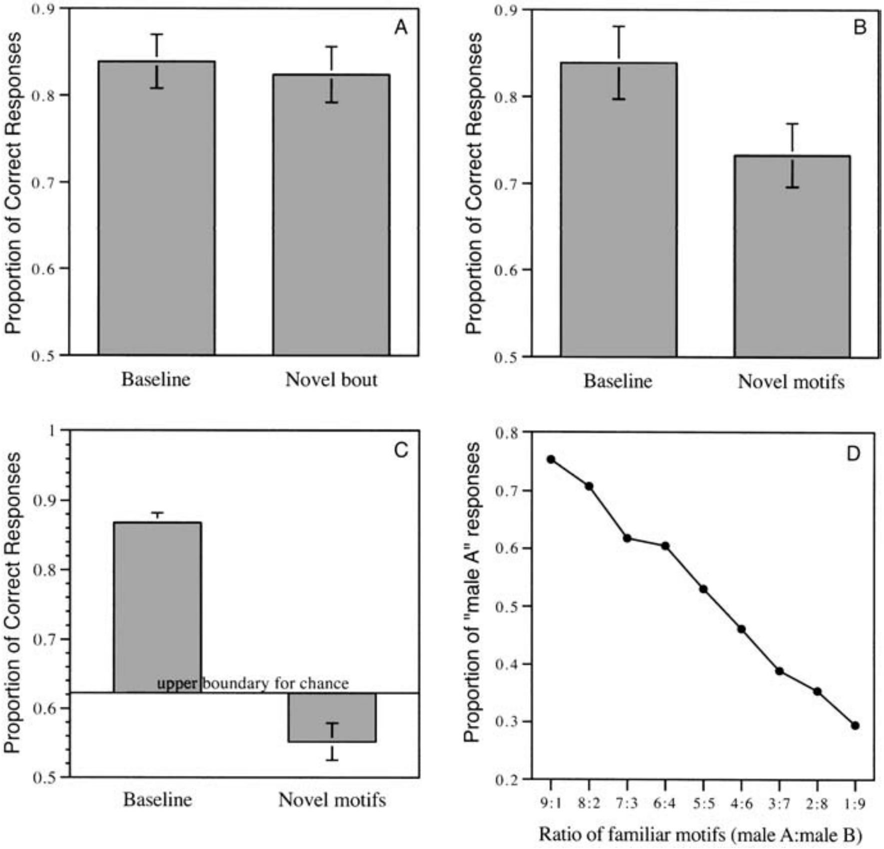

FIGURE 7.4. Vocal-recognition behavior in European starlings. (**A**) Mean (±SEM) proportion of correct responses given during asymptotic performance on an operant recognition task ("Baseline") and during initial transfer to novel songs containing familiar motifs ("Novel bout"). (**B** and **C**) Mean (±SEM) proportion of correct responses during transfer from the baseline training to novel songs from the same singers composed of "Novel motifs." Data in (**B**) show the transfer when the subjects were exposed to the training and test songs outside of the operant apparatus, whereas the data in (**C**) show the results for the same transfer after controlling for this experience. Note that in (**C**) recognition of the novel motifs falls below chance. (**D**) Data showing the close (and approximately linear) relationship between the statistical distribution of familiar motifs from two different singers and song recognition.

classify conspecific songs, they do so by memorizing large numbers of unique song components (i.e., motifs) and then organizing subsets of these motifs into separate classes. As a cognitive strategy, classifying songs according to their component structure represents a parsimonious method of dealing with these complex acoustic signals. Because individual starlings tend to possess unique motif repertoires, disjoint sets of motifs will generally correspond to individual identity. Therefore, attending to the motif structure captures a significant portion of the individual variation in the signal.

The behavioral data inform a number of hypotheses regarding the neural mechanisms of vocal recognition in starlings. First, the representational architecture of forebrain auditory regions should reflect the segmentation of song at the level of the motif. When similarity/difference computations are invoked, as they must be for recognition to proceed, the input to such computations should be some representation of the motif. Determination of the precise acoustic information corresponding to such a representation has yet to be accomplished but is amenable to behavioral study in the laboratory. Investigations of these representations at the neural level are likely to tell a great deal about the manner in which complex auditory objects are represented and processed. Second, the behavioral strategy employed by starlings for vocal recognition suggests that the upper bound on the capacity for accurate recognition is constrained by the memory capacity of the system for specific motifs. If the representation of a motif is coded by dynamic temporal and spatial patterns of connectivity among local feature detectors, then these memory constraints may derive directly from the perceptual mechanisms for coding complex stimuli.

3.3.4. Songbird Auditory System: Anatomy

The basic plan of the passerine auditory system follows a general reptile-bird pattern of connections (Ulinski and Margoliash 1990; Carr 1992; Carr and Code 2000). In birds, the auditory nerve projects to two cochlear nuclei, the nucleus magnocellularis and the nucleus angularis. These nuclei project in turn to second-order olivary nuclei, to the lemniscal nuclei, and contralaterally to the central nucleus of the nucleus mesencephalicus lateralis dorsalis (MLd), the avian analog of the inferior colliculus. The central nucleus of the MLd projects to a major target in the medial portion of the dorsal thalamus, the nucleus ovoidalis (Ov), and to a lesser extent to the subjacent nucleus semiluminaris parovoidalis (SPO) (Karten 1968) and a region surrounding the Ov-SPO complex referred to as the "shell" (Durand et al. 1992). There are a wealth of connections and patterns of connections in the more peripheral auditory structures not reviewed here. In addition to the primary projections to the auditory telencephalon, auditory fibers also project from the thalamus to the hypothalamus, providing one possible route for interaction between auditory, neuromodulatory (Li and Sakaguchi 1997), and endocrine systems (Durand et al. 1992).

Nissl and Golgi preparations of male zebra finch brains (Bonke et al. 1979b) confirm the general pattern of organization observed in other species (Karten 1968) and demonstrate that the caudal medial portion of the avian telencephalon is composed of five cytoarchitectonic subregions—L1, L2a, L2b, L3 and L—called the field L complex (Fortune and

Margoliash 1992). The field L complex is the primary telencephalic target for auditory information arriving via several parallel pathways from the Ov complex in the thalamus (Fig. 7.3). Neurons in the shell of Ov also project to the caudomedial portion of the neostriatum (NCM) (Durand et al. 1992; Vates et al. 1996). The subregions of field L are densely interconnected and project to the NCM and reciprocally to the lateral portion of the caudal ventral hyperstriatum (clHV). The NCM and clHV share reciprocal connections with the caudal medial portion of HV (cmHV). Figure 7.3 provides a schematic for this complicated pattern in songbirds.

Neurons in clHV, L1, and L3 also project to the neostriatum immediately ventral to HVc, referred to as the "shelf" (Kelley and Nottebohm 1979). Neurons in the shelf are auditory (Scheich et al. 1979a; Müller and Leppelsack 1985), and it has been proposed that the shelf is a source of auditory input to HVc. The projections of the shelf into HVc are extremely sparse, however (L. Katz, cited in Margoliash 1987; Vates et al. 1996). One proposition is that dendrites in caudal HVc, which extend into the shelf, receive auditory input from shelf axons. These hypotheses are longstanding but have never been confirmed, in part because the small size of the structures and their physical proximity to each other hinders independent manipulation of the shelf and HVc. Auditory information may also enter the VMP via the clHV-NIf-HVc pathway (Fig. 7.3). BOS selective auditory responses and correlated activity in NIf have recently been demonstrated (Janata and Margoliash 1999). NIf activity may be necessary for HVc auditory responses (Boco and Margoliash 2001). NIf receives input from clHV and possibly also from field L via the dendritic projections of a distinct class of NIf neurons lying along the dorsorostral border of NIf with L1 (Fortune and Margoliash 1995). Identification of sources of auditory input into the VMP is of particular interest because it provides a neural substrate for hypotheses about the origin of the well-known BOS selective responses in these structures. The auditory inputs to HVc also provide the basis for speculations that expand the functions of AFP structures beyond their classic role in juvenile song learning and into adult song perception (see Section 2.3.3).

The song system is associated with a series of cytoarchitectonically indistinct structures. These structures, such as the shelf ventral to HVc and medial MAN, are physically close to cytoarchitectonically distinct song-system nuclei and have patterns of connections similar to song-system nuclei. Components of this indistinct system have also been observed in several species of nonpasserine birds (Brauth et al. 1987; Brauth and McHale 1988; Korzeniewska and Güntürkün 1990; Fortune and Margoliash 1995; see also Margoliash et al. 1994). Although the functions of this pathway remain unknown, one possibility is that the song system arose as a series of specializations of the indistinct pathway in relation to selection pressures associated with vocal learning. If this is true, then the indistinct pathway in oscines and other birds may mediate more general aspects of vocal behavior, such as calling behavior in relation to complex social interactions (Margoliash et al. 1994), or other reward/reinforcement contingencies involving vocal perception in adults. Such speculations are consistent with emerging homologies between the AFP and cortico-basal ganglia loops in mammals (Luo and Perkel 1999).

3.3.5. Songbird Auditory System: Physiology

In European starlings, Ov is tonotopically organized, with best frequencies decreasing ventrally (Bigalke-Kunz et al. 1987). Neurons throughout the auditory telencephalon also show tonotopic organization (Leppelsack and Schwartzkopff 1972; Rubsamen and Dorrscheidt 1986), although in much more complex patterns. In starlings, roughly 11 different regions can be identified on the basis of the direction of the tonotopic gradient and tuning curve bandwidth (Haüsler 1997; Capsius and Lepplesack 1999), and similar patterns are observed in zebra finches (Gehr et al. 1999). These tonotopically defined regions appear to respect anatomically defined regions of the field L complex.

Relatively few studies have examined responses in the telencephalic regions using complex acoustic stimuli. Neurons in L1 and L3 have lower response rates to tone bursts than those in L2 and show greater selectivity to species-specific vocalizations (Leppelsack and Vogt 1976; Bonke et al. 1979a; Müller and Leppelsack 1985; Theunissen and Doupe 1998). This selectivity is borne out by the complexity of the spatial-temporal receptive fields (STRFs) for many neurons within field L. Indeed, more reliable estimates of the STRF are derived from responses to conspecific vocalizations than tone pips (Theunissen et al. 2000; cf. Schäfer et al. 1992). This general pattern of increasing response selectivity from field L2 to the higher-order areas continues into NCM and cHV (Müller and Leppelsack 1985), suggesting that these regions are involved in the extraction of complex features. Early data from white-crowned sparrows are consistent with this in showing a small subset of neurons in the NCM that are selective for specific directions of FM in a common trill element of conspecific song (Leppelsack 1983). Recent preliminary data (Grace and Theunissen 2000; Gentner and Margoliash 2001) support the idea that cHV in particular is involved in the extraction and/or representation of complex features in showing highly selective responses in this region to specific features in behaviorally relevant conspecific songs.

Neurons in NCM are broadly responsive to conspecific stimuli and respond to the repeated presentation of conspecific song in a stimulus-specific manner (Chew et al. 1995; Stripling et al. 1997). The repeated presentation of a single conspecific song elicits a rapid modulation in the initial firing rate of NCM neurons (Stripling et al. 1997). If the same song is repeated on the order of 200 times, this initial modulation of the firing rate is no longer observed when that same song is presented on subsequent trials.

This is true even though the initial response modulation can still be observed for other conspecific songs (Chew et al. 1995, 1996; Stripling et al. 1997). These stimulus-specific changes in the response properties of NCM neurons have led to the hypothesis that NCM may play an important role in individual vocal recognition (Chew et al. 1996). Consistent with this idea is the fact that many neurons in NCM (and cHV) show a rapid up-regulation of the immediate early gene (IEG) *zenk* in response to the presentation of conspecific songs (Mello et al. 1992) that is tuned to the acoustics of particular conspecific song syllables (Ribiero et al. 1998). Interestingly, the genomic response also habituates to the repeated presentation of the same conspecific song (Mello et al. 1995) and is elevated during specific components of the vocal-recognition task described above in starlings (Gentner et al. 1999). The mammalian homolog to *zenk* is required for expression of late LTP and long-term memories in mice (Jones et al. 2001). This suggests that *zenk* expression in NCM and cHV may be related to learning about conspecific songs and implicates these structures in concomitant processes.

3.4. The Nonhuman Primate Model

3.4.1. Vocal-Recognition Behavior

Various forms of vocal recognition are also prominent among many species of primates. In humans, *Homo sapiens*, the ability to recognize the sex and the identity of a talker is anecdotally obvious. Acoustic cues to a talker's sex are present in both the fundamental frequency (reflecting larynx size) and vocal tract length (reflecting body size). Recent data on speech perception suggest that acoustic cues to individual recognition within sexes are due to supralaryngeal vocal tract filtering caused by anatomical variation between talkers (Bachorowski and Owren 1999). Individual vocal recognition (and kin recognition) has also been demonstrated in rhesus macaques, *Macaca mulatta* (Rendall et al. 1996), and acoustic cues related to vocal tract filtering have been suggested as the basis for this behavior (Owren et al. 1997; Rendall et al. 1998). Vocal recognition is also apparent in the referential call systems of many primates (e.g., Seyfarth et al. 1980a, 1980b; Hauser 1998; Rendall et al. 1999; Fitch and Hauser, Chapter 3). In these cases, rather than having signals that are associated with specific individuals or groups of individuals, particular calls are used to refer, for example, to different classes of predators, food types, or various other behavioral events. This aspect of primate vocal communication has been reviewed elsewhere (Cheney and Seyfarth 1990; Ghazanfar and Hauser 1999), and we do not address it further except to point out that such behaviors provide another set of natural categories for vocalizations that is amenable to future neuroethological study.

A third class of vocal-recognition behavior has been studied extensively in Japanese macaques, following the observation that several subtypes of the "coo" vocalization in this species can be defined on the basis of acoustic variation (Green 1975). Much of this work has focused on two particular subtypes of coo vocalizations, the smooth-early-high (SEH) and smooth-late-high (SLH), so-called because of the relative position of the peak of one frequency component that sweeps up and then down over the course of the call. Although Japanese macaque mothers can discriminate the coos of their young from others (Pereira 1986), the role of individual recognition cues in the coo vocalizations has not been well-studied. Nevertheless, the SEH and SLH call types function in different behavioral contexts (Green 1975), and Japanese macaques possess a species-specific bias for discriminations involving these coos when the relevant variation is in the relative timing of the FM peak (Zoloth et al. 1979).

Although initial data suggested that the SEH and SLH calls are perceived categorically (May et al. 1989), more recent data from the field indicate that many adult female coo vocalizations have FM peaks within the ambiguous zone between "early" and "late" prototypes (Owren and Casale 1994). The categorical boundaries determined in laboratory tests do not coincide with natural variation in the distribution of calls in the field. Nonetheless, all data to date show clear evidence that the coo calls are perceived as perceptually distinct classes (if not categorically), and several studies now substantiate the notion that the relative position of the FM peak within the coo is the most salient cue to discrimination among different coos (May et al. 1988; Le Prell and Moody 2000). Amplitude cues also appear to function in call discrimination (Le Prell and Moody 1997). Thus, the stimulus dimensions involved in "real-world" classification and/or categorization of coo calls may involve a more complex stimulus space than that suggested by the original categorical perception studies.

Interestingly, lesions to the left, but not the right, superior temporal gyrus impair discrimination between the SEH and SLH coos by Japanese macaques, suggesting a human language-like hemispheric dominance (Heffner and Heffner 1984). Similarly, behavioral tests among free-ranging rhesus macaques, *Macaca mulatta*, indicate a left hemisphere dominance for the processing of conspecific compared to heterospecific vocalizations that is present in adults but not infants (Hauser and Andersson 1994). Apart from their comparative value in understanding the evolutionary origin of human cognition, the lateralization of specific cognitive functions can in theory allow for identification of at least some of the brain regions mediating these behaviors. Comparisons of activation in each hemisphere during tasks that call on lateralized functions, using either fMRI or immediate early gene-expression techniques, should yield regions of differential activity in the appropriate hemisphere worthy of closer study.

3.4.2. Primate Auditory System: Cortical Anatomy and Physiology

The mammalian auditory cortex extends across the superior temporal plane and onto the adjacent superior temporal gyrus. The primary auditory cortex in primates can be divided into separate core and belt areas on the basis of

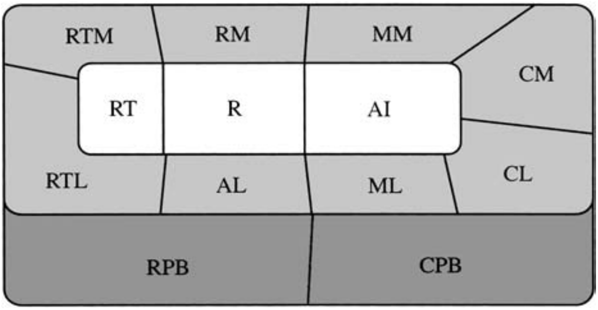

FIGURE 7.5. Schematic layout of primate auditory cortex. Areas shown in white represent the core regions of the auditory cortex, those shown in light gray the lateral belt, and those in dark gray the para-belt. Most of the areas in the core and lateral-belt area are interconnected. AI: primary auditory cortex; R: rostral area; RT: rostrotemporal area; CL: caudolateral area; CM: caudomedial area; MM: middle medial area; RM: rostromedial area; RTM: medial rostrotemporal area; RTL: lateral rostrotemporal area; AL: anterolateral area; ML: middle lateral area; RPB: rostral para-belt; CPB: caudal para-belt. (Based on Kaas and Hackett 2000).

cytoarchitecture and connectivity, with a third para-belt region positioned laterally (Fig. 7.5). Moving rostral to caudal, the core is composed of three distinct regions: the rostral temporal field RT, the rostral field R, and most caudally, AI. Both AI and R have well-defined cochlear representations, with caudorostral tonotopic gradients of best frequencies running from high to low in AI and low to high in R. The tonotopic organization of RT is less well-studied, but it appears that best frequencies are arranged from high to low as one moves rostrally (reviewed in Kaas and Hackett 2000). All three of these core areas are densely interconnected ipsilaterally and project colosally to homotopic regions in the opposite hemisphere. All three regions receive dense inputs from the medial geniculate complex, the principal nucleus of the thalamus. In the common marmoset, *Calithrix jacchus*, subpopulations of AI cells show selective responses to conspecific vocalizations as compared with synthetic variations with the same spectral but different temporal characteristics (Wang et al. 1995). This suggests that, at least in marmosets, AI cells may be sensitive to more than narrow-band frequency parameters.

Immediately surrounding the core is an area containing approximately four to eight auditory regions (Rauschecker 1998; Kaas and Hackett 2000), each with distinct cochlear representations, referred to as the belt (Fig. 7.5). Most of the thalamic input into these regions arises in the dorsal and medial division of the medial geniculate complex. It has been argued that the distribution of these thalamic projections and those to the core areas reflects the functional separation of spatial-localization and pattern-recognition mechanisms (Romanski et al. 1999; Rauschecker and Tian 2000) analogous to the so-called "what" and "where" pathways in visual processing. Although neurons in belt regions can be driven using tonal stimuli, and there are cochleotopic representations in several regions, many neurons seem to prefer specific bandwidths for frequency-centered sound bursts, independent of intensity (Rauschecker et al. 1995). The bandwidth tuning of cells in the caudolateral (CL), mediolateral (ML), and anterior later (AL) belt regions varies along an axis that is orthogonal to the cochleotopic organization.

An additional auditory processing region just lateral to the belt, and referred to as the "para-belt", has been identified histologically (Hackett et al. 1998), but neither its anatomical extent nor physiological response properties have been fully examined. In contrast to the dense reciprocal connections among the core and belt areas, the para-belt does not have direct access to information in the core but rather appears to be driven primarily by input from the belt (Kaas and Hackett 2000). The nature of this input is not fully resolved. However, both the general pattern of connections among the core, belt, and para-belt regions in primates and the change in tuning properties across these regions resemble that among field L, NCM, and cmHV in songbirds described above. Lateral-belt neurons are tuned for both direction and rate of frequency modulation (Rauschecker 1997), and, at least in macaques, many of these cells also respond vigorously to conspecific calls or components thereof (Rauschecker et al. 1995). Although responses to species-specific calls can also be observed in AI, the selectivity is significantly less pronounced than that in the lateral-belt areas (see Rauschecker 1998). Moreover, early recordings in squirrel monkeys, *Saimiri sciureus*, showing large percentages of cells responsive to conspecific calls and subpopulations therein with modest nonlinear response properties (Wollberg and Newman 1972; Newman and Wollberg 1973; Winter and Funkenstein 1973) were mostly done in the superior temporal gyrus and therefore were likely to lie within the lateral belt or para-belt. Thus, there appears to be a general pattern of increasingly complex receptive field properties that coincides with the transition from primary auditory cortex to postsynaptic regions. This can be observed in humans as well, where the dorsolateral superior temporal gyrus and planum temporale (putative belt) are more strongly activated by FM tones than noise, and the superior temporal sulcus (putative para-belt) is more activated by various speech stimuli, including pseudowords and reversed speech, than by FM tones and other nonvocal sounds (Binder et al. 1994, 2000; Belin et al. 2000).

4. Summary and Conclusions

We have discussed organizational schemes and mechanisms that underlie the coding of complex stimuli, including feature detectors and combination sensitivity, which are likely to provide the basis for representations of acoustically complex vocal communication signals. Although we have attempted to maintain a focus on perceptual mechanism, the nature of vocal communication systems necessitates an integrative approach to production and perception. Our discussion of motor-sensory linkages in the well-studied oscine song system addresses these concerns from both empirical and theoretical perspectives.

Much of the auditory cortex in mammals and auditory forebrain in birds is mapped tonotopically, but a number of additional acoustic stimulus parameter mappings have been described in birds and mammals, including pitch periodicity, frequency, intensity, spatial location, duration, amplitude, and frequency modulation (see Ehret 1997; Schulze and Langner 1997). The structure of these mappings may change significantly in awake, alert animals (Evans and Whitfield 1964; Pfingst et al. 1977; Dave et al. 1998a; Schmidt and Konishi 1998; Capsius and Leppelsack 1999; cf. Recanzone et al. 2000). Patterns of organization in numerous other systems—including barn owls, bats, electric fish, songbirds, and others—suggest that a breakdown in cochleotopic organization of responses often coincides with mappings for more complex stimulus parameters. This may represent the convergence of multiple acoustically simpler maps onto some regions (i.e., a stimulus reconstruction) or the extraction of encoded information along dimensions that have not been represented in prior mappings (i.e., emergent properties). The neuroethological approach provides a logical and theoretically sound basis for investigation of such emergent properties in complex signals.

The neuroethological approach to vocal communication is also informative of the likely functional outputs of a perceptual system. These outputs form the basis for higher cognition. We have described how vocal-recognition behavior, and more explicitly the organization of vocal communication signals into behaviorally relevant classes or categories, can be used to develop hypotheses about the perceptual and cognitive demands placed on the central nervous system.

Principles in neuroethology have been difficult to identify. Most reviews of neuroethological research are organized around a series of case examples. This can incorrectly reinforce the conclusion that principles of neuroethological research are not forthcoming. In contrast, for example, systems neuroscience has identified a number of features of CNS organization, such as mappings, lateral interactions, population dynamics, and network reconfiguration, which help to organize research and can be considered principles. Neuroethological principles do exist, such as those described in this paper, but they have been difficult to demonstrate because of the scope of work required for a comparative analysis of brain and behavior. Only now, after some 40 years of research, and only in the most extensively studied behaviors, such as vocal communication, are these principles emerging. Somewhat akin to neurobiology, neuroethology aims to explain how molecular, cellular, and systems-level phenomena result in behaviors such as perception and cognition. But by the very nature of its underlying definitions, neuroethology aims to do more, for it provides the nexus for integration of the multiple levels of analysis articulated by Tinbergen's classic four questions. Ultimate and proximate explanations are incomplete until they are united. The task for neuroethology is enormous, but the principles and patterns of organization we have described for vocal communication are examples of how successful this approach can be and how much promise future research holds.

Acknowledgments. We thank Andrea Megela-Simmons for valuable comments on the manuscript. This work was supported by grants from the NIH to DM (MH59831 and MH60276) and TQG (DC00389).

References

Adret P (1993) Operant conditioning, song learning and imprinting to taped song in the zebra finch. Anim Behav 46:149–159.

Adret-Hausberger M, Jenkins PF (1988) Complex organization of the warbling song in starlings. Behaviour 107:138–156.

Alexander RD (1962) Evolutionary change in cricket acoustical communication. Evolution 16:443–467.

Anderson SE, Dave AS, Margoliash D (1996) Template-based automatic recognition of birdsong syllables from continuous recordings. J Acoust Soc Am 100: 1209–1219.

Aubin T, Jouventin P, Hildebrand C (2000) Penguins use the two-voice system to recognize each other. Proc R Soc Lond B Biol Sci 267:1081–1087.

Bachorowski JA, Owren MJ (1999) Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. J Acoust Soc Am 106:1054–1063.

Bankes SC, Margoliash D (1993) Parametric modeling of the temporal dynamics of neuronal responses using connectionist architectures. J Neurophysiol 69:980–991.

Barlow HB (1972) Single units and sensation: A neuron doctrine for perceptual psychology? Perception 1:371–394.

Barlow HB (1995) The neuron doctrine in perception. In: Gazzaniga M (ed) The Cognitive Neurosciences. Cambridge, MA: MIT Press, pp. 415–435.

Bateson P (1979) How do sensitive periods arise and what are they for? Anim Behav 27:470–486.

Beecher MD (1991) Success and failures of parent–offspring recognition in animals. In: Hepper PG (ed) Kin Recognition. Cambridge, UK: Cambridge University Press.

Beecher MD, Campbell SE, Burt J (1994) Song perception in the song sparrow: Birds classify by song type but not by singer. Anim Behav 47:1343–1351.

Beer C (1982) Conceptual issues in the study of communication. In: Kroodsma DE, Miller EH (eds) Acoustic Communication in Birds, Vol 2. Song Learning and Its Consequences. New York: Academic Press, pp. 279–310.

Beguin N, LeBoucher G, Kreutzer M (1998) Sexual preferences for mate song in female canaries (*Serinus canaria*). Behaviour 135:1185–1196.

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. Nature 403:309–312.

Bentley DR, Hoy RR (1972) Genetic control of the neuronal network generating cricket song patterns. Anim Behav 20:478–492.

Best CT, McRoberts GW, Sithole NM (1988) Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. J Exp Psychol Hum Percept Perform 14:345–360.

Bigalke-Kunz B, Rubsamen R, Dorrscheidt G (1987) Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. J Comp Physiol A 161:255–265.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 10:512–528.

Binder JR, Rao SM, Hammeke TA, Yetkin FZ, Jesmanowicz A, Bandettini PA, Wong EC, Estkowski LD, Goldstein MD, Haughton VM, Hyde JS (1994) Functional magnetic resonance imaging of human auditory cortex. Ann Neurol 35: 662–672.

Boco T, Margoliash D (2001) NIf is a major source of auditory and spontaneous drive to HVc. Soc Neurol Abstr 27:381.2

Bodnar DA, Bass AH (1997) Temporal coding of concurrent acoustic signals in auditory midbrain. J Neurosci 17:7553–7564.

Bodnar DA, Bass AH (1999) Midbrain combinatorial code for temporal and spectral information in concurrent acoustic signals. J Neurophysiol 81:552–563.

Bonke D, Scheich H, Langner G (1979a) 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. J Comp Physiol 132:243–255.

Bonke DA, Bonke D, Scheich H (1979b) Connectivity of the auditory forebrain nuclei in the guinea fowl (*Numida meleagris*). Cell Tissue Res 200:101–121.

Bottjer SW, Johnson F (1997) Circuits, hormones, and learning: Vocal behavior in songbirds. J Neurobiol 33:602–618.

Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science 224:901–903.

Boughman JW, Wilkinson GS (1998) Greater spear-nosed bats discriminate group mates by vocalizations. Anim Behav 55:1717–1732.

Braaten RF (2000) Multiple levels of representation of song by European starlings (*Sturnus vulgaris*): Open-ended categorization of starling song types and differential forgetting of song categories and exemplars. J Comp Psychol 114:61–72.

Bradbury JW, Vehrencamp SL (1998) Principles of Animal Communication. Sunderland, MA: Sinauer.

Bradbury JW, Vehrencamp SL (2000) Economic models of animal communication. Anim Behav 59:259–268.

Brainard MS, Doupe AJ (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature 404:762–766.

Brainard MS, Doupe AJ (2001) Postlearning consolidation of birdsong: Stabilizing effects of age and anterior forebrain lesions. J Neurosci 21:2501–2517.

Brauth SE, McHale CM (1988) Auditory pathways in the budgerigar. II. Intratelencephalic pathways. Brain Behav Evol 32:193–207.

Brauth SE, McHale CM, Brasher CA, Dooling RJ (1987) Auditory pathways in the budgerigar. I. Thalamo-telencephalic projections. Brain Behav Evol 30:174–199.

Bregman AS (1990) Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press.

Brenowitz EA (1991) Altered perception of species-specific song by female birds after lesions of a forebrain nucleus. Science 251:303–305.

Brenowitz EA, Margoliash D, Nordeen K (eds) (1997) The neurobiology of birdsong. J Neurobiol 33:495–516.

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

Brosch M, Schreiner CE (2000) Sequence sensitivity of neurons in cat primary auditory cortex. Cereb Cortex 10:1155–1167.

Bruns V, Schmieszek E (1980) Cochlear innervation in the greater horseshoe bat: Demonstration of an acoustic fovea. Hear Res 3:27–43.

Burt JM, Lent KL, Beecher MD, Brenowitz EA (2000) Lesions of the anterior forebrain song control pathway in female canaries affect song perception in an operant task. J Neurobiol 42:1–13.

Capranica RR (1965) The evoked vocal response of the bullfrog. Research Monograph No. 33. Cambridge, MA: MIT Press.

Capranica RR (1978) Auditory processing in anurans. Fed Proc Fed Am Soc Exp Biol 37:2324–2328.

Capsius B, Leppelsack H-J (1999) Response patterns and their relationship to frequency analysis in auditory forebrain centers of a songbird. Hear Res 136:91–99.

Carr CE (1992) Evolution of the central auditory system in reptiles and birds. In: Webster DB, Fay RR, Popper AN (eds) The Evolutionary Biology of Hearing. New York: Springer-Verlag, pp. 511–544.

Carr CE, Code RA (2000) The central auditory system of reptiles and birds. In: Dooling RJ, Fay RR, and Popper AN (eds) Comparative Hearing: Birds and Reptiles. New York: Springer-Verlag, pp. 197–248.

Casseday JH, Covey E (1995) Mechanisms for analysis of auditory temporal patterns in the brainstem of echolocating bats. In: Covey E, Hawkins HL, Port RF (eds) Neural Representations of Temporal Patterns. New York: Plenum Press, pp. 25–51.

Casseday JH, Ehrlich D, Covey E (1994) Neural tuning for sound duration: Role of inhibitory mechanisms in the inferior colliculus. Science 264:847–850.

Casseday JH, Covey E, Grothe B (1997) Neural selectivity and tuning for sinusoidal frequency modulations in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. J Neurophysiol 77:1595–1605.

Chaiken M, Böhner J, Marler P (1993) Song acquisition in European starlings, *Sturnus vulgaris*: A comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. Anim Behav 46:1079–1090.

Cheney DL, Seyfarth RM (1990) How Monkeys See the World. Chicago: University of Chicago Press.

Cherry EC, Taylor WK (1954) Some further experiments on the recognition of speech, with one and two ears. J Acoust Soc Am 26:554–559.

Chew SJ, Mello CV, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses to a repeated conspecific song are long lasting and require two periods of protein synthesis in the songbird forebrain. Proc Natl Acad Sci USA 92:3406–3410.

Chew SJ, Vicario DS, Nottebohm F (1996) A large-capacity memory system that recognizes the calls and songs of conspecifics. Proc Natl Acad Sci USA 93:1950–1955.

Covey E (2000) Neural population coding and auditory temporal pattern analysis. Physiol Behav 69:211–220.

Crawford JD (1997) Feature-detecting auditory neurons in the brain of a sound-producing fish. J Comp Physiol A 180:439–450.

Cynx J (1990) Experimental determination of a unit of song production in the zebra finch (*Taeniopygia guttata*). J Comp Psychol 104:3–10.

Cynx J, Nottebohm F (1992) Role of gender, season, and familiarity in discrimination of conspecific song by zebra finches (*Taeniopygia guttata*). Proc Natl Acad Sci USA 89:1368–1371.

Dabelsteen T, Pedersen SB (1993) Song-based species discrimination and behaviour assessment by female blackbirds, *Turdus merula*. Anim Behav 45:759–771.

Damasio A, Tranel D, Damasio H (1990) Face agnosia and the neural substrate of memory. Annu Rev Neurosci 13:89–109.

Dave AS, Margoliash D (2000) Song replay during sleep and computational rules for sensorimotor vocal learning. Science 290:812–816.

Dave AS, Yu AC, Margoliash D (1998a) Behavioral state modulation of auditory activity in a vocal motor system. Science 282:2250–2254.

Dave A, Yu AC, Gilpin JJ, Margoliash D (1998b) Methods for chronic neuronal ensemble recordings in singing birds. In: Nicolelis MAL (ed) Methods for Neuronal Ensemble Recordings. Boca Raton: CRC Press, pp. 101–120.

Dawkins R, Krebs JR (1978) Animal signals: Information or manipulation? In: Krebs JR, Davies NB (eds) Behavioural Ecology. Oxford: Blackwell Scientific Publications, pp. 282–309.

deCharms RC, Zador A (2000) Neural representation and the cortical code. Annu Rev Neurosci 23:613–647.

Del Negro C, Gahr M, Leboucher G, Kreutzer M (1998) The selectivity of sexual responses to song displays: Effects of partial chemical lesion of the HVC in female canaries. Behav Brain Res 96:151–159.

Del Negro C, Kreutzer M, Gahr M (2000) Sexually stimulating signals of canary (*Serinus canaria*) songs: Evidence for a female-specific auditory representation in the HVc nucleus during the breeding season. Behav Neurol 114:526–542.

DeVoogd TJ, Cardin JA, Szekely T, Büki J, Newman SW (1996) Relative volume of lMAN in female warbler species varies with the number of songs produced by conspecific males. Soc Neurosci Abstr 22:1401.

Dickinson A (1980) Contemporary Animal Learning Theory. New York: Cambridge University Press.

Dooling RJ, Searcy MH (1980) Early perceptual selectivity in the swamp sparrow. Dev Psychobiol 13:499–506.

Dooling RJ, Brown SD, Klump GM, Okanoya K (1992) Auditory perception of conspecific and heterospecific vocalizations in birds: Evidence for special processes. J Comp Psychol 1:20–28.

Dooling RJ, Lohr B, Dent ML (2000) Hearing in birds and reptiles. In: Dooling RJ, Fay RR, Popper AN (eds) Comparative Hearing: Birds and Reptiles. New York: Springer-Verlag, pp. 308–359.

Doupe AJ (1997) Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. J Neurosci 17:1147–1167.

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: Common themes and mechanisms. Annu Rev Neurosci 22:567–631.

Durand S, Tepper J, Cheng M (1992) The shell region of the nucleus ovoidalis: A subdivision of the avian auditory thalamus. J Comp Neurol 323:495–518.

Eaton RC (1983) Is the Mauthner cell a vertebrate command neuron? In: Ewert JP, Capranica RR, Ingle DJ (eds) Advances in Vertebrate Neuroethology. New York: Plenum Press, pp. 629–636.

Eens M, Pinxten M, Verheyen RF (1989) Temporal and sequential organization of song bouts in the European starling. Ardea 77:75–86.

Eens M, Pinxten M, Verheyen RF (1991) Organization of song in the European starling: Species-specificity and individual differences. Belg J Zool 121:257–278.

Ehret G (1997) The auditory cortex. J Comp Physiol A 181:547–557.

Eichenbaum HB, Davis JL (eds) (1998) Neuronal Ensembles: Strategies for Recording and Decoding. New York: John Wiley and Sons.

Evans FE, Whitfield IC (1964) Classification of unit responses in the auditory cortex of the unanesthetized and unrestrained cat. J Physiol (Lond) 171:476–493.

Ewert JP, Burghagen H, Schürg-Pfeiffer E (1983a) Neuroethological analysis of the innate releasing mechanism for prey-catching in toads. In: Ewert JP, Capranica RR, Ingle DJ (eds) Advances in Vertebrate Neuroethology. New York: Plenum Press, pp. 413–475.

Ewert JP, Capranica RR, Ingle DJ (1983b) Advances in Vertebrate Neuroethology. New York: Plenum Press.

Falls JB (1982) Individual recognition by sound in birds. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds. New York: Academic Press, pp. 237–278.

Falls JB (1987) Does song deter territorial intrusion in White-throated Sparrows (*Zonotrichia albicollis*)? Can J Zool 66:206–211.

Falls JB, Brooks RJ (1975) Individual recognition by song in white-throated sparrows. II: Effects of location. Can J Zool 53:1412–1420.

Farries MA, Perkel DJ (2000) Electrophysiological properties of avian basal ganglia neurons recorded in vitro. J Neurophysiol 84:2502–2513.

Fay RR (1998) Auditory stream segregation in goldfish (*Carassius auratus*). Hear Res 120:69–76.

Feng AS, Ratnam R (2000) Neural basis of hearing in real-world situations. Annu Rev Psychol 51:699–725.

Feng AS, Schellart NAM (1999) Central auditory processing in fish and amphibians. In: Popper AN, Fay RR (eds) Comparative Hearing: Fish and Amphibians. New York: Springer-Verlag, pp. 218–268.

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hear Res 151:167–187.

Fitch WT (2000) The evolution of speech: A comparative review. Trends Cogn Sci 4:258–267.

Fitzpatrick DC, Suga N, Olsen JF (1998) Distribution of response types across entire hemispheres of the mustached bat's auditory cortex. J Comp Neurol 391:353–365.

Fodor JA (1983) The Modularity of Mind. Cambridge, MA: MIT Press.

Fortune ES, Margoliash D (1992) Cytoarchitectonic organization and morphology of cells in the field L complex in male zebra finches (*Taeniopygia guttata*). J Comp Neurol 325:388–404.

Fortune ES, Margoliash D (1995) Parallel pathways and convergence onto HVc and adjacent neostriatum of adult zebra finches (*Taeniopygia guttata*). J Comp Neurol 360:413–441.

Fowler CA (1996) Listeners do hear sounds, not tongues. J Acoust Soc Am 99:1730–1741.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. Science 291:312–316.

Fujita I, Tanaka K, Ito M, Cheng K (1992) Columns for visual features of objects in monkey inferotemporal cortex. Nature 360:343–346.

Fuzessery ZM, Feng AS (1982) Frequency selectivity in the anuran auditory midbrain: Single unit responses to single and multiple tone stimulation. J Comp Physiol 146:471–484.

Fuzessery ZM, Feng AS (1983) Mating call selectivity in the thalamus and midbrain of the leopard frog (*Rana p. pipiens*): Single and multiunit analyses. J Comp Physiol 150:333–344.

Gallistel CR (1990) The Organization of Learning. Cambridge, MA: MIT Press.

Gao E, Suga N (1998) Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. Proc Natl Acad Sci USA 95:12663–12670.

Gao E, Suga N (2000) Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: Role of the corticofugal system. Proc Natl Acad Sci USA 97:8081–8086.

Gehr DD, Capsius B, Grabner P, Gahr M, Leppelsack HJ (1999) Functional organization of the field-L-complex of adult male zebra finches. Neuroreport 10:375–380.

Gehr DD, Komiya H, Eggermont JJ (2000) Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. Hear Res 150:27–42.

Gentner TQ (1999) Behavioral and neurobiological mechanisms of song perception among European starlings. Unpublished Ph.D. Thesis, Johns Hopkins University, Baltimore, MD.

Gentner TQ, Hulse SH (1998) Perceptual mechanisms for individual recognition in European starlings (*Sturnus vulgaris*). Anim Behav 56:579–594.

Gentner TQ, Hulse SH (2000) Perceptual classification based on the component structure of song in European starlings. J Acoust Soc Am 107:3369–3381.

Gentner TQ, Margoliash D (2001) Behaviorally relevant selectivity of auditory neurons in the cmHV of European starlings. Sixth Int Congr Neuroethol Abstr:146.

Gentner TQ, Hulse SH, Ball GF (1999) IEG ZENK expression in songbirds during individual vocal recognition. Soc Neurosci Abstr 25:624.

Gentner TQ, Hulse SH, Bentley GE, Ball GF (2000) Individual vocal recognition and the effect of partial lesions to HVc on discrimination, learning, and categorization of conspecific song in adult songbirds. J Neurobiol 42:117–133.

Gentner TQ, Hulse SH, Duffy D, Ball GF (2001) Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. J Neurobiol 46:48–58.

Gerhardt HC (1982) Sound pattern recognition in some North American treefrogs (*Anura: Hylidae*): Implications for mate choice. Am Zool 22:581–595.

Gerhardt HC, Doherty JA (1988) Acoustic communication in the gray treefrog, *Hyla versicolor*: Evolutionary and neurobiological implications. J Comp Physiol A 162:261–278.

Ghazanfar AA, Hauser MD (1999) The neuroethology of primate vocal communication: Substrates for the evolution of speech. Trends Cogn Sci 3:377–384.

Godard R (1991) Long-term memory for individual neighbors in a migratory songbird. Nature 350:228–229.

Gold JI, Knudsen EI (2000) A site of auditory experience-dependent plasticity in the neural representation of auditory space in the barn owl's inferior colliculus. J Neurosci 20:3469–3486.

Grace JA, Theunissen FE (2000) Processing of natural and synthetic sounds in the avian auditory forebrain. Soc Neurosci Abstr 26:2030.

Green S (1975) Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): A field study. In: Rosenblum LA (ed) Primate Behavior. New York: Academic Press, pp. 1–102.

Gross CG (1992) Representation of visual stimuli in inferior temporal cortex. Philos Trans R Soc Lond B Biol Sci 335:3–10.

Gurney ME (1981) Hormonal control of cell form and number in the zebra finch song system. J Neurosci 1:658–673.

Hackett TA, Stepniewska I, Kaas JH (1998) Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 394:475–495.

Hamilton KS, King AP, Sengelaub DR, West MJ (1997) A brain of her own: A neural correlate of song assessment in a female songbird. Neurobiol Learn Mem 68: 325–332.

Hausberger M (1997) Social influences on song acquisition and sharing in the European starling (*Sturnus vulgaris*). In: Snowden C, Hausberger M (eds) Social Influences on Vocal Development. Cambridge, UK: Cambridge University Press.

Hausberger M, Cousillas H (1995) Categorization in birdsong: From behavioural to neuronal responses. Behav Process 35:83–91.

Hauser MD (1998) Functional referents and acoustic similarity: Field playback experiments with rhesus monkeys Anim Behav 56:1309–1310.

Hauser MD, Andersson K (1994) Left hemisphere dominance for processing of vocalizations in adult, but not infant rhesus monkeys: Field experiments. Proc Natl Acad Sci USA 91:3946–3947.

Haüsler U (1997) Measurement of short-time spatial activity patterns during auditory stimulation in the starling. In Syka J (ed) Acoustical Signal Processing in the Central Auditory System. New York: Plenum Press, pp. 85–91.

Hebb DO (1949) The Organization of Behavior: A Neuropsychological Theory. New York: Wiley.

Heffner HE, Heffner RS (1984) Temporal lobe lesions and perception of species-specific vocalizations by macaques. Science 226:75–76.

Heiligenberg WF (1977) Principles of Electrolocation and Jamming Avoidance in Electric Fish: A Neuroethological Approach. Berlin: Springer-Verlag.

Heiligenberg WF (1991) Neural Nets in Electric Fish. Cambridge, MA: MIT Press.

Herrnstein RJ (1990) Levels of stimulus control: A functional approach. Cognition 37:133–166.

Hessler NA, Doupe AJ (1999) Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. J Neurosci 19:10461–10481.

Hiebert SM, Stoddard PK, Arcese P (1989) Repertoire size, territory acquisition and reproductive success in the song sparrow. Anim Behav 37:266–273.

Holekamp KE, Boydston EE, Szykman M, Graham I, Nutt KJ, Birch S, Piskiel A, Singh M (1999) Vocal recognition in the spotted hyaena and its possible implications regarding the evolution of intelligence. Anim Behav 58:383–395.

Hosino T, Okanoya K (2000) Lesion of a higher-order song nucleus disrupts phrase level complexity in Bengalese finches. Neuroreport 11:2091–2095.

Hoy RR, Hahn J, Paul RC (1977) Hybrid cricket auditory behavior: Evidence for genetic coupling in animal communication. Science 195:82–84.

Hulse SH, MacDougall-Shackleton SA, Wisniewski AB (1997) Auditory scene analysis by songbirds: Stream segregation of birdsong by European starlings (*Sturnus vulgaris*). J Comp Psychol 111:3–13.

Insley SJ (2000) Long-term vocal recognition in the northern fur seal. Nature 406:404–405.

Janata P, Margoliash D (1999) Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. J Neurosci 19:5108–5118.

Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. Proc Natl Acad Sci USA 94:4097–4102.

Jones EG (2000) Cortical and subcortical contributions to activity-dependent plasticity in primate somatosensory cortex. Annu Rev Neurosci 23:1–37.

Jones MW, Errington ML, French PJ, Fine A, Bliss TV, Garel S, Charnay P, Bozon B, Laroche S, Davis S (2001) A requirement for the immediate early gene Zif268 in the expression of late LTP and long-term memories. Nat Neurosci 4:289–296.

Jouventin P, Aubin T, Lengagne T (1999) Finding a parent in a king penguin colony: The acoustic system of individual recognition. Anim Behav 57:1175–1183.

Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci USA 97:11793–11799.

Kanwal JS, Fitzpatrick DC, Suga N (1999) Facilitatory and inhibitory frequency tuning of combination-sensitive neurons in the primary auditory cortex of mustached bats. J Neurophysiol 82:2327–2345.

Karten HJ (1968) The ascending auditory pathway in the pigeon (*Columba liva*). II. Telencephalic projections of the nucleus ovoidalis thalami. Brain Res 11:134–153.

Katz LC, Gurney ME (1981) Auditory responses in the zebra finch's motor system for song. Brain Res 211:192–197.

Kelley D, Nottebohm F (1979) Projections of a telencephalic auditory nucleus—field L—in the canary. J Comp Neurol 183:455–470.

Kilgard MP, Merzenich MM (1998) Plasticity of temporal information processing in the primary auditory cortex. Nat Neurosci 1:727–731.

Kilgard MP, Pandya PK, Vazquez J, Gehi A, Schreiner CE, Merzenich MM (2001) Sensory input directs spatial and temporal plasticity in primary auditory cortex. J Neurophysiol 86:326–338.

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

Kluender KR, Diehl RL, Killeen PR (1987) Japanese quail can learn phonetic categories. Science 237:1195–1197.

Klump G (2000) Sound localization in birds. In: Dooling RJ, Fay RR, Popper AN (eds) Comparative Hearing: Birds and Reptiles. New York: Springer-Verlag, pp. 249–307.

Knudsen EI (1985) Experience alters the spatial tuning of auditory units in the optic tectum during a sensitive period in the barn owl. J Neurosci 5:3094–3109.

Kobayashi K, Uno H, Okanoya K (2001) Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches. Neuroreport 12:353–358.

Konishi M (1970) Evolution of design features in the coding of species-specificity. Am Zool 10:67–72.

Konishi M (1978) Auditory environment and vocal development in birds. In: Walk RD, Pick HLJ (eds) Perception and Experience. New York: Plenum, pp. 105–118.

Korzeniewska E, Güntürkün O (1990) Sensory properties and afferents of the N. dorsolateralis posterior thalami of the pigeon. J Comp Neurol 292:457–479.

Krebs J, Ashcroft R, Webber M (1978) Song repertoires and territory defense in the great tit. Nature 271:539–542.

Kuhl PK, Miller JD (1975) Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. Science 190:69–72.

Kupfermann I, Weiss KR (1978) The command neuron concept. Behav Brain Sci 1:3–39.

Kuwabara N, Suga N (1993) Delay lines and amplitude selectivity are created in subthalamic auditory nuclei: The brachium of the inferior colliculus of the mustached bat. J Neurophysiol 69:1713–1724.

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

Lasky RE, Syrdal-Lasky A, Klein RE (1975) VOT discrimination by four to six and a half month old infants from Spanish environments. J Exp Child Psychol 20:215–225.

Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. Nature 401:788–791.

Lehrman DW (1953) A critique of Konrad Lorenz's theory of instinctive behavior. Q Rev Biol 28:337–369.

Lengagne T, Aubin T, Jouventin P, Lauga J (2000) Perceptual salience of individually distinctive features in the calls of adult penguins. J Acoust Soc Am 107:508–516.

Leppelsack HJ (1983) Analysis of song in the auditory pathway of song birds. In: Evert JP, Capranica BR, Ingle DJ (eds) Advances in Vertebrate Neuroethology. New York: Plenum Press, pp. 783–799.

Leppelsack HJ, Schwartzkopff J (1972) Properties of acoustic neurons in the caudal neostriatum of birds. J Comp Physiol A 80:137–140.

Leppelsack HJ, Vogt M (1976) Response to auditory neurons in the forebrain of a song bird to stimulation with species-specific sounds. J Comp Physiol A 107:263–274.

Le Prell CG, Moody DB (1997) Perceptual salience of acoustic features of Japanese monkey coo calls. J Comp Psychol 111:261–274.

Le Prell CG, Moody DB (2000) Factors influencing the salience of temporal cues in the discrimination of synthetic Japanese monkey (*Macaca fuscata*) coo calls. J Exp Psychol Anim Behav Process 26:261–273.

Leroy SA, Wenstrup JJ (2000) Spectral integration in the inferior colliculus of the mustached bat. J Neurosci 20:8533–8541.

Lettvin JY, Maturana HR, McCulloch WS, Pitts WH (1959) What the frog's eye tells the frog's brain. Proc Inst Radio Eng NY 47:1940–1951.

Lewicki MS (1996) Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. J Neurosci 16:5855–5863.

Lewicki MS, Konishi M (1995) Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. Proc Natl Acad Sci USA 92:5582–5586.

Li R, Sakaguchi H (1997) Cholinergic innervation of the song control nuclei by the ventral paleostriatum in the zebra finch: a double-labeling study with retrograde fluorescent tracers and choline acetyltransferase immunohistochemistry. Brain Res 763:239–246.

Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. Cognition 21:1–36.

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code. Psychol Rev 74:431–461.

Lind H, Dabelsteen T, McGregor PK (1997) Female great tits can identify mates by song. Anim Behav 52:667–671.

Lindblom B (1991) The status of phonetic gestures. In: Mattingly IG, Studdert-Kennedy M (eds) Modularity and the Motor Theory of Speech Perception. Hillsdale, NJ: Lawrence Erlbaum, pp. 7–24.

Logan JS, Lively SE, Pisoni DB (1989) Training Japanese listeners to identify /r/ an /l/. J Acoust Soc Am 85:137–138.

Long KD, Kennedy G, Balaban E (2001) Transferring an inborn auditory perceptual predisposition with interspecies brain transplants. Proc Natl Acad Sci USA, 98:5862–5867.

Lotto AJ, Kluender KR, Holt LL (1997) Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). J Acoust Soc Am 102:1134–1140.

Luo M, Perkel DJ (1999) Long-range GABAergic projection in a circuit essential for vocal learning. J Comp Neurol 403:68–84.

MacDougall-Shackleton SA, Hulse SH, Ball GF (1998a) Neural bases of song preferences in female zebra finches (*Taeniopygia guttata*). Neuroreport 9:3047–3052.

MacDougall-Shackleton SA, Hulse SH, Gentner TQ, White W (1998b) Auditory scene analysis by European starlings (*Sturnus vulgaris*): Perceptual segregation of tone sequences. J Acoust Soc Am 103:3581–3587.

Manley GA (1990) Peripheral Hearing Mechanisms in Reptiles and Birds. Berlin: Springer-Verlag.

Manley JA, Müller-Preuss P (1978) Response variability in the mammalian auditory cortex: An objection to feature detection? Fed Proc Fed Am Soc Exp Biol 37:2355–2359.

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

Margoliash D (1986) Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. J Neurosci 6:1643–1661.

Margoliash D (1987) Neural plasticity in birdsong learning. In: Rauschecker JP, Marler P (eds) Imprinting and Cortical Plasticity. New York: Wiley, pp. 23–54.

Margoliash D (1997) Functional organization of forebrain pathways for song production and perception. J Neurobiol 33:671–693.

Margoliash D (2002) Offline learning and the role of autogenous speech: New suggestions from birdsong research. Speech Communication, in press.

Margoliash D, Fortune ES (1992) Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. J Neurosci 12:4309–4326.

Margoliash D, Konishi M (1985) Auditory representation of autogenous song in the song-system of white-crowned sparrows. Proc Natl Acad Sci USA 82:5997–6000.

Margoliash D, Fortune ES, Sutter ML, Yu AC, Wren-Hardin BD, Dave A (1994) Distributed representation in the song system of oscines: Evolutionary implications and functional consequences. Brain Behav Evol 44:247–264.

Marler P, Sherman V (1983) Song structure without auditory feedback: Emendations of the auditory template hypothesis. J Neurosci 3:517–531.

Marr D (1982) Vision. San Francisco: Freeman.

Martin KA (1994) A brief history of the “feature detector.” Cereb Cortex 4:1–7.

May B, Moody DB, Stebbins WC (1988) The significant features of Japanese macaque coo sounds: A psychophysical study. Anim Behav 36:1432–1444.

May B, Moody DB, Stebbins WC (1989) Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. J Acoust Soc Am 85:837–847.

McCasland JS (1987) Neuronal control of bird song production. J Neurosci 7:23–39.

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

McComb K, Moss C, Sayialel S, Baker L (2000) Unusually extensive networks of vocal recognition in African elephants. Anim Behav 59:1103–1109.

McCulloch S, Boness DJ (2000) Mother-pup vocal recognition in the gray seal (*Halichoerus grypus*) of Sable Island, Nova Scotia, Canada. J Zool 251:449–455.

McDonald MV (1989) Function of song in Scott’s seaside sparrow, *Ammodramus maritimus peninsulae*. Anim Behav 38:468–485.

McKibben JR, Bass AH (1998) Behavioral assessment of acoustic parameters relevant to signal recognition and preference in a vocal fish. J Acoust Soc Am 104:3520–3533.

Mello CV, Nottebohm F, Clayton D (1995) Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene’s response to that song in zebra finch telencephalon. J Neurosci 15:6919–6925.

Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. Proc Natl Acad Sci USA 89:6818–6822.

Miller PJO, Bain DE (2000) Within-pod variation in the sound production of a pod of killer whales, *Orcinus orca*. Anim Behav 60:617–628.

Mittman DH, Wenstrup JJ (1995) Combination-sensitive neurons in the inferior colliculus. Hear Res 90:185–191.

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

Mooney R (2000) Different subthreshold mechanisms underlie song selectivity in identified HVc neurons of the zebra finch. J Neurosci 20:5420–5436.

Morris JS, Friston KJ, Dolan RJ (1998) Experience-dependent modulation of tonotopic neural responses in human auditory cortex. Proc R Soc Lond B Biol Sci 265:649–657.

Mudry KM, Constantine-Paton M, Capranica RR (1977) Auditory sensitivity of the diencephalon of the leopard frog (*Rana p. pipiens*). J Comp Physiol 114:1–13.

Müller CM, Leppelsack HJ (1985) Feature extraction and tonotopic organization in the avian forebrain. Exp Brain Res 59:587–599.

Narins PM, Capranica RR (1976) Sexual differences in the auditory system of the tree frog *Eleutherodactylus coqui*. Science 192:378–380.

Nearey TM (1997) Speech perception as pattern recognition. J Acoust Soc Am 101:3241–3254.

Nelken I, Prut Y, Vaadia E, Abeles M (1994a) Population responses to multifrequency sounds in the cat auditory cortex: One- and two-parameter families of sounds. Hear Res 72:206–222.

Nelken I, Prut Y, Vaadia E, Abeles M (1994b) Population responses to multifrequency sounds in the cat auditory cortex: Four-tone complexes. Hear Res 72:223–236.

Nelson DA (1989) Song frequency as a cue for recognition of species and individuals in the field sparrow (*Spizella pusilla*). J Comp Psychol 103:171–176.

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

Nicolelis MAL (ed) (2001) Advances in Neural Population Coding. Progress in Brain Research, Vol 130. Amsterdam: Elsevier.

Nieder A, Klump GM (1999) Adjustable frequency selectivity of auditory forebrain neurons recorded in a freely moving songbird via radiotelemetry. Hear Res 127:41–54.

Nordeen KW, Nordeen EJ (1993) Long-term maintenance of song in adult zebra finches is not affected by lesions of a forebrain region involved in song learning. Behav Neural Biol 59:79–82.

Nottebohm F, Arnold AP (1976) Sexual dimorphism in vocal control areas of the songbird brain. Science 194:211–213.

Nottebohm F, Stokes TM, Leonard CM (1976) Central control of song in the canary, *Serinus canarius*. J Comp Neurol 165:457–486.

Nudo RJ, Wise BM, SiFuentes F, Milliken GW (1996) Neural substrates for the effects of rehabilitative training on motor recovery after ischemic infarct. Science 272:1791–1794.

Ohlemiller KK, Kanwal JS, Suga N (1996) Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. Neuroreport 7:1749–1755.

Okanoya K, Dooling RJ (1990) Song syllable perception in song sparrows (*Melospiza melodia*) and swamp sparrows (*Melospiza georgiana*): An approach from animal psychophysics. Bull Psychon Soc 28:221–224.

Okanoya K, Yamaguchi A (1997) Adult Bengalese finches (Lonchura striata var. domestica) require real-time auditory feedback to produce normal song syntax. J Neurobiol 33:343–356.

Okanoya K, Tsumaki S, Honda E (2000) Perception of temporal properties in self-generated songs by Bengalese finches (*Lonchura striata var. domestica*). J Comp Psychol 114:239–245.

O'Loghlen AL, Beecher MD (1997) Sexual preferences for mate song types in female song sparrows. Anim Behav 53:835–841.

O'Loghlen AL, Beecher MD (1999) Mate, neighbor and stranger songs: A female song sparrow perspective. Anim Behav 58:13–20.

Olsen JF (1994) Medial geniculate neurons in the squirrel monkey sensitive to intercomponent delays that categorize species-specific calls. Assoc Res Otolaryngol Abstr 21:84.

Olsen JF, Suga N (1991a) Combination-sensitive neurons in the medial geniculate body of the mustached bat: Encoding of relative velocity information. J Neurophysiol 65:1254–1274.

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

O'Neill WE (1985) Responses to pure tones and linear FM components of the CF-FM biosonar signal by single units in the inferior colliculus of the mustached bat. J Comp Physiol A 157:797–815.

O'Neill WE (1995) The bat auditory cortex. In: Popper AN, Fay RR (eds) Hearing by Bats. New York: Springer-Verlag, pp. 416–480.

O'Neill WE, Suga N (1979) Target range-sensitive neurons in the auditory cortex of the mustache bat. Science 203:69–73.

Owings DH, Morton ES (1998) Animal Vocal Communication: A New Approach. Cambridge, UK: Cambridge University Press.

Owren MJ, Casale TM (1994) Variations in fundamental frequency peak position in Japanese macaque (*Macaca fuscata*) coo calls. J Comp Psychol 108:291–297.

Owren MJ, Seyfarth RM, Cheney DL (1997) The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions. J Acoust Soc Am 101:2951–2963.

Peek FW (1972) An experimental study of the territorial function of vocal and visual display in the male red-winged blackbird (*Agelaius phoeniceus*). Anim Behav 20:112–118.

Pereira MR (1986) Maternal recognition of juvenile offspring coo vocalizations in Japanese macaques. Anim Behav 34:935–937.

Perkel DJ, Farries MA (2000) Complementary "bottom-up" and "top-down" approaches to basal ganglia function. Curr Opin Neurobiol 10:725–731.

Pfingst BE, O'Connor TA, Miller JM (1977) Response plasticity of neurons in auditory cortex of the rhesus monkey. Exp Brain Res 29:393–404.

Pons TP, Garraghty PE, Mishkin M (1991) Serial and parallel processing of tactual information in somatosensory cortex of rhesus monkeys. J Neurophysiol 68:518–527.

Popper AN, Fay RR (eds) (1995) Hearing by Bats. New York: Springer-Verlag.

Portfors CV, Wenstrup JJ (1999) Delay-tuned neurons in the inferior colliculus of the mustached bat: Implications for analyses of target distance. J Neurophysiol 82:1326–1338.

Pytte CL, Suthers RA (1999) A bird's own song contributes to conspecific song perception. Neuroreport 10:1773–1778.

Rauschecker JP (1997) Processing of complex sounds in the auditory cortex of cat, monkey, and man. Acta Otolaryngol Suppl 532:34–38.

Rauschecker JP (1998) Cortical processing of complex sounds. Curr Opin Neurobiol 8:516–521.

Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci USA 97:11800–11806.

Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268:111–114.

Rauske PL, Margoliash D (1999) Does behavioral state modulate sensorimotor properties in HVc? Soc Neurosci Abstr 25:624.

Recanzone GH (1998) Rapidly induced auditory plasticity: The ventriloquism aftereffect. Proc Natl Acad Sci USA 95:869–875.

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13:87–103.

Recanzone GH, Guard DC, Phan ML (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. J Neurophysiol 83:2315–2331.

Rendall D, Rodman PS, Emond RE (1996) Vocal recognition of individuals and kin in free-ranging rhesus monkeys. Anim Behav 51:1007–1015.

Rendall D, Owren MJ, Rodman PS (1998) The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. J Acoust Soc Am 103:602–614.

Rendall D, Seyfarth RM, Cheney DL, Owren MJ (1999) The meaning and function of grunt variants in baboons. Anim Behav 57:583–592.

Ribeiro S, Cecchi GA, Magnasco MO, Mello CV (1998) Toward a song code: Evidence for a syllabic representation in the canary brain. Neuron 21:359–371.

Rieke F, Warland D, Van Steveninck RDR, Bialek W (1996) Spikes, Exploring the Neural Code. Cambridge, MA: MIT Press.

Riquimaroux H, Gaioni SJ, Suga N (1991) Cortical computational maps control auditory perception. Science 251:565–568.

Roeder KD (1966) Auditory system of noctuid moths. Science 154:1515–1521.

Rolls ET (1992) Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. Philos Trans R Soc Lond B Biol Sci 335:11–20.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–1136.

Rose G, Capranica RR (1983) Temporal selectivity in the central auditory system of the leopard frog. Science 219:1087–1089.

Rose GJ, Kawasaki M, Heiligenberg W (1988) "Recognition units" at the top of a neuronal hierarchy? Prepacemaker neurons in *Eigenmannia* code the sign of frequency differences unambiguously. J Comp Physiol A 162:759–772.

Rubsamen R, Dorrscheidt GJ (1986) Tonotopic organization of the auditory forebrain in a songbird, the European starling. J Comp Physiol A 158:639–646.

Ryan MJ (1998) Sexual selection, receiver biases, and the evolution of sex differences. Science 281:1999–2003.

Sakai M, Suga N (2001) Plasticity of the cochleotopic (frequency) map in specialized and nonspecialized auditory cortices. Proc Natl Acad Sci USA 98:3507–3512.

Sayigh LS, Tyack PL, Wells RS, Solow AR, Scott MD, Irvine AB (1999) Individual recognition in wild bottlenose dolphins: A field test using playback experiments. Anim Behav 57:41–50.

Schäfer M, Rubsamen R, Dorrscheidt GJ, Knipschild M (1992) Setting complex tasks to single units in the avian auditory forebrain. II. Do we really need natural stimuli to describe neuronal response characteristics? Hear Res 57:231–244.

Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: Implications for vocal learning. J Neurosci 11:2896–2913.

Scharff C, Nottebohm F, Cynx J (1998) Conspecific and heterospecific song discrimination in male zebra finches with lesions in the anterior forebrain pathway. J Neurobiol 36:81–90.

Scheich H, Bonke BA, Bonke D, Langer HJ (1979a) Functional organization of some auditory nuclei in the guinea fowl demonstrated by the 2-deoxyglucose technique. Cell Tissue Res 204:17–27.

Scheich H, Langner G, Bonke D (1979b) 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.

Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. Nat Neurosci 1:513–518.

Schulze H, Langner G (1997) Periodicity coding in the primary auditory cortex of the Mongolian gerbil (*Meriones unguiculatus*): Two different coding strategies for pitch and rhythm? J Comp Physiol A 181:651–663.

Schwarz DW, Tomlinson RW (1990) Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*). J Neurophysiol 64:282–298.

Searcy WA (1996) Sound-pressure levels and song preferences in female red-winged blackbirds (*Agelaius phoeniceus*) (Aves, Emberizidae). Ethology 102:187–196.

Searcy WA, Yasukawa K (1996) Song and female choice. In: Kroodsma DE, Miller EH (eds) Ecology and Evolution of Acoustic Communication in Birds. Ithaca, NY: Cornell University Press, pp. 454–473.

Seyfarth RM, Cheney DL, Marler P (1980a) Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. Science 210:801–803.

Seyfarth RM, Cheney DL, Marler P (1980b) Vervet monkey alarm calls: Semantic communication in a free-ranging primate. Anim Behav 28:1070–1094.

Shea SD, Rauske PL, Margoliash D (2001) Identification of HVc projection neurons in extracellular records by antidromic stimulation. Soc Neurosci Abstr 27:381.6.

Simoncelli E, Olshausen A (2001) Natural image statistics and neural representation. Annu Rev Neurosci 24:1193–1215.

Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18:555–586.

Sinnott JM (1980) Species-specific coding in bird song. J Acoust Soc Am 68:494–497.

Smith DG, Reid FA (1979) Roles of the song repertoire in red-winged blackbirds. Behav Ecol Sociobiol 5:279–290.

Smith WJ (1997) The behavior of communicating, after twenty years. In: Owings DH, Beecher MD, Owings DH, Thompson NS (eds) Perspectives in Ethology, Vol 12: Communication. New York: Plenum Press, pp. 7–53.

Sohrabji F, Nordeen EJ, Nordeen KW (1990) Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. Behav Neural Biol 53:51–63.

Solis MM, Brainard MS, Hessler NA, Doupe AJ (2000) Song selectivity and sensorimotor signals in vocal learning and production. Proc Natl Acad Sci USA, 97:11836–11842.

Stoddard PK (1996) Vocal recognition of neighbors by territorial passerines. In: Kroodsma DE, Miller EH (eds) Ecology and Evolution of Acoustic Communication in Birds. Ithaca, NY: Cornell University Press, pp. 356–374.

Stoddard PK, Beecher MD, Loesche P, Campbell SE (1992) Memory does not constrain individual recognition in a bird with song repertoires. Behaviour 122:274–287.

Stripling R, Volman SF, Clayton DF (1997) Response modulation in the zebra finch neostriatum: Relationship to nuclear gene expression. J Neurosci 17:3883–3893.

Suga N, Horikawa J (1986) Multiple axes for representation of echo delays in the auditory cortex of the mustached bat. J Neurophysiol 55:776–805.

Suga N, O'Neill WE, Manabe T (1978) Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat. Science 200:778–781.

Suga N, O'Neill WE, Manabe T (1979) Harmonic-sensitive neurons in the auditory cortex of the mustache bat. Science 203:270–274.

Suga N, Zhang Y, Yan J (1997) Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. J Neurophysiol 77:2098–2114.

Suga N, Yan J, Zhang Y (1998) The processing of species-specific sounds by the ascending and descending auditory systems. In: Poon PWF, Brugge JF (eds) Central Auditory Processing and Neural Modeling. New York: Plenum Press.

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

Sutter ML, Schreiner CE (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. J Neurophysiol 65:1207–1226.

Takahashi T, Moiseff A, Konishi M (1984) Time and intensity cues are processes independently in the auditory system of the owl. J Neurosci 4:1781–1786.

Taniguchi I, Niwa H, Wong D, Suga N (1986) Response properties of FM-FM combination-sensitive neurons in the auditory cortex of the mustached bat. J Comp Physiol A 159:331–337.

Theunissen FE, Doupe AJ (1998) Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches. J Neurosci 18:3786–3802.

Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of non-linear auditory neurons obtained using natural sounds. J Neurosci 20:2315–2331.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. Science 292:290–293.

Tinbergen N (1963) On aims and methods of ethology. Z Tierpsychol 20:410–433.

Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat Neurosci 4:832–838.

Tsuzuki K, Suga N (1988) Combination-sensitive neurons in the ventroanterior area of the auditory cortex of the mustached bat. J Neurophysiol 60:1908–1923.

Ulinski PS (1984) Design features in vertebrate sensory systems. Am Zool 24:717–731.

Ulinski PS, Margoliash D (1990) Neurobiology of the reptile-bird transition. In: Jones EG, Peters A (eds) Cerebral Cortex. New York: Plenum, pp. 217–265.

Vallet E, Kreutzer M (1995) Female canaries are sexually responsive to special song phrases. Anim Behav 49:1603–1610.

van Dijk P, Wit HP, Segenhout JM (1997) Dissecting the frog inner ear with Gaussian noise. II. Temperature dependence of inner ear function. Hear Res 114:243–251.

Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*Taenopygia guttata*). J Comp Neurol 366:613–642.

Venkatachalam S, Fee MS, Kleinfeld D (1999) Ultra-miniature headstage with 6-channel drive and vacuum-assisted micro-wire implantation for chronic recording from the neocortex. J Neurosci Methods 90:37–46.

Vicario DS, Yohay KH (1993) Song-selective auditory input to a forebrain vocal control nucleus in the zebra finch. J Neurobiol 24:488–505.

Vicario DS, Naqvi NH, Raksin JN (2001) Behavioral discrimination of sexually dimorphic calls by male zebra finches requires an intact vocal motor pathway. J Neurobiol 47:109–120.

Volman S (1993) Development of neural selectivity for birdsong during vocal learning. J Neurosci 13:4737–4747.

von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. Science 224:1260–1262.

von Helversen O, von Helversen D (1994) Forces driving coevolution of song and song recognition in grasshoppers. Fortschr Zool 39:253–284.

Vu ET, Mazurek ME, Kuo Y-C (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. J Neurosci 14:6924–6934.

Vu ET, Kuo Y, Chance FS (1995) Effects of lesioning nucleus interfacialis on adult zebra finch song. Soc Neurosci Abstr 21:964.

Wang X, Kadia SC (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. J Neurophysiol 86:2616–2620.

Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics. J Neurophysiol 74:2685–2706.

Wang Y, Fujita I, Murayama Y (2000) Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. Nat Neurosci 3:807–813.

Wanker R, Apcin J, Jennerjahn B, Waibel B (1998) Discrimination of different social companions in spectacled parrotlets (*Forpus conspicillatus*): Evidence for individual vocal recognition. Behav Ecol Sociobiol 43:197–202.

Weary DM, Krebs JR (1992) Great tits classify songs by individual voice characteristics. Anim Behav 43:283–287.

Wenstrup JJ (1999) Frequency organization and responses to complex sounds in the medial geniculate body of the mustached bat. J Neurophysiol 82:2528–2544.

Wenstrup JJ, Grose CD (1995) Inputs to combination-sensitive neurons in the medial geniculate body of the mustached bat: The missing fundamental. J Neurosci 15:4693–4711.

Wenstrup JJ, Leroy SA (2001) Spectral integration in the inferior colliculus: Role of glycinergic inhibition in response facilitation. J Neurosci 21:RC124.

Wenstrup JJ, Mittmann DH, Grose CD (1999) Inputs to combination-sensitive neurons of the inferior colliculus. J Comp Neurol 409:509–528.

Werker JF, Tees RC (1984) Cross language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behav Dev 7:49–63.

Wiley HB, Hatchwell BJ, Davies NB (1991) Recognition of individual males' songs by female Dunnocks: A mechanism increasing the number of copulatory partners and reproductive success. Ethology 88:145–153.

Williams H, Mehta N (1999) Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. J Neurobiol 39:14–28.

Williams H, Nottebohm F (1985) Auditory responses in avian vocal motor neurons: A motor theory for song perception in birds. Science 229:279–282.

Williams H, Vicario DS (1993) Temporal patterning of song production: Participation of nucleus uvaeformis of the thalamus. J Neurobiol 24:903–912.

Winter P, Funkenstein HH (1973) The effects of species-specific vocalization on the discharge of auditory cortical cells in the awake squirrel monkey (*Saimiri sciureus*). Exp Brain Res 18:489–504.

Wollberg Z, Newman JD (1972) Auditory cortex of squirrel monkey: Response patterns of single cells to species-specific vocalizations. Science 175:212–214.

Wyttenbach RA, May ML, Hoy RR (1996) Categorical perception of sound frequency by crickets. Science 273:1542–1544.

Yan J, Suga N (1999) Corticofugal amplification of facilitative auditory responses of subcortical combination-sensitive neurons in the mustached bat. J Neurophysiol 81:817–824.

Yasukawa K (1981) Song repertoires in the red-winged blackbird (*Agelaius phoeniceus*): A test of the Beau Geste hypothesis. Anim Behav 29:114–125.

Yu A, Margoliash D (1996) Temporal hierarchical control of singing in birds. Science 273:1871–1875.

Zhang Y, Suga N (2000) Modulation of responses and frequency tuning of thalamic and collicular neurons by cortical activation in mustached bats. J Neurophysiol 84:325–333.

Zhang Y, Suga N, Yan J (1997) Corticofugal modulation of frequency processing in bat auditory system. Nature 387:900–903.

Zoloth SR, Peterson MR, Beecher MD, Green SG, Marler P, Moody DB, Stebbins WC (1979) Species-specific perceptual processing of vocal sounds by monkeys. Science 204:870–873.