

Gradual Emergence of Song Selectivity in Sensorimotor Structures of the Male Zebra Finch Song System

Petr Janata and Daniel Margoliash

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

Birdsong is a model system for understanding how motor and sensory information interact to coordinate behavior. Neurons in one potential site of sensorimotor integration, the forebrain nucleus HVC, have premotor activity during singing and auditory activity during playback of the bird's own song. It is not known whether the high degree of selectivity for learned features of song observed during playback arises in HVC or also in structures afferent to HVC. We recorded in anesthetized adult zebra finches from two structures afferent to HVC: either the nucleus interfacialis (NIf) or the L1 subdivision of the field L complex, and simultaneously from a second electrode in HVC. Correlations in the bursting pattern of ongoing activity of HVC

and NIf recordings were observed; these helped to localize the first electrode to NIf recording sites. Most NIf neurons exhibited song-selective responses, but as a population, they were less selective than were HVC neurons. Most L1 neurons were not song-selective. NIf neurons have also been reported to have premotor activity during singing; thus, NIf is another potential site of auditory-motor interactions in the song system. Evidence gathered to date suggests that those brain areas in the passerine forebrain that are recruited during song production also display the most selective learned auditory responses.

Key words: sensorimotor integration; bird song; HVC; NIf; complex sounds; simultaneous recordings; passerine; forebrain

Elucidation of the neural mechanisms of sensorimotor learning, such as vocal learning, requires an understanding of where and how motor and sensory information that are important for the behavior under consideration converge and interact. Beyond human speech, vocal learning may be common in some orders of mammals (Reiss and McCowan, 1993; Esser, 1994; Snowden and Hausberger, 1997; Boughman, 1998) and is well represented in birds (Nottebohm, 1972). Although song learning and the bird song system are extensively studied, surprisingly little is known about the physiological basis for sensory input into the song system. Here, we report that auditory responsiveness to learned features of song emerges gradually within a forebrain sensorimotor hierarchy.

The song system forebrain nucleus HVC is known to be auditory (Katz and Gurney, 1981) and confers song selectivity onto its efferent targets (Doupe and Konishi, 1991; Vicario and Yohay, 1993). Neurons in the HVC are selective for the bird's own song (BOS) when compared with responses to conspecific songs (CON) or BOS played in reverse (REV) (Margoliash, 1993, 1986; Margoliash and Fortune, 1992). This selectivity reflects sensitivity to learned features of song and arises during sensorimotor learning (Volman, 1993; Whaling et al., 1997). The site or sites where such selectivity emerges, however, is not well established.

HVC receives from the thalamic nucleus uvaefornis (Uva), from nucleus interfacialis (NIf), which also receives from Uva (Nottebohm et al., 1982), and possibly from multiple other targets (Margoliash et al., 1994). HVC, NIf, and Uva have singing-related

activity (McCasland and Konishi, 1981; McCasland, 1987; Williams and Vicario, 1993; Yu and Margoliash, 1996). Uva is probably not auditory, as judged by both its apparent homology with the pigeon DLPc and its sources of afferent input (Wild, 1994). The sensory status of NIf has not been studied extensively (Williams, 1989).

Auditory input to HVC may result from direct and indirect input from the thalamorecipient forebrain auditory structure field L (Kelley and Nottebohm, 1979; Fortune and Margoliash, 1995; Vates et al., 1996; Mello et al., 1998). Although some field L neurons have somewhat weaker responses to REV than to BOS, most of them respond comparably well to CON and BOS (Margoliash, 1986; Lewicki and Arthur, 1996). Thus, song-selective responses could, in principle, arise *de novo* within HVC, in a subset of field L neurons that project directly or indirectly to HVC, or in structures afferent to HVC other than field L.

In this report, we investigate the functional properties of NIf neurons. In zebra finches, the medial aspect of NIf comprises a wedge of cells rostral to the L2a subdivision of the field L complex; laterally, NIf also includes a thin plate extending dorsally along the rostral border of L2a (Fortune and Margoliash, 1992). Tracer studies indicate NIf may receive auditory input via the caudolateral ventral hyperstriatum (clHV), which projects to NIf and is reciprocally connected to several subdivisions of the field L complex (Vates et al., 1996). Golgi and retrograde labeling studies further demonstrate dendritic extensions of NIf neurons into the L1 subdivision of the field L complex (Fortune and Margoliash, 1992, 1995); these too might be sources of auditory input to NIf. Given that NIf may provide auditory information to HVC and serve as a site of sensorimotor integration preceding HVC, we decided to investigate whether NIf has song-selective neurons. The difficult morphology of NIf and its location within the field L complex, which is auditory, motivated our search for a better methodological approach than single electrode recordings.

Received Sept. 11, 1998; revised March 25, 1999; accepted March 26, 1999.

This work was supported by National Institutes of Health Grants F32 NS10395 (to P.J.) and RO1 NS25677 (to D.M.). We thank an anonymous reviewer for several helpful criticisms of an earlier version of this manuscript.

Correspondence should be addressed to Dr. Daniel Margoliash, Department of Organismal Biology and Anatomy, 1027 East 57th Street, University of Chicago, Chicago, IL 60637.

Copyright © 1999 Society for Neuroscience 0270-6474/99/195108-11\$05.00/0

MATERIALS AND METHODS

Stimuli. The exemplar of BOS was chosen by examining 50–200 bouts of singing. A recording with a high signal-to-noise ratio that contained the number of motifs (fixed sequences of syllables) typical for that bird (usually two to three) was selected as the exemplar. REV, the songs of several conspecifics, and 250 msec bursts (10 msec on–off ramps) of broadband noise served as the basic stimulus set. Reversed songs of conspecifics were occasionally presented. Stimuli were presented in pseudorandom order (5–40 repetitions). Additionally, tones ranging in 500 Hz steps from 1000 to 6500 Hz (250 msec, 10 msec on–off ramps) were sometimes presented, also in pseudorandom order.

Stimuli were sampled and output with 16-bit resolution at 20 kHz, using eight-pole elliptical anti-aliasing low-pass 10 kHz filters. In the present experiments, we used a Raven R2 speaker (Zalytron Industries, Mineola, NY) situated ~2 m from the bird in a walk-in, double-walled sound isolation chamber (Industrial Acoustics Corp., Bronx, NY). Otherwise, the sound system, free-field sound conditions, and sound calibration have been described previously (Margoliash and Fortune, 1992). The stimuli were presented at a root-mean squared amplitude of 65–70 dB sound pressure level.

Electrophysiological recording. At least 2 d (typically three or more) before the electrophysiological recording session, adult male zebra finches (Magnolia Bird Farms, Anaheim, CA) were deprived of food and water for 1 hr and then anesthetized with an intramuscular injection of 50 μ l of modified Equithesin (0.85 gm of chloral hydrate, 0.21 gm of pentobarbital, 2.2 ml of 100% ethanol, and 8.6 ml of propylene glycol, to a total volume of 20 ml with dH₂O). The birds were immobilized in a stereotactic frame consisting of ear bars and a beak holder that held the head at a 45° angle, the top layer of the skull was removed, and a pin was implanted caudal to the bifurcation of the midsagittal sinus.

On the days of recordings, birds were food and water deprived for 1 hr and anesthetized with three doses of 20% urethane (40, 30, and 30 μ l) administered intramuscularly over a 1 hr period. A bird was placed on a cushion, and the head was immobilized by fastening the implanted pin to a frame. The bottom layer of the skull was removed so as to allow access to both HVC and NIf. Locations of electrode penetrations targeting NIf were initially made stereotactically relative to the bifurcation of the midsagittal sinus and subsequently adjusted based on observed responses to auditory stimuli at different depths of the electrode. For example, L2a, which forms the caudal border of NIf and is the primary telencephalic recipient of auditory afferents from the thalamus, was readily recognized by its robust auditory responses and tonotopy.

NIf is an anatomically complex three-dimensional structure (Nottebohm et al., 1982; Fortune and Margoliash, 1992). To aid in determining electrode position during the experiments, simultaneous recordings were made from an electrode positioned in HVC and the NIf-directed electrode in 9 of the 11 birds studied. The patterns of ongoing activity in NIf, but not L1 dorsal to NIf, showed strong temporal synchrony with ongoing HVC activity (see Results). Such observations were used in part during the experiment to indicate when electrodes were in NIf. However, for analyses, units were assigned to NIf, L1, or the NIf/L1 border based solely on postmortem histology (see below). In several of the initial experiments, electrolytic lesions (5–10 μ A for 5–10 sec) were made directly at the locations at which the auditory responses of single units or multiunit clusters had been characterized. Given the small size and sheet-like geometry of NIf, we were concerned that its normal activity patterns might be disrupted significantly after a lesion anywhere within its boundaries. Therefore, these experiments usually yielded data from only one or two units along a single penetration. After gaining more experience targeting NIf, we usually made several penetrations through NIf and made fiduciary lesions outside of its borders.

The activities of neurons in L1 encountered along penetrations targeted at NIf were recorded to serve as a control comparison for NIf recordings. Previous studies (Margoliash, 1986; Lewicki and Arthur, 1996) have found few, if any, song-selective responses within the field L complex. Because the primary focus of these experiments was to record from NIf, we did not, however, systematically sample responses throughout L1.

NIf-targeted electrodes were made from 25 μ m diameter W/Pt wire encased in 80 μ m quartz fiber (Uwe Thomas Recording, Marburg, Germany). The fibers were pulled on a laser puller (P-2000; Sutter Instrument Co., Novato, CA), and tips were fashioned by grinding on a beveller (BV-10; Sutter Instrument Co.). Final tip diameters were ~5–10 μ m, and impedances ranged from 0.5 to 2.5 M Ω at 1 kHz. HVC recordings were made either with fiber electrodes or with etched 0.003 inch

Pt/Ir wire (A-M Systems Inc., Everett, WA), which was then insulated with solder glass (Corning, Corning, NY).

Signals from the electrodes were amplified, filtered (300 Hz to 5 kHz bandpass; M. Walsh Electronics, San Dimas, CA), digitized (ATMIO16x card; National Instruments, Austin, TX) at 20 kHz with 16-bit resolution, and saved to computer disk for off-line analyses.

Reconstruction of recording sites. After the electrophysiological recordings, the bird was anesthetized deeply with 50 μ l of Nembutal, injected intracardially with 50 μ l of heparin, and exsanguinated with 0.9% saline, followed by 10% formalin. After at least 2 d in 10% formalin, the brain was transferred to a 30% sucrose–10% formalin solution. Parasagittal sections (50 μ m) were prepared by frozen-sectioning on a microtome and stained with cresyl violet using standard procedures.

For each bird, a map of recording site positions relative to anatomical landmarks was drawn using a camera lucida. Recording site locations were specified based on the positions of electrolytic lesions and visible electrode tracks in the Nissl material. Viewing the Nissl material under dark-field conditions aided greatly in determining the borders of NIf. Summary maps were created by expressing the recording locations for each bird in terms of standardized coordinates. Two standardized sections were created. In the more lateral representation, NIf is elongated ventrodorsally in a sheet along the rostral edge of L2a. In the more medial representation, the ventrodorsal aspect of NIf is greatly reduced, whereas the rostrocaudal extent along the dorsal medullary lamina (LMD) is slightly increased (see Fig. 1). Each of the individual camera lucida tracings was assigned to one or the other of the projection planes.

For each projection plane, a standardized coordinate system was established as follows. The junction of the LMD, L2a, and NIf served as the origin. The tracing of each section that contained recording sites in NIf and/or L1 was aligned so that the origins of the standard and the individual section overlapped and the border of L2a and NIf ran along the 90° line (y-axis). The paths of the lamina hyperstriatica (LH) and the border of NIf were then traced onto the standardized coordinate frame. The visually estimated average of the many individual NIf border and LH path tracings served as the “standard” NIf border and LH in the schematic diagrams. The standardized coordinate system was then traced onto the drawings of individual sections. The position of each recording location was described by the angle and distance from origin. The distances from the origin to the edge of NIf and LH along the line intersecting the recording site were also noted for each section so that recording site positions could be specified relative to these anatomical landmarks. For NIf locations, the location in the standardized coordinate frame was determined by computing the ratio of the distance to the recording site and the edge of NIf, and multiplying this value by the distance to the standard NIf edge. L1 positions were determined in the same way, except that the distance to LH was used rather than the distance to the edge of NIf. For sites identified as lying along the NIf/L1 border, ratios were computed using the edge of NIf.

Data analysis. Recordings that appeared to represent the activity of one to three single units, as based on visual inspection, were processed with a spike-sorting algorithm (Lewicki, 1994) to separate the times of spike events for each unit. Peristimulus time histograms (PSTHs) were created by assigning spike times to 10 msec bins and averaging the counts in each bin across multiple repetitions of the stimulus. Unless stated otherwise, the activity of single neurons is expressed as the instantaneous rate of firing (spikes per second) minus the baseline firing rate. The baseline firing rate was estimated as the average firing rate in the 1 sec preceding stimulus onset. The strength of the response to any given stimulus was expressed as the mean instantaneous firing rate over the course of the stimulus.

Recordings in which the activity of multiple neurons was evident but which could not be spike-sorted reliably were treated as multiunit activity. In these cases, the raw waveforms were full-wave rectified and averaged over the stimulus repetitions into nonoverlapping 10 msec bins. The mean of the full-wave rectified signal from 1 sec preceding stimulus onset was subtracted from each bin to give a baseline-corrected instantaneous firing rate. The stimulus response strength was expressed as the mean instantaneous firing rate over the course of the stimulus. This treatment of the data follows previously established procedures (Margoliash, 1986).

A unit was classified as auditory if the mean activity level during any stimulus in the stimulus set was significantly different (either excited or inhibited) from baseline activity (paired *t* test; *p* < 0.05 level). Mean firing rates elicited by the different classes of stimuli were compared with two-tailed paired *t* tests. When multiple exemplars of CON were used to

probe the responses of a neuron, the average response to CON was used in these *t* tests.

We adopted two of the several measures of song selectivity that have been used in previous studies (Margoliash, 1986; Margoliash et al., 1994; Volman, 1996; Solis and Doupe, 1997; Theunissen and Doupe, 1998). First, we expressed selectivity for BOS as a simple ratio of the response to the comparison stimulus, e.g., CON, and the response to BOS (Margoliash, 1986; Margoliash and Fortune, 1992; Margoliash et al., 1994). When the response to BOS is excitatory and greater than the response to the comparison stimulus (as is almost always the case in HVC in anesthetized birds; Margoliash, 1986), the ratio assumes values less than one. If the response to the comparison stimulus is greater, the value of the ratio is greater than one. Difficulties in interpretation may arise with the ratio measure if the response value to BOS or both stimuli is negative (inhibitory response).

The selectivity of any given neuron for BOS relative to other stimuli was also expressed as a *d'* value, defined as

$$\frac{2(\bar{A} - \bar{B})}{\sqrt{\sigma_A^2 + \sigma_B^2}}$$

where *A* refers to BOS and *B* refers to the comparison stimulus (Theunissen and Doupe, 1998). The *d'* metric differs from the ratio metric in two fundamental aspects. First, if *A* > *B*, the value of *d'* will be positive, regardless of whether the actual response to BOS was excitatory or inhibitory. The more positive *d'* is, the more song-selective the response of a neuron is said to be. In the case of inhibitory responses, however, a more strongly inhibited response to BOS will result in a more negative *d'* value. A second important property of the *d'* metric is that decreased variability in response strength across trials increases the *d'* value.

Differences in song selectivity of NIf and L1 were assessed with *t* tests comparing the distributions of simple ratios and distributions of *d'* values observed in each brain region for both the REV:BOS and CON:BOS comparisons. When responses of a unit to multiple (two to three) different exemplars of CON were collected, the song selectivity metric values were averaged together across exemplars before being entered into the distribution. The distribution of HVC single-unit responses to BOS, REV, and CON from an earlier data set (Margoliash et al., 1994) were compared with the NIf response distribution. The HVC data in the previous experiments were collected using glass-coated Pt/Ir electrodes also in urethane-anesthetized zebra finches.

In eight of nine birds for which NIf recordings were obtained, simultaneous recordings were obtained from NIf (16 single units, 14 sites) and HVC single or multiunits. In seven birds, simultaneous recordings were also obtained from L1 (19 single units, 14 sites) and HVC single or multiunit activity. To compare the temporal structure in the activity of a NIf or L1 site with that at an HVC site, the single-trial histogram and/or rectified waveforms constructed for each trial were cross-correlated. The bin size of the single-trial histograms and rectified waveforms was 1 msec. To remove effects of different mean firing rates, the mean firing rate of each channel in the time interval to be cross-correlated was subtracted from the single-trial histograms before the cross-correlation. The cross-correlations were normalized by the product of the auto-correlations on each channel at a time-lag of 0 msec. The single-trial cross-correlograms were then averaged. Cross-correlations were computed for lags of up to 500 msec but are only plotted for ± 50 msec for lack of any side peaks at longer lags. Estimates of the cross-correlation in the ongoing activity were derived by pooling the cross-correlations for all epochs of 1 sec preceding stimulus presentations.

The observed cross-correlograms were compared with “shuffled” cross-correlograms to determine the significance of peaks in the observed cross-correlations, as well as the contribution of stimulus-locked responses to the cross-correlations. Shuffled cross-correlograms were constructed by pseudorandomly pairing single-trial histograms from different repetitions of the stimulus as inputs to the cross-correlation function. Shuffled cross-correlograms of ongoing activity were similarly computed by pseudorandomly pairing all prestimulus epochs. No pairing was used twice, and the actual pairings (from the experiments) were excluded from the shuffled data set. Thus, a maximum number of permutations of $N(N - 1)$ was calculated (where *N* is the number of repetitions), with an upper limit (for large *N*) of 500.

For each 1 msec bin (τ) between -50 and 50 msec, we tested whether the observed cross-correlation value was significantly larger than the corresponding shuffled cross-correlation value using a one-tailed *t* test. The α criterion in each *t* test was adjusted using the Bonferroni correction for multiple comparisons (the total number of time bins).

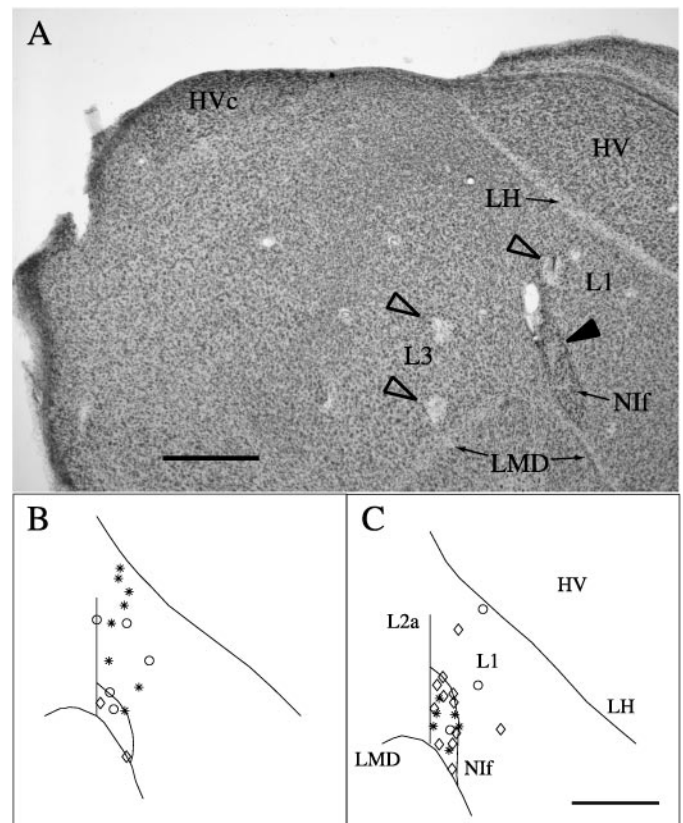


Figure 1. Anatomical locations of recording sites. *A*, Cresyl violet-stained material in a parasagittal section showing sites of electrolytic lesions in L3, L1 (open arrowheads), and NIf (filled arrowhead). NIf is the darkly staining oval nucleus at the bottom right. HVC is at the top left. The lesion in NIf marks the site from which the data illustrated in Figure 2, *E* and *F*, were acquired. Scale bar, 500 μ m. *B*, A map of NIf and L1 single-unit recording sites on a standardized schematic of a parasagittal section through the medial aspect of NIf. Asterisks denote recording sites that were either 0–50 μ m from a lesion site or along an electrode track with two lesions visible in the same section. Circles denote recording sites along electrode tracks marked with a single lesion and other indication of the orientation of the track, or along electrode tracks without lesions in the same section as a parallel track marked by two lesions. Sites along tracks in sections not possessing lesions, but adjacent to sections with lesions, are marked by diamonds. *C*, A map of NIf and L1 single-unit recording sites on a standardized schematic of a parasagittal section through the lateral aspect of NIf. Same symbols as in *B*. Scale bar (in *C*): *B*, *C*, 500 μ m.

RESULTS

Responses to complex auditory stimuli

Auditory responses were recorded from 17 single neurons at 15 sites in the NIf of nine birds (Fig. 1). For the purpose of comparison, the activity of 26 single neurons at 19 locations in the L1 region of the field L complex of 10 birds was also recorded. Additionally, in five birds, eight neurons at seven locations were identified as falling on the NIf/L1 border. Because the primary goal of these experiments was to assess auditory responses in NIf, we searched exclusively for units showing an auditory response. Only those sites that showed a statistically significant response ($p < 0.05$) to some of the probe stimuli (e.g., BOS, reverse BOS, or a variety of conspecific songs) compared with a 1 sec prestimulus baseline were included in the statistical comparisons among stimuli. Of the 17 neurons that were localized to NIf based on inspection of the histological material, 16 were auditory.

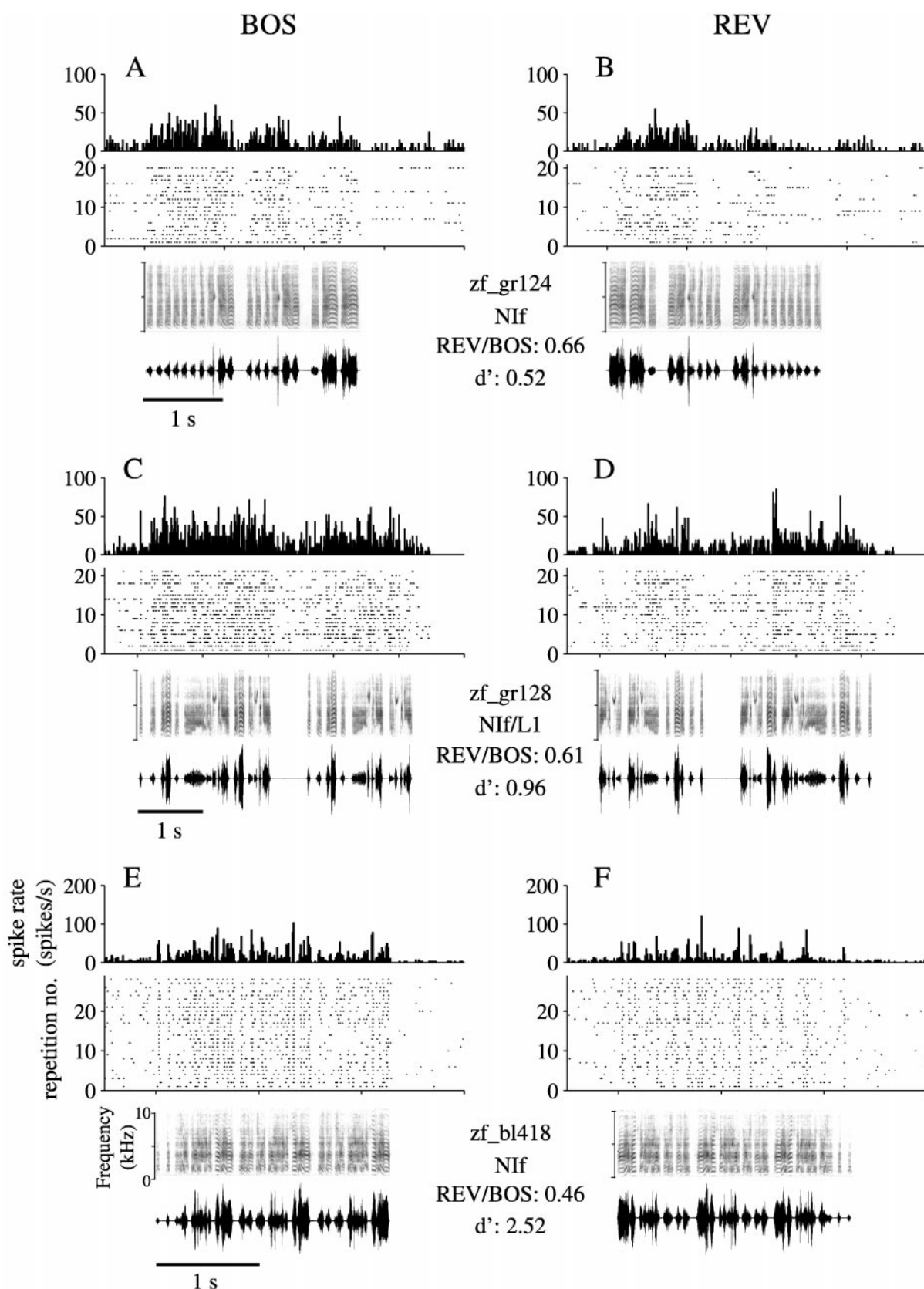


Figure 2. Responses of NIf single units to complex auditory stimuli. The activity patterns from three birds of two single NIf neurons and one neuron on the border of NIf and L1 are illustrated in response to playback of the BOS (left column) and REV (right column). For each unit, a PSTH is shown above a raster plot, which indicates the times of spike occurrences arranged as dots in horizontal rows. The PSTHs and rasters are aligned with the spectrographs and oscillographs shown below the raster. Two values of two song selectivity metrics, the REV/BOS ratio and d' , are indicated below each bird's name between the BOS and REV columns. The recordings for bird zf_bl418 (E, F) were acquired from the site of the NIf lesion shown in Figure 1.

Table 1. Summary of selectivity for BOS over CON in areas L1 and NIf

#CON exemplars	L1			NIf		
	1	2	3	1	2	3
BOS > CON for all CON	2			6	1	5
BOS > CON for many CON			1 (2)		1 (1)	1 (2)
BOS \approx CON	9		2 (1)			
			5 (0)			

A tally of the responses of units in L1 and NIf to BOS and CON. Units were tested with one, two, or three exemplars of CON, as indicated by the different columns. BOS > CON for all CON represents the number of units whose responses were significantly stronger (two-tailed t test; $p < 0.05$) to BOS than to CON for every exemplar of CON tested. For BOS > CON for many CON and BOS \approx CON, the numbers in parentheses indicate the number of CON exemplars for which the BOS response was stronger. Note that the preponderance of L1 cells are in BOS \approx CON, whereas the NIf cells are mostly in BOS > CON for all CON.

Auditory responses of two representative NIf neurons and one NIf/L1 neuron from three birds are illustrated in Figure 2. The location of the neuron whose data are shown in Figure 2, *E* and *F*, is marked by an electrolytic lesion shown in Figure 1*A*. The *left column* in Figure 2 illustrates responses to BOS, and the *right column* shows responses to REV. The density of dots in the raster plots indicates that responses to REV were weaker than responses to BOS. The d' measures and the REV/BOS ratios comparison are shown for each unit. By both the ratio and d' criteria, the unit shown in Figure 2, *E* and *F*, is the most selective of the three.

Most (12 of 16) NIf neurons showed tonically elevated firing rates throughout the stimulus, whereas the remaining neurons showed primarily phasic activity. Generally, phasic responses were emitted throughout the stimulus. Of the 12 neurons that exhibited tonic activity during the stimulus, six showed distinct phasic peaks in addition to the tonic activity. In response to REV, the phasic components of the response that were present during BOS were often absent. Responses to REV also tended to be stronger at the beginning of the stimulus. Responses to tone bursts were assayed in a few of the initial experiments. Because NIf neurons were generally unresponsive to tones (cf. Williams, 1989) and no frequency tuning was observed, the stimulus set was focussed on the more complex stimuli.

Across all NIf neurons, the mean spike rate in response to BOS was significantly higher than the mean spike rate in response to REV (9.77 and 4.61 spikes/sec, respectively; $df = 15$; $t = 4.09$; $p < 0.001$). Individually, 10 of 16 neurons showed significantly larger responses to BOS than to REV (two-tailed t test; $p < 0.05$). The population of NIf neurons responded more strongly to BOS than to CON (4.79 spikes/sec for CON; $df = 13$; $t = 5.28$; $p < 0.0002$). Individually, 12 of 14 neurons responded significantly more strongly to BOS than to all presented exemplars of CON, and the remaining two responded more strongly to BOS than to one of two, and two of three, of the CON exemplars tested (Table 1). The enhanced response to BOS relative to CON is particularly important in that this difference is likely to result from the song learning experience (see Discussion).

In contrast to NIf neurons, there was no statistically significant difference in mean firing levels in response to BOS (6.69 spikes/sec) and REV (6.06 spikes/sec) for L1 neurons ($df = 23$; $t = 0.62$; NS). Individually, 5 of 24 L1 neurons responded significantly more strongly to BOS than to REV. Mean firing levels were similar for BOS and CON (7.44 spikes/sec) ($df = 18$; $t = 0.08$; NS). Of 19 L1 neurons, only three showed a significant degree of BOS selectivity (Table 1). Responses of L1 neurons were some-

what heterogeneous. In two L1 neurons, the mean firing rate during stimulus playback did not reach the criterion established for a significant response. In one of these cases, the temporal structure of the activity changed during stimulus playback, although the overall firing rate did not increase. In 27% (7 of 26) of the sites, responses were inhibitory to at least one of the stimuli, with the inhibition lasting throughout the stimulus. In 65% (17 of 26) of the neurons, responses to auditory stimuli were excitatory.

Song selectivity of NIf, L1, and HVC neurons

The distributions of song selectivity, expressed as the response to either REV or CON divided by the response to BOS (the ratio measure), were compared for HVC, NIf, and L1 single neurons (Fig. 3). For the purpose of comparison, the selectivities of a large sample of HVC neurons that were recorded in our laboratory previously using the same recording techniques (Margoliash and Fortune, 1992; Margoliash et al., 1994) are plotted above the distributions of selectivities for NIf and L1 neurons. HVC neurons were the most song-selective, with a mean REV/BOS ratio of 0.17 and a mean CON/BOS ratio of 0.21. The mean REV/BOS ratio of NIf neurons was 0.64, and the mean CON/BOS ratio was 0.54. The NIf and HVC distributions were significantly different from each other for both REV (two-tailed t test; $df_{NIf} = 15$; $df_{HVC} = 98$; $p < 0.0005$) and CON (two-tailed t test; $df_{NIf} = 13$; $df_{HVC} = 90$; $p < 0.04$).

Song selectivity, as assessed by the ratio measure, varied most in L1. The *bottom left panel* in Figure 3 shows that the REV/BOS ratios of several neurons were ≤ 0 , indicating a high degree of song selectivity. In five cases, such ratios were attributable to neurons that showed inhibition during presentation of BOS and weak to moderate excitation during presentation of REV. In another case, the response to BOS was excitatory and the response to REV was inhibitory. In two cases, the responses to BOS were extremely weak and did not reach the criterion for an auditory response. These units were included in the parametric comparisons of mean firing rates described above because either the response to REV or CON met the auditory response criterion. However, the song selectivity ratios were grossly inflated (>10) in these cases, and the data were removed as outliers before comparing the NIf and L1 ratio distributions. Overall, the mean REV/BOS ratio for L1 was 0.60. This value was not significantly different from that observed in NIf (two-tailed t test; $df_{NIf} = 15$; $df_{L1} = 22$; NS). The mean CON/BOS ratio for L1 was 1.07, which showed a trend toward being significantly different from the mean of NIf (two-tailed t test; $df_{NIf} = 13$; $df_{L1} = 16$; $p < 0.06$). These data also demonstrate the limitations of the ratio measure when applied to responses that are not dominated by an excitatory response to BOS.

The HVC recordings in the present study were primarily multiunit in nature. Relatively few sites were sampled given that the same HVC recording location was maintained while responses at multiple NIf and L1 sites were assayed. Although we did not try to sample HVC systematically in the present experiments, when single-unit HVC recordings were obtained, the REV/BOS and CON/BOS ratios were strongly biased toward BOS (REV/BOS = 0.16; $n = 13$ observations, 11 units, 11 sites, 5 birds; CON/BOS = 0.30; $n = 19$ observations, 10 units, 10 sites, 4 birds). This trend is consistent with previous results showing a selectivity for BOS in zebra finch HVC neurons (Margoliash and Fortune, 1992; Margoliash et al., 1994).

Song selectivity was also described in terms the distributions of d' scores for NIf and L1 single units (Fig. 4). (For the d' measure,

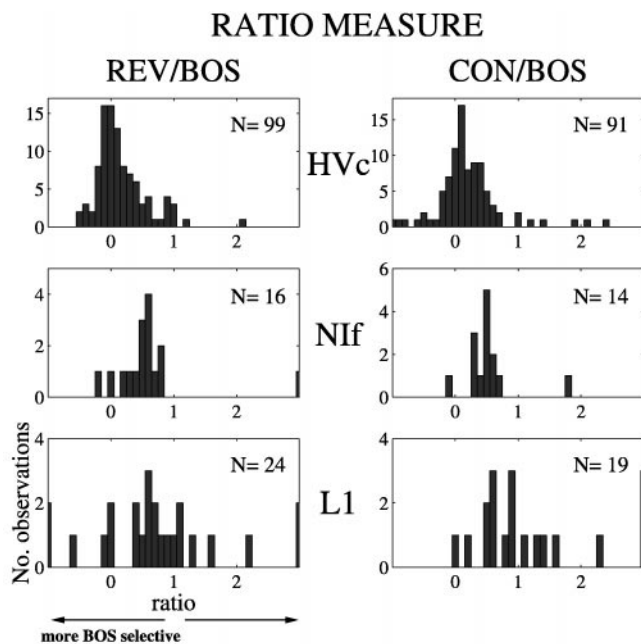


Figure 3. Comparison of song selectivity of HVC, NIf, and L1 single units. Song selectivity was quantified by a ratio expressed as the mean firing rate in response to either REV or CON divided by the mean firing rate in response to BOS. A value of 1 indicates an equally strong response to both stimuli. Values <1 indicate a stronger response to BOS, whereas values >1 indicate a stronger response to the stimulus being compared. Data for HVC are the same as those published in Figure 3b of Margoliash et al. (1994). The number of units contributing to each histogram is indicated in the top right corner of each plot. In the case of CON/BOS, multiple exemplars of CON may have been presented while recording any given unit. Thus, each entry in the histogram reflects that average response to CON for each unit.

the stronger the response is to BOS, the more positive is the d' value; see Materials and Methods.) The distributions for NIf neurons (Fig. 4, top panels) are displaced toward positive values when BOS is compared with REV (mean of 1.39) and CON (mean of 1.50). In contrast, the distributions of L1 neuronal responses (Fig. 4, bottom panels) are centered around zero for both REV (mean of -0.17) and CON (mean of -0.23). The distributions of d' scores are significantly different between NIf and L1 sites for BOS selectivity with respect to both REV ($p < 0.0002$) and CON ($p < 0.0002$). The compelling statistical significance obtained for these comparisons using the d' measure compared with marginal or no significance using the ratio measure highlights the greater power of the d' measure in detecting a reliable difference. Thus, the d' measure is particularly valuable when applied to heterogeneous responses of field L neurons to songs.

Neurons on the NIf/L1 border

As a group, the neurons localized to the border of NIf and L1 did not show a significant difference in the comparisons of mean firing rates of REV to BOS, and CON to BOS (9.69, 7.57, 9.83 spikes/sec, for BOS, REV, CON, respectively). The REV/BOS ratio was 0.70, whereas the CON/BOS ratio was 0.89. The mean d' scores were 0.98 and 0.39 for REV and CON, respectively. Only 4 of 12 observations of CON had a CON/BOS d' value >1.0 , and only 2 of 12 had a CON/BOS ratio <0.5 .

The lack of BOS selectivity of the population of “border” neurons distinguishes them from NIf neurons and resembles the

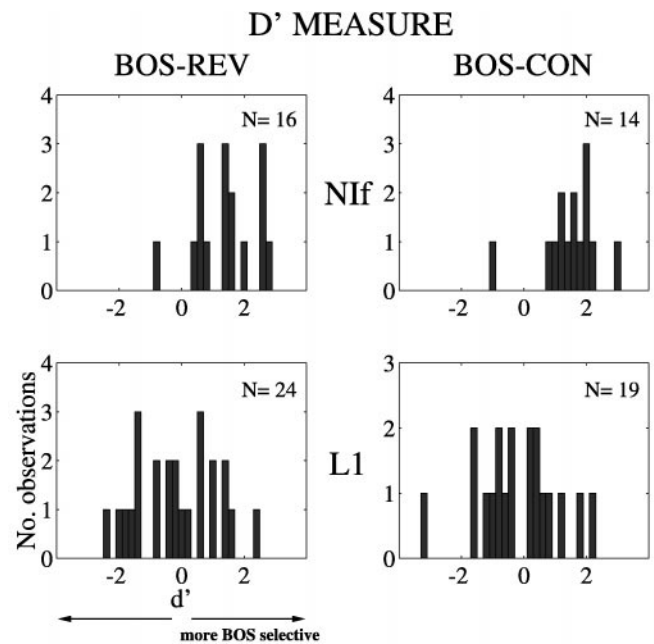


Figure 4. Song selectivity of NIf and L1 neurons as measured using an index of discriminability, d' (see Materials and Methods). The left panels depict comparisons of REV and BOS responses, and the right panels show comparisons of CON and BOS responses. A d' value of zero indicates no difference between responses to the two stimuli, whereas a positive d' value indicates a stronger response to BOS. The distributions for L1 neurons are centered around zero but are shifted to more positive values for both REV and CON in the case of NIf neurons, indicating selectivity of NIf but not L1 neurons to BOS.

lack of selectivity found for field L neurons. Nevertheless it is difficult to conclude that the group of border neurons did not include some or many neurons from NIf, because a morphologically distinct class of NIf neurons is distributed along the rostral border of NIf (Fortune and Margoliash, 1995), and the physiological properties of this class of neurons have yet to be identified.

Correlation between NIf and HVC activity

A characteristic feature of NIf single-unit recording sites and the multiunit activity around them was a correlation of the NIf activity with the activity recorded simultaneously in HVC. As electrodes approached NIf through L1, the ongoing multiunit activity would switch from showing no apparent correlation with the activity on the HVC electrode to showing approximately synchronous transient bursts of activity on both electrodes. Such correlated activity would persist across several hundred micrometers (depending on the entry point into NIf) and disappear again after reaching the ventral or ventrocaudal border of NIf. The pattern of correlated activity was easy to observe audiovisually on-line and proved to be of considerable utility in targeting NIf. This was particularly important given the unusual morphology of NIf.

When viewing traces of NIf and HVC activity, the correlation of activity was visually most apparent on long or intermediate scales of time. For example, Figure 5 illustrates the correlation in ongoing activity of NIf and HVC on three time scales. Note the occurrence of brief bursts of activity that co-occur on the two channels. The most expanded view (Fig. 5C) shows that the bursts do not occur simultaneously but rather that the HVC activity tends to lag behind the NIf activity by a few milliseconds.

To quantify this phenomenon, we characterized the joint ac-

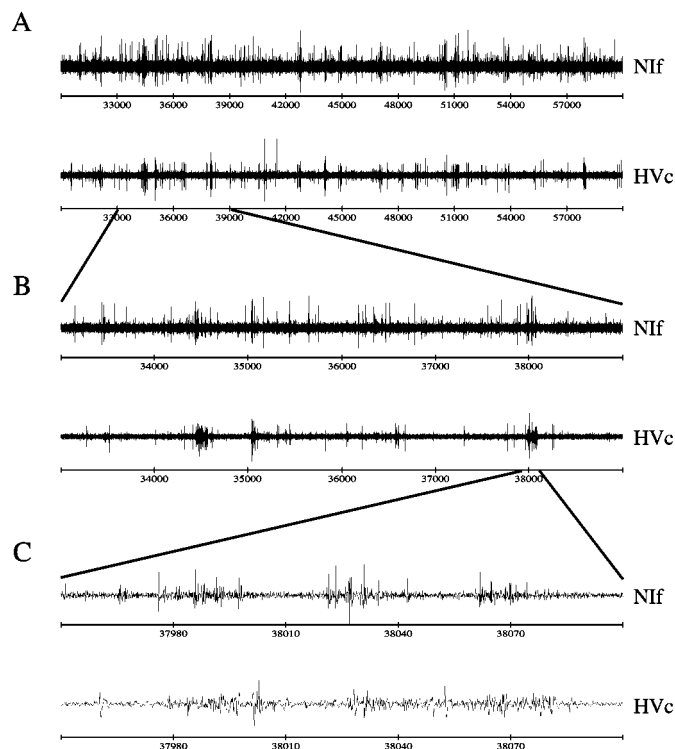


Figure 5. Correlated firing in NIf and HVC ongoing activity. *A*, Thirty seconds of simultaneously recorded ongoing multiunit activity in NIf and HVC. Note the appearance of correlated bursts of activity. *B*, A closer view of a 6 sec excerpt from *A*. *C*, A 150 msec detail of the correlated burst on the right in *B*. The HVC activity lags slightly behind the NIf activity.

tivity in recordings from NIf and HVC, or field L and HVC, using cross-correlations. Average cross-correlograms of HVC multiunit activity with two NIf and two L1 single units are shown in Figure 6. For each unit, the cross-correlations of both ongoing activity and activity during playback of BOS are shown. The *thin lines* indicate the observed cross-correlations, and the *thick lines* are shuffled cross-correlations (see Materials and Methods). The peak in the cross-correlogram of ongoing activity for the NIf unit from bird zf_bl411 was rather broad (± 10 msec width at half-maximum) and centered around zero, indicating that, on average, action potentials in this neuron occurred at the same time as bursts of activity at the HVC recording site. In contrast, NIf unit from another bird tended to lead HVC activity by ~ 5 msec. The two L1 units shown tended to lead HVC activity slightly in the ongoing activity. As expected, the shuffled cross-correlograms of 1 sec periods of ongoing activity preceding stimuli had no peaks.

The correlations in ongoing activity at short (< 10 msec) lags that were observed for both the NIf and L1 units persisted during playback of BOS. Relatively symmetrical negative side bands around the central peak in the cross-correlograms appeared during BOS playback in three of the units shown (Fig. 6). In total, negative side bands were observed in two L1 units and in six NIf units.

Although both field L and NIf neurons could exhibit correlated activity with HVC, such correlations were prevalent in only the NIf data set. Of the 16 NIf neurons recorded simultaneously with HVC, 14 showed a significant correlation ($p < 0.0005$ after Bonferroni correction) in the ongoing activity for at least one value of τ . In three of these cases, the significance of the cross-

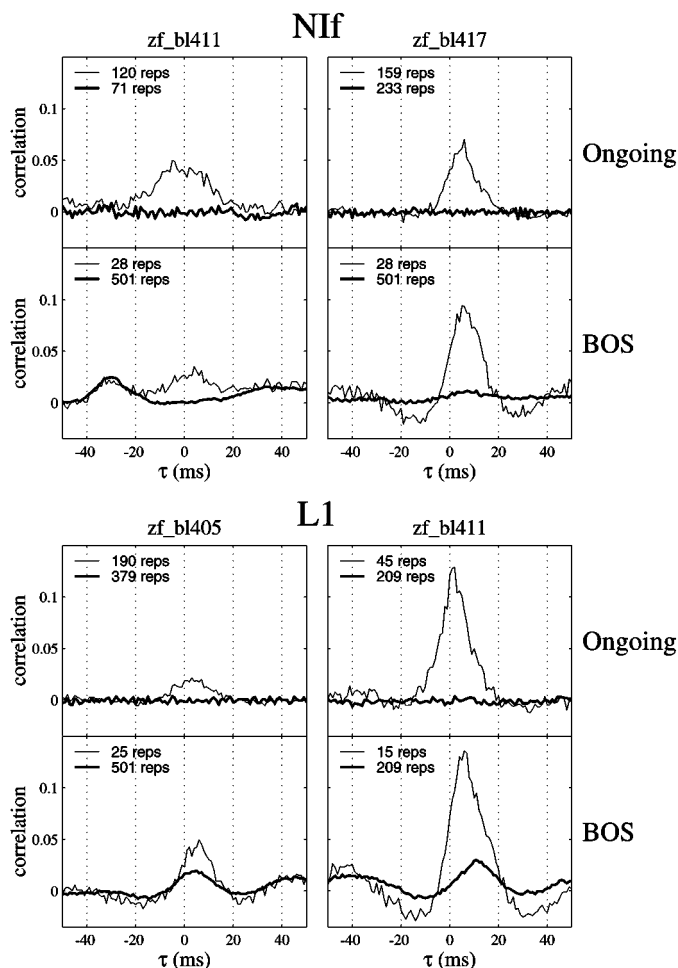


Figure 6. Cross-correlograms of single-unit activity in NIf or L1 and multiunit activity in HVC. In each panel, the *thin line* indicates the measured cross-correlation, and the *thick line* indicates the shuffled cross-correlation that results from cross-correlating epochs of activity recorded at different times on each channel (see Materials and Methods). For each unit in each bird, the *top panel* of each pair of panels shows the cross-correlation in the ongoing activity in the 1 sec epochs preceding stimuli. The *bottom panel* shows the cross-correlation in the activity recorded during playback of BOS. The peaks in the shuffled cross-correlograms computed for BOS playback epochs indicate constant timing relationships between activity on the two channels across multiple presentations of BOS. Negative values of τ indicate that HVC activity leads either the NIf or L1 activity.

correlograms did not appear to be very robust, with only three or fewer time lags reaching significance (Fig. 7, unit numbers 4, 10, and 13), whereas in the other cases numerous lags (typically clustered close together) were significant. The cross-correlograms of seven of the units were centered around zero, whereas six of them were shifted toward positive (NIf leading) lags. Of the original 19 L1 units that were recorded simultaneously with HVC, four were excluded from the analysis because the extremely low ongoing firing rates precluded computation of cross-correlograms for the “ongoing activity” epochs. Of the remaining 15 units, seven showed a significant peak in the cross-correlogram, although only three showed significant peaks at more than three time lags.

Significant peaks that were present in the ongoing activity disappeared during playback of BOS in five NIf units (Fig. 7, unit numbers 2, 3, 4, 13, and 16). The three L1 units that were most

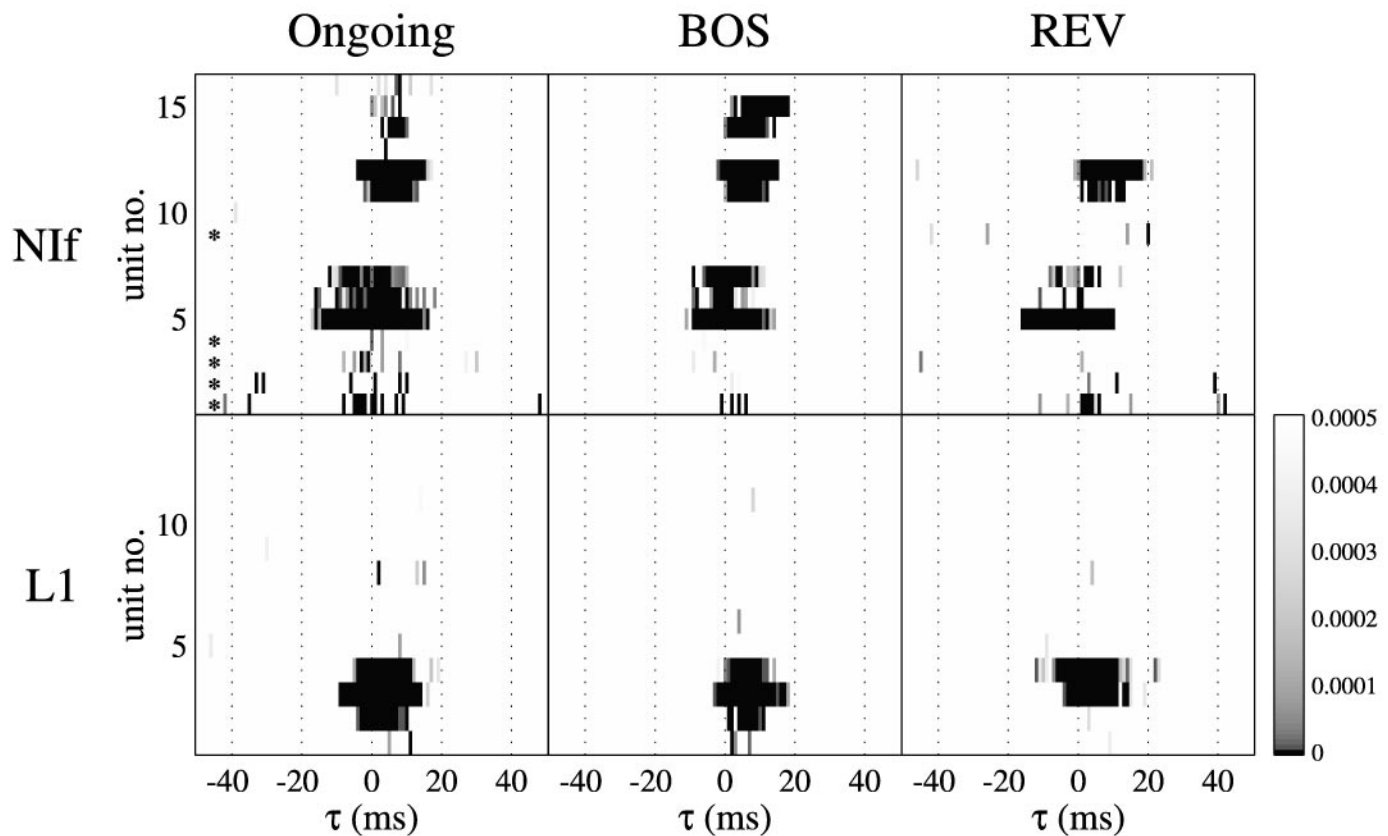


Figure 7. Summary of cross-correlations observed between NIf or L1 single units and HVC activity. Each row in a panel corresponds to the cross-correlogram for one unit. Within each row, a vertical band indicates that the cross-correlation value at that particular τ was significantly greater in the cross-correlogram of the observed activity than in the shuffled activity cross-correlogram (see Materials and Methods). Darker shades of gray indicate a more highly statistically significant result (lower p value). The values on the color bar range from $p = 0$ to the Bonferroni-corrected significance criterion cutoff value of $p = 0.0005$. The asterisks in the top left panel mark rows that represent cross-correlograms between NIf single units and HVC single units. For NIf, units numbers 7 and 12 correspond to cross-correlograms shown in Figure 6 for zf_bl411 and zf_bl417, respectively. For L1, units numbers 3 and 4 correspond to zf_bl405 and zf_bl411, respectively.

significantly correlated with HVC ongoing activity remained so. The four previously excluded L1 units exhibited sufficient numbers of spikes to be analyzed for significant correlations during BOS but failed to show any. During playback of REV, the significant correlations observed for two NIf units (unit numbers 14 and 15) disappeared completely and for three others (unit numbers 6, 7, and 11) appeared to be weakened. The correlation disappeared for one of the L1 units (unit number 2) during REV.

We noted that some of the energy in the cross-correlograms during the playback of BOS was attributable to stimulus-locked activity. For example, the NIf unit from bird zf_bl411 showed an off-center peak at approximately -30 msec (HVC leading) in the cross-correlogram (Fig. 6). The peak at -30 msec could be explained entirely by activity that was locked to the onset of the stimulus because shuffling of the repetitions on each channel before performing the cross-correlation did not affect the amplitude of the peak. Stimulus-locked responses also accounted for some of the features in the cross-correlograms of L1 neurons (Fig. 6). All three of the L1 units that showed significant large peaks in the cross-correlograms also had a peak in the shuffled cross-correlograms within the same range of lags as the peak in the observed cross-correlogram.

Stimulus-locked activity was more prevalent in field L than in NIf. To quantify the degree of stimulus-locked activity, the magnitudes of the peaks relative to the baseline in the shuffled

cross-correlograms were quantified as the number of SDs from the median value across all lags to the maximum value of the peak. For example, with regard to the data shown in Figure 6, these values were 2.7 for the L1 neuron from bird zf_bl405, 2.5 for the L1 neuron from bird zf_bl411, and 2.8 for the third L1 neuron (data not shown). In contrast, of the seven NIf neurons that showed peaks comprised of numerous significant lags (Fig. 7, NIf unit numbers 5, 6, 7, 11, 12, 14, and 15), only one showed a peak in the shuffled cross-correlogram that was >2.5 SDs away from baseline, and only two others showed SDs >2.0 . These measurements agreed with the subjective impression that the shuffled cross-correlograms indicated little stimulus-locked activity in the NIf/HVC correlation.

The degree of correlation of NIf and HVC ongoing activity was not significantly correlated with song selectivity ($r = 0.30$; NS) as assayed by comparing CON and BOS with the d' measure. The lack of a relationship was because of the fact that most NIf neurons were song-selective, yet they displayed a broad range of cross-correlations. The relationship of the correlation between L1 and HVC activity and song selectivity of L1 neurons was slightly stronger but insignificant ($r = 0.40$; NS). This relationship was dominated by two of the three field L neurons, which showed a high degree of correlation and were also song-selective. Potentially, these field L neurons could be the class of neurons tenta-

tively identified as projecting to HVC (Fortune and Margoliash, 1995).

DISCUSSION

We have shown that, in adult male zebra finches, NIf neurons respond to complex auditory stimuli and show a preference for the bird's own song. As with other oscine passerines, song acquisition in zebra finches involves obligatory sensory acquisition and sensorimotor learning phases (Immelmann, 1969; Price, 1979; Eales, 1985). During sensorimotor learning, motor activity controlling singing is honed based on comparisons of auditory feedback from vocal production with an acquired sensory memory of tutor songs, within the bounds of species-specific constraints on the phonology and sequencing of song elements (Konishi, 1978; Marler, 1997). The physiological nature of the representation(s) of the memory and the mechanisms for the comparison are not well established, but it has been shown that sensorimotor learning sculpts response properties of auditory neurons in the HVC of developing birds (Volman, 1993). Ultimately, this process results in a sensory representation for learned acoustic features of BOS in the adult bird (Margoliash, 1983, 1986).

A sensorimotor hierarchy

Earlier attempts to determine whether song selectivity first emerged within HVC or arose at an earlier stage of the ascending auditory pathway focussed on the responses of field L neurons and comparisons with HVC neurons (Leppelsack, 1983; Margoliash, 1986; Lewicki and Arthur, 1996). Coupled with the presumption that the "shelf" ventral to HVC (which receives input from field L subdivisions) is a source of auditory input to HVC, the general failure of the earlier studies to find song selectivity in the field L complex upheld the notion that song selectivity arises suddenly within HVC. This conclusion was troubling, however, because hierarchical organization of sensory systems and gradual emergence of highly selective responses is a well established principle in a number of other systems (Suga, 1989; Takahashi, 1989; Heiligenberg, 1991; Tanaka, 1996). In contrast, we have found that NIf neurons are both auditory and song-selective, although, on average, their selectivity is not as strong as that observed for HVC neurons. Because NIf neurons also exhibit premotor activity (McCasland, 1987), song selectivity can now be viewed in the context of a sensorimotor hierarchy.

Neurons in the L1 and L3 subdivisions of the field L complex may respond selectively to complex sounds or species-typical calls (Leppelsack and Vogt, 1976; Scheich et al., 1979; Langner et al., 1981; Muller and Leppelsack, 1985), although they typically do not exhibit BOS-selective responses. Our finding that relatively few (5 of 24) L1 neurons responded more strongly to BOS than to REV is similar to the incidence reported by Lewicki and Arthur (1996). BOS selectivity has been observed only in premotor structures that are known to be or are thought to be recruited during singing (McCasland and Konishi, 1981; Doupe and Konishi, 1991; Vates et al., 1997). Thus, the emergence of selectivity for BOS can be thought of in terms of a hierarchy of "sensorimotor song filters," in which the highly selective sensory responses are coupled to a specific behavioral context and associated motor programs.

Auditory input to the song system

To understand how song selectivity emerges, it is necessary to identify the sources of auditory input to the song system. It is

commonly assumed (Vates et al., 1996) that HVC dendrites that extend beyond the ventral border of HVC receive auditory input from the shelf, which in turn receives from field L (Kelley and Nottebohm, 1979). The shelf is a relatively cell-free zone through which course fibers from many structures, however, and it has yet to be established that HVC dendrites contact axons of field L neurons or other auditory neurons in the shelf. Small injections of biocytin into the shelf suggest a sparse projection from the shelf into HVC (Mello et al., 1998). Injections of retrograde tracers into HVC label field L neurons, suggesting a direct projection of field L onto HVC (Fortune and Margoliash, 1995). In all these examples, the putative auditory projection to HVC was relatively sparse, which complicates the interpretation. The interpretation was further aggravated in cases that suffered a difficulty in interpretation of control injections (for example, because of fibers-of-passage).

In contrast, the projection of NIf is unidirectional, directly into HVC and robust (Nottebohm et al., 1982; Fortune and Margoliash, 1995), so our finding of BOS-selective auditory responses in NIf unambiguously demonstrates appropriately selective auditory input into HVC proper. Thus, the strong song selectivity of HVC need not arise *de novo* from unselective inputs but could derive, in part, from local processing that combines input from NIf, which is moderately strong in its song selectivity, with other auditory inputs (whose song selectivity has yet to be determined). At least one other source of input to HVC, the medial magnocellular nucleus of the anterior neostriatum (mMAN), is probably auditory and song-selective, but this may be a form of feedback because auditory activity in mMAN probably depends on auditory activity in HVC (Vates et al., 1997).

There are two likely sources of auditory input to NIf. One arises from somata located in lateral regions of caudal HV whose axons terminate within NIf. Iontophoretic injections of biotinylated dextran amine (BDA) into NIf result in retrogradely labeled cells in cHV, and BDA injections into cHV label fibers in NIf. Lateral aspects of the caudal HV also send projections to the shelf caudal to HVC and make reciprocal connections with the field L complex (Vates et al., 1996). Caudal HV neurons in starlings also respond to auditory stimuli (Muller and Leppelsack, 1985; Capsius and Leppelsack, 1996). Indeed, on several approaches to NIf, we found caudal HV neurons to respond to BOS, REV, and CON (data not shown). Our data set of HV neurons is not sufficient, however, to assess whether a substantial population of HV neurons shows any song selectivity in its responses. Additionally, it is premature to conclude that those HV neurons that are auditory also project to NIf.

Second, in a manner analogous to HVC receiving auditory information from the shelf area, NIf may receive auditory input via dendritic arborizations within the adjoining L1 subdivision of the field L complex (Fortune and Margoliash, 1992, 1995). Within NIf, two classes of neurons have been observed, both of which project to HVC. Type 5 cells are large and oblong with thick proximal dendrites lacking spines and medium dendrites with spines more distally (Fortune and Margoliash, 1992). Cells in the other class have fusiform somata and are found along the entire rostral border of NIf; these cells have dendrites extending into L1 (Fortune and Margoliash, 1995). These dendrites could receive input from L1 neurons or from other areas that send axons to L1 such as L3, L2a, and cHV (Vates et al., 1996). Indeed, neurons near the NIf/L1 border had heterogeneous physiological proper-

ties, but we have no evidence as to the morphological characteristics of these neurons.

Nature of the correlated NIf and HVC activity

In our experiments, a useful and distinguishing feature of NIf was the correlation of its ongoing activity with that of the ongoing activity in HVC. The correlated activity was particularly prominent in multiunit traces and evident in ~80% of the NIf neurons sampled. In contrast, only three of the 15 L1 neurons recorded simultaneously with HVC exhibited a correlation in ongoing activity that could be tested. Although the sample of neurons is small, the correlation of ongoing activity in NIf and HVC does not appear to be predictive of song selectivity.

In principle, two possibilities could account for the observed correlations in the ongoing activity of NIf and HVC. HVC ongoing activity might be driven by NIf. The correlated activity could also be a response to a common input to both structures. We found some peaks in the cross-correlograms with a positive lag (NIf leading) but others that were centered around zero (Fig. 7). The breadth of the peaks also indicated considerable variability in the correlations. Had we observed correlations between NIf and HVC that were characterized by narrow, tall peaks, with NIf leading, this would support the idea that activity in NIf drives activity in HVC. However, failing to observe such correlations limits our conclusions. As a population, NIf projects broadly and nontopographically onto HVC (Fortune and Margoliash, 1995). Too little is known about the projections of single neurons, however, to estimate the contribution of a single NIf neuron to the total input of an HVC neuron and consequently its effect on the measured cross-correlations. Additionally, when multiunit data enter into the cross-correlation, the strength of the correlation may be influenced by the degree of correlation among neurons recorded on one of the two electrodes, in this case HVC, as well as the correlation between neurons on the different electrodes (Bedenbaugh and Gerstein, 1997).

By an alternate account, the thalamic nucleus Uva, which projects to both NIf and HVC, may be the source of the correlated activity. Under this scenario, the correlated ongoing bursting could be viewed as a component of the motor activity in NIf and HVC, because Uva is integral to the motor control pathway in the song system (Williams and Vicario, 1993; Striedter and Vu, 1998). Uva is not thought to be auditory, so the lack of an increase in correlated activity during auditory stimulation is consistent with this possibility.

The correlated activity would also be observed if NIf and HVC received common input from auditory structures. This explanation could account for the correlation in ongoing activity observed in a small number of L1 sites. Viewed within the context of a sensory hierarchy, it is surprising that if auditory structures provided sufficient common input to NIf and HVC to drive them synchronously during ongoing activity, that the correlation would remain relatively unchanged during auditory stimulation (Fig. 6). A more parsimonious explanation is that the correlated activity represents a motor patterning component of the system, to which an auditory component is added.

REFERENCES

Bedenbaugh P, Gerstein GL (1997) Multiunit normalized cross correlation differs from the average single-unit normalized correlation. *Neural Comput* 9:1265–1275.
 Boughman JW (1998) Vocal learning by greater spear-nosed bats. *Proc R Soc Lond B Biol Sci* 265:227–233.
 Capsius B, Leppelsack HJ (1996) Influence of urethane anesthesia on

neural processing in the auditory cortex analogue of a songbird. *Hear Res* 96:59–70.
 Doupe AJ, Konishi M (1991) Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad Sci USA* 88:11339–11343.
 Eales LA (1985) Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Anim Behav* 33:1293–1300.
 Esser KH (1994) Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. *NeuroReport* 5:1718–1720.
 Fortune ES, Margoliash D (1992) Cytoarchitectonic organization and morphology of cells of 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.
 Heiligenberg W (1991) Neural nets in electric fish. Cambridge, MA: MIT.
 Immelmann K (1969) Song development in the zebra finch and other estrilid finches. In: *Bird vocalizations* (Hinde RA, ed), pp 61–77. Cambridge, UK: Cambridge UP.
 Katz LC, Gurney ME (1981) Auditory responses in the zebra finch's motor system for song. *Brain Res* 211:192–197.
 Kelley DB, Nottebohm F (1979) Projections of a telencephalic auditory nucleus-field L in the canary. *J Comp Neurol* 183:455–469.
 Konishi M (1978) Auditory environment and vocal development in birds. In: *Perception and experience* (Walk RD, Pick HLJ, eds), pp 105–118. New York: Plenum.
 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.
 Leppelsack HJ (1983) Analysis of song in the auditory pathway of song birds. In: *Advances in vertebrate neuroethology* (Ewert JP, Capranica RR, Ingle DJ, eds), pp 783–799. New York: Plenum.
 Leppelsack HJ, Vogt M (1976) Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. *J Comp Physiol* 107:263–274.
 Lewicki MS (1994) Bayesian modeling and classification of neural signals. *Neural Comput* 6:1005–1030.
 Lewicki MS, Arthur BJ (1996) Hierarchical organization of auditory temporal context sensitivity. *J Neurosci* 16:6987–6998.
 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, Fortune ES (1992) Temporal and harmonic combination-sensitive neurons in the zebra finch's HVC. *J Neurosci* 12:4309–4326.
 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 (1997) Three models of song learning: evidence from behavior. *J Neurobiol* 33:501–516.
 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.
 Mello CV, Vates GE, Okuhata S, Nottebohm F (1998) Descending auditory pathways in the adult male zebra finch (*Taeniopygia guttata*). *J Comp Neurol* 395:137–160.
 Muller CM, Leppelsack HJ (1985) Feature extraction and tonotopic organization in the avian auditory forebrain. *Exp Brain Res* 59:587–599.
 Nottebohm F (1972) The origins of vocal learning. *Am Nat* 106:116–140.
 Nottebohm F, Kelley DB, Paton JA (1982) Connections of vocal control nuclei in the canary telencephalon. *J Comp Neurol* 207:344–357.
 Price PH (1979) Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93:268–277.
 Reiss D, McCowan B (1993) Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *J Comp Psychol* 107:301–312.
 Scheich H, Langner G, Bonke D (1979) Responsiveness of units in the

- auditory neostriatum of the Guinea fowl (*Numida meleagris*) to species-specific calls and synthetic stimuli. II. Discrimination of iambus-like calls. *J Comp Physiol* 132:257–276.
- Snowdon CT, Hausberger M (1997) Social influences on vocal development. Cambridge, UK: Cambridge UP.
- Solis MM, Doupe AJ (1997) Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. *J Neurosci* 17:6447–6462.
- Striedter GF, Vu ET (1998) Bilateral feedback projections to the forebrain in the premotor network for singing in zebra finches. *J Neurobiol* 34:27–40.
- Suga N (1989) Principles of auditory information-processing derived from neuroethology. *J Exp Biol* 146:277–286.
- Takahashi TT (1989) The neural coding of auditory space. *J Exp Biol* 146:307–322.
- Tanaka K (1996) Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109–139.
- 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.
- 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. *J Comp Neurol* 366:613–642.
- Vates GE, Vicario DS, Nottebohm F (1997) Reafferent thalamo-“cortical” loops in the song system of oscine songbirds. *J Comp Neurol* 380:275–290.
- Vicario DS, Yohay KH (1993) Song-selective auditory input to a forebrain vocal control nucleus in the zebra finch. *J Neurobiol* 24:488–505.
- Volman SF (1993) Development of neural selectivity for birdsong during vocal learning. *J Neurosci* 13:4737–4747.
- Volman SF (1996) Quantitative assessment of song-selectivity in the zebra finch “high vocal center.” *J Comp Physiol [A]* 178:849–862.
- Whaling CS, Solis MM, Doupe AJ, Soha JA, Marler P (1997) Acoustic and neural bases for innate recognition of song. *Proc Natl Acad Sci USA* 94:12694–12698.
- Wild JM (1994) Visual and somatosensory inputs to the avian song system via nucleus uvaeformis (Uva) and a comparison with the projections of a similar thalamic nucleus in a nonsongbird, *Columba livia*. *J Comp Neurol* 349:512–535.
- Williams H (1989) Multiple representations and auditory-motor interactions in the avian song system. *Ann NY Acad Sci* 563:148–164.
- Williams H, Vicario DS (1993) Temporal patterning of song production: participation of nucleus uvaeformis of the thalamus. *J Neurobiol* 24:903–912.
- Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. *Science* 273:1871–1875.