# Distributed Time-Domain Representations in the Birdsong System

# Minireview

#### Daniel Margoliash

Department of Organismal Biology and Anatomy The University of Chicago Chicago, Illinois 60637

Vocal communication systems are common throughout the animal kingdom, and learned vocalizations such as human speech or bird song are of particular interest. How are vocalizations represented in the brain? Neural codes represent solutions to information-processing problems posed by an animal's behavioral requirements. In vocal communication, sounds are used as symbols, embodying certain parameters or features that have special significance for a given species. For example, the song of male crickets will attract females, but only if it embodies certain species-specific characteristics. Physical attributes that affect sound transmission through the acoustic habitat may affect the design of vocalizations, as in the case of long-distance communication, but the use of symbols suggests a certain arbitrariness as to the parameters of the symbols. Additionally, where vocal communication is invoked as part of reproductive behavior, it evolves under sexual selection, which can favor solutions that are not optimal in the engineering sense. The idiosyncratic, species-specific features of bird songs-essential diagnostic tools for bird watchers—provide examples of this type of system.

In many vertebrates, the ascending auditory system is organized tonotopically (spatial representation of sound frequency), and the descending motor system for vocalizations is ultimately organized spatially based on muscles of the vocal organ and vocal tract. The representations of basic units of behavior (e.g., songs) in the higher motor centers, however, do not map continuously onto spatial coordinates of either system. An equivalent result will generally obtain for higher-order representations of other behaviors and in other sensory systems (e.g., visual objects). The question of how information is represented at the sensorimotor level of the vocal-motor system is therefore of considerable interest. One possibility is that information about components of vocal behavior is represented spatially, perhaps with local domains representing individual time or spectral components of the vocal signal. A good example of this type of organization is the representation of different components of pulse and echo signals in distinct areas of the bat auditory cortex, with continuous spatial representation of "information-bearing" parameters within each cortical area (Suga, 1990). Another possibility is that information about higher-level aspects of the behavior is not represented spatially but globally and in a distributed fashion. In this case, the unifying dimension could be time, represented by the synchronous activity of populations of neurons. This minireview focuses on recent evidence suggesting the latter form of organization in a sensorimotor component of the vocal-motor ("song") system of song birds.

### The Birdsong System

The forebrain pathways for song comprise three or more separate streams of information flow (reviewed by Margoliash, 1997). The ascending auditory system (green, Figure 1) sends auditory input to HVc and NIf, and perhaps other song system nuclei, via a complex web of interconnections that starts with field L, the primary thalamorecipient structure. The nuclei Uva, NIf, HVc, and RA are connected sequentially and unidirectionally, forming a descending pathway for song production ultimately projecting to the brainstem hypoglossal nucleus, which in turn projects to the avian vocal organ, the syrinx (red, Figure 1). Large bilateral lesions of Uva affect timing of elements within song, and animals with large bilateral lesions of HVc and RA attempt to sing but fail to produce songs. (The status of NIf in song production is uncertain.) The anterior forebrain pathway (AFP) starts at the level of HVc and ultimately projects to RA (blue, Figure 1). Lesions of AFP nuclei have little effect on adult song but severely disrupt the development of song (Bottjer et al., 1984). (Logically, mMAN and DMP should be included in these pathways [see Figure 1], although the role of these nuclei in song development or song production is unclear.)

The two forebrain nuclei necessary for singing, HVc and RA, have contrasting and complementary patterns of connections: HVc receives multiple inputs and has two outputs, whereas RA sends multiple outputs and has two inputs (Figure 1). Another organizational feature that distinguishes HVc from many of the other nuclei is the apparent lack of topography. The auditory pathways are tonotopically organized, and the AFP and RA are organized in a topography that ultimately relates to the muscles of the syrinx (see Figure 1). In contrast, spatial topography has yet to be delineated for the HVc. Small injections of HVc label fibers throughout its efferent targets (RA and area X), with individual HVc fibers ramifying extensively within those targets. The same injections also label somata and fibers throughout HVc. Small injections of structures afferent to HVc label fibers throughout HVc, without any evidence of small sprays of boutons common in topographically organized projections. The position of HVc in the motor and sensory pathways, the high degree of convergence onto HVc and divergence of HVc's outputs, and the apparent absence of topography within HVc suggest it as a candidate site for higher-level representations within the song system.

## Auditory Representations of Song

Birds learn song by reference to auditory feedback; hence, sites of sensorimotor interactions are also candidate sites where experience modifies circuitry, i.e., sites of learning. The auditory response properties of neurons in the adult song system of finches and sparrows (reviewed by Margoliash et al., 1994) confirm this reasoning. Both HVc and RA neurons exhibit song-selective auditory response properties in urethane-anesthetized birds, but recent evidence suggests that the RA responses are uncovered by anesthesia—they are not commonly observed in awake animals (see Margoliash,



#### Figure 1. Connections of the Song System

The song system of oscine passerine birds can be subdivided into functional streams. In this representation, auditory information (green) ascends and enters the song system at HVc and NIf. "Motor" structures (red) are those in the forebrain necessary for singing (HVc and RA) and the midbrain and brainstem structures that receive from RA. "Motor contributory" structures also have a role in singing but are not essential. Anterior forebrain pathways are involved in feedforward and feedback connections and have been implicated in song development. Many song system structures are topography failed (-). A list of abbreviations can be found in Margoliash et al., 1994.

1997 for a discussion of this unresolved issue). HVc neurons have similar song-selective properties in both urethanized and awake conditions. In each individual bird, neurons in HVc have highly specific responses to that bird's own song (BOS). Some of these neurons are temporal combination sensitive (TCS), responding only when presented with two to six notes or syllables of a sequence drawn from BOS. Thus, remarkable higherorder time domain processing is observed in the auditory response properties of these HVc neurons. TCS units can exhibit exquisite sensitivity to acoustic parameters of BOS. For example, white-throated sparrow songs were modeled with a zero-crossing technique that extracted cycle-by-cycle frequency and amplitude information. When the slight ( $\sim$ 10–20 Hz) frequency modulations that give bird whistles their appealing vibrato were suppressed or eliminated, leaving amplitude information unaltered, some units gave weaker or no response (Margoliash, 1983). In contrast, TCS units were quite insensitive to changes in the duration of individual components of the requisite sequence of sounds or to changes in the intervals between individual components. The TCS response was maintained even when units integrated over abnormal intervals of many hundreds of milliseconds. The juxtaposition of heightened sensitivity and special processing for temporal sequences of component sounds and fine frequency modulation of each component, with insensitivity to component and interval duration, is consistent with the behaviorally derived sensitivities of finches and sparrows to acoustic parameters of song and emphasizes the symbolic nature of these learned signals.

Non-TCS HVc neurons also exhibit stronger responses

to BOS than to conspecific (same-species) songs. When BOS is presented reversed in time, these units exhibit weaker responses than to normal BOS, implicating timevarying features of the acoustics in specifying the neuronal response. HVc neurons in adults even prefer BOS over the songs the birds were tutored with early in life. The acoustic parameters of song are learned, and auditory properties of HVc neurons in juvenile whitecrowned sparrows appear to track the ever-changing features of plastic song during the sensorimotor learning phase of song development (Volman, 1993). Thus, there is a tight linkage between the sensory and motor properties of HVc neurons throughout ontogeny. There is little insight to date into the cellular and circuit basis for this linkage, a central issue in the neural mechanisms of sensorimotor learning in birdsong development.

Several aspects of the organization of auditory responses in HVc suggest a distributed representation of song (reviewed by Margoliash et al., 1994). Virtually all HVc neurons respond more strongly to BOS than to conspecific songs, no matter which conspecific song, which part of HVc one records from, and whether the neurons are phasically or tonically responsive to song. Intracellular recordings show that HVc neurons have song-selective inputs and suggest that local convergence from HVc neurons (not just integration of afferent input) is important in producing the song-selective and TCS properties of HVc neurons (Lewicki and Konishi, 1995). To date, song-selective properties akin to those observed in HVc have yet to be reported in structures afferent to HVc, for example in the subdivisions of field L (Figure 1). The connectivity of structures leading up to HVc is still being delineated, however, and physiological analysis of these structures is quite incomplete. The conclusion that selectivity for BOS appears en masse at the level of HVc remains tentative. Such a conclusion would stand in striking contrast to the sensory hierarchies that have been described in other systems (e.g., see Konishi, 1990; Heiligenberg, 1991).

Another aspect of auditory activity in HVc that suggests a distributed representation is the similarity of temporal features of response properties across the population of HVc neurons. When presented with BOS, most HVc multiunit recording sites (and single units extracted from the multiunits) exhibit strong excitatory peaks of activity for some syllables and weaker activity for other syllables (Sutter and Margoliash, 1994). This suggests that excitation and inhibition across HVc exhibit a considerable degree of synchronization during presentation of BOS. Across recording sites, the variation in response to BOS could just be noise. Alternatively, information could be encoded in the variation around the overall pattern; for example, a spatial representation for syllables that depends on time could exist. Such a representation could only be detected if both spatial and temporal properties of HVc were assessed simultaneously, which has yet to be reported. When a few sites in HVc were recorded simultaneously, strong temporal synchronization of activity was observed in the spontaneous bursts that are characteristic of ongoing activity in HVc, even in sites recorded from opposite ends of the nucleus. This could arise from the high degree of local (intrinsic) connectivity observed in HVc,



Figure 2. Terminology for Zebra Finch Song Song is represented by spectrographs, with frequency on the ordinate versus time on the abcissa and with the signal power denoted by gray levels. Components of songs of zebra finches include "notes," the smallest subdivision; "syllables," which comprise one or more notes; and "motifs," which comprise one or more syllables. The color bands underneath the song denote individual notes, syllables, and motifs of the same type.

from correlated input, or from both. The synchronization of spontaneous and driven activity throughout HVc can be compared to the stimulus response-dependent synchronization of unit activity recently demonstrated in mammalian auditory cortex (deCharms and Merzenich, 1996). These and other studies suggest that precise correlation of spike timing is a general mechanism for representing information across populations of neurons.

Sutter and Margoliash (1994) observed that a single narrow (25 ms) window of time accounted for the strongest peak of excitation for a majority or plurality of HVc units, depending on the bird. Across individuals, the timing of the peak of excitation could not be predicted by any surface acoustic features of the associated syllable of maximum synchronization (SMS), including amplitude, duration, morphology, and sequence position (in which motif and in which part of the motif the SMS occurred). Presumably, the temporal pattern of excitation and inhibition during presentation of BOS is a consequence of the patterns of connections established during sensorimotor learning, which will vary across birds learning to sing different songs. Apparently, all HVc neurons are premotor (see below) and many are also auditory, but in adult birds, individual HVc neurons do not exhibit any obvious similarity in their auditory and motor responses. Perhaps such similarities might emerge in the activity of the population of neurons, especially if examined during the period of vocal learning.

#### Song Production

Studies of information coding along the descending motor pathway for song give additional insight into the organization of HVc and support the notions of a timedomain distributed representation. Vu et al. (1994) electrically stimulated song system nuclei with small currents ( $<5 \mu$ A) while zebra finches were singing. When applied to HVc, such currents induced cessation of singing followed rapidly by recommencement of singing starting with a new motif (group of syllables; see Figure 2 for terminology of song). When applied to RA, such currents induced distortions of the syllable the bird was singing but produced no disruption of the temporal pattern of song. This suggested that timing of larger units of song is organized at the level of HVc or its afferents, and coding of syllable morphology is organized at the level of RA and its efferents.

In a complementary approach, Yu and Margoliash (1996) developed a technique to record single neurons from singing zebra finches. In HVc, neurons had relatively tonic premotor activity patterns starting prior to each syllable of song. Each neuron had a different pattern of activity associated with each different syllable

type. The relative position of the syllable within the sequence of syllables (e.g., which syllable followed or preceded the target syllable) did not affect the activity pattern of HVc neurons: HVc neurons were insensitive to temporal context. Interestingly, all HVc neurons were active for all syllables of song. Neurons recorded in RA had a contrasting pattern of activity. During singing, RA neurons exhibited short, phasic bursts of activity whose temporal pattern was extremely tightly regulated. All RA neurons were also active during singing but not necessarily for all syllables of song. Thus, a spatial topography of excitation and inhibition that depends on vocal units may be present in RA but not in HVc. The temporal pattern of activation of RA neurons depended only on the identity of individual elements (notes) within a syllable, and RA neurons were insensitive to the temporal context of notes. The different temporal scales of coding in HVc and RA were highlighted in those cases in which a bird sang two syllables that had some notes in common but had other notes that were different (e.g., ABCD and ABEF). HVc neurons exhibited different activity profiles for the two syllables, including the parts of the syllables that had common notes, whereas RA neurons exhibited similar burst patterns for similar notes and dissimilar burst patterns for dissimilar notes.

Based on these observations. Yu and Margoliash (1996) concluded that the population of HVc neurons codes for syllable identity, whereas the population of RA neurons codes for note identity. Temporal sequencing and different scales of time are hierarchically organized in the bird forebrain, matching the hierarchical organization of song (Figure 2). In this scheme, the temporal sequence of notes of a syllable results from an interaction of HVc input with RA local circuits. HVc "plays" RA (in the sense of a musical instrument) at the syllable level, and RA in turn transforms these commands into the appropriate sequence of notes. This permits HVc to have considerable information about syllables with little information about notes. A similar interaction is envisioned for Uva (and NIf) input to HVc organizing sequences of syllables into motifs and introductory notes and motifs into songs.

The differences in the activity patterns of HVc and RA neurons gives some insight into the form of the interaction that must occur at the level of RA. RA neurons could exhibit up to  $\sim$ 10 separate bursts for a single complex syllable lasting up to  $\sim$ 400 ms. The number of spikes and the spike pattern of each burst was sufficiently well regulated (and different) to form a "signature" of the behavior, so that in favorable cases it was possible to "read" the singing behavior from individual spike trains.

In some cases, the temporal precision between particular pairs of spikes within each burst was regulated at least within the timing precision with which the raw signals were originally sampled (20 kHz, or 50 µs/sample). A plausible scheme to explain these observations posits that the RA network is reconfigured for each new burst, changing the relative coupling between RA neurons, which have distinct subthreshold and suprathreshold oscillatory activity (Mooney, 1992). Reconfiguration could occur perhaps via the action of neuromodulators, as has been shown in lobster stomatogastric ganglion. The final burst pattern would reflect primarily the action of local RA circuitry, thereby achieving the precision in burst patterning. It is noteworthy that the timing of burst onsets of RA neurons was less precisely regulated than the timing of within-burst patterns. Potentially, the relative lack of precision in burst onsets could arise as an artifact from time/frequency resolution limitations in the Fourier transform signal-processing techniques used to align the spike data with the notes. If so, signal-processing techniques that better determine note boundaries might demonstrate improvement in the timing of burst onsets. Alternatively, and consistent with the proposed scheme, lack of precision in the timing of burst onsets could be biologically significant, arising from the relative temporal imprecision of HVc input. Perhaps burst onsets are then brought into better temporal register at the level of the hypoglossal nucleus.

## Conclusions

How vocalizations are represented in the auditory system is a long-standing problem. Both gnostic "recognition" neurons and distributed representations have been proposed as possible mechanisms. In the birdsong system, there is evidence for both. Neurons with highly selective auditory properties are commonly observed, with some of these exhibiting among the most complex requirements for temporal sequences of sounds ever reported. The mapping of auditory representations onto motor representations does not appear to be organized in a spatially topographic fashion, with subpopulations of neurons whose firing rate specifies different features of the vocalization, but rather appears to be distributed in space and highly synchronized in time. Synchronous activity has been observed in a number of sensorimotor systems, but its significance is not well understood. The analysis of how birds learn songs may contribute to resolution of this problem.

#### Selected Readings

Bottjer, S.W., Miesner, E.A., and Arnold, A.P. (1984). Science 224, 901–903.

deCharms, R.C., and Merzenich, M.M. (1996). Nature *381*, 610–613. Heiligenberg, W.F. (1991). Neural Nets in Electric Fish (Cambridge, MA: Massachusetts Institute of Technology Press).

Konishi, M. (1990). Cold Spring Harb. Symp. Quant. Biol. 55, 575-584.

Lewicki, M.S., and Konishi, M. (1995). Proc. Natl. Acad. Sci. USA 92, 5582–5586.

Margoliash, D. (1983). J. Neurosci. 3, 1039-1057.

Margoliash, D. (1997). J. Neurobiol. 33, 671-693.

Margoliash, D., Fortune, E.S., Sutter, M.L., Yu, A.C., Wren-Hardin, B.D., and Dave, A. (1994). Brain Behav. Evol. 44, 247–264.

Mooney, R. (1992). J. Neurosci. 12, 2464-2477.

Suga, N. (1990). Neural Networks 3, 3-21.

Sutter, M.L., and Margoliash, D. (1994). J. Neurophys. 72, 2105–2123. Volman, S. (1993). J. Neurosci. 13, 4737–4747.

Vu, E.T., Mazurek, M.E., and Kuo, Y.-C. (1994). J. Neurosci. 14, 6924–6934.

Yu, A.C., and Margoliash, D. (1996). Science 273, 1871-1875.