

Singing-Related Neural Activity in a Dorsal Forebrain–Basal Ganglia Circuit of Adult Zebra Finches

Neal A. Hessler and Allison J. Doupe

Keck Center for Integrative Neuroscience, Departments of Physiology and Psychiatry, University of California, San Francisco, California 94143-0444

The anterior forebrain pathway (AFP) of songbirds, a specialized dorsal forebrain–basal ganglia circuit, is crucial for song learning but has a less clear function in adults. We report here that neurons in two nuclei of the AFP, the lateral magnocellular nucleus of the anterior neostriatum (LMAN) and Area X, show marked changes in neurophysiological activity before and during singing in adult zebra finches. The presence of modulation before song output suggests that singing-related AFP activity originates, at least in part, in motor control nuclei. Some neurons in LMAN of awake birds also responded selectively to

playback of the bird's own song, but neural activity during singing did not completely depend on auditory feedback in the short term, because neither the level nor the pattern of this activity was strongly affected by deafening. The singing-related activity of neurons in AFP nuclei of songbirds is consistent with a role of the AFP in adult singing or song maintenance, possibly related to the function of this circuit during initial song learning.

Key words: birdsong; songbird; LMAN; Area X; feedback; deafening; reafferent; variability

The singing of birds is a temporally complex behavior (see Fig. 1A) that is crucial in courtship and territorial contexts (for review, see Catchpole and Slater, 1995). Like human speech, song is learned early in an individual's life, by listening to and copying the vocalizations of adults. In both songbirds and humans, vocal output is strongly dependent on access to auditory feedback of the individual's own vocalizations, especially during learning and to a somewhat lesser extent in the mature individual (Konishi, 1965; Nordeen and Nordeen, 1992; Leonardo and Konishi, 1999). Finally, like humans, songbirds have evolved specialized forebrain circuitry for producing complex vocalizations.

In songbirds, this circuitry, which is not found in closely related species that do not learn to sing, is known as the song system (for review, see Brenowitz, 1997; Doupe and Kuhl, 1999). Nuclei in the motor pathway (see Fig. 1B, gray) are essential throughout life for normal song production. Their activity during singing is highly patterned (McCasland, 1987; Yu and Margoliash, 1996), and disruption of their activity alters or abolishes song output (Nottebohm et al., 1976; Vu et al., 1994). In contrast, a second circuit of forebrain song nuclei has a less clear function in adult motor production but, like hearing, plays a particularly critical role during learning. This anterior forebrain pathway (AFP) (see Fig. 1B, black) links the motor pathway nuclei HVC and robust nucleus of the archistriatum (RA), but via an indirect loop through basal ganglia (Area X), thalamus [the medial portion of the dorsolateral thalamus (DLM)], and a cortex-like nucleus, the lateral magnocellular nucleus of the anterior neostriatum

(LMAN) (Okuhata and Saito, 1987; Bottjer et al., 1989; Luo and Perkel, 1999a,b). Lesions or pharmacological inactivation of AFP nuclei markedly disrupt initial song learning and production but have no obvious effect on stable "crystallized" song of adult zebra finches (Bottjer et al., 1984; Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Basham et al., 1996). Moreover, in anesthetized birds, neurons in these nuclei respond more strongly to playback of the bird's own song than to songs of conspecific birds, and this neural selectivity develops over the course of song learning (Doupe, 1997; Solis and Doupe, 1997, 1999). Thus, results of both lesion and physiological studies support a model in which the AFP is involved in the evaluation of self-produced vocalizations that occur as the bird learns to produce a copy of the song memorized early in life.

Although the AFP clearly is crucial during juvenile song learning, its function in adult finches is less apparent. Recent experiments, however, have raised several possibilities. First, the maintenance of adult zebra finch song by auditory feedback (Nordeen and Nordeen, 1992; Leonardo and Konishi, 1999) could be mediated by the AFP. Furthermore, the AFP may be involved in the perception and classification of conspecific songs (Scharff et al., 1998) and in the modulation of behavior by social context (Jarvis et al., 1998; Hessler and Doupe, 1999). To investigate these possible functions, we undertook to describe in detail the activity of the AFP nuclei in awake adults. We report here that LMAN and Area X are strongly active during singing and that this activity has many features in common with the singing-related activity of the motor nuclei HVC and RA (McCasland, 1987; Yu and Margoliash, 1996). Moreover, although we found that neurons in some birds can show selective auditory responses to playback of songs, initial experiments on the effects of manipulating hearing during singing indicate that much of the singing-related activity does not depend on hearing. These results suggest that the AFP, although it is not required for adult song production, nevertheless has some function in this complex learned

Received July 6, 1999; revised Aug. 26, 1999; accepted Sept. 13, 1999.

This work was supported by the Merck Fund, the EJLB Foundation, and National Institutes of Health (Grants MH55987 and NS34835 to A.J.D. and NS09913 to N.A.H.). We thank Adria Arteseros, Laszlo Bocskai, Robin Booth, Gabriela Carrillo, Cooper Roddey, Marc Schmidt, and Amy Tam for technical assistance, and Michael Brainard, Mimi Kao, and Todd Troyer for comments on this manuscript. We also thank Mimi Kao for providing data from several recordings in RA.

Correspondence should be addressed to Neal A. Hessler, Department of Physiology, Box 0444, 513 Parnassus Avenue, University of California, San Francisco, CA 94143-0444. E-mail: neal@keck.ucsf.edu.

Copyright © 1999 Society for Neuroscience 0270-6474/99/1910461-21\$05.00/0

behavior, perhaps serving in part as an “efference copy” of singing motor commands from the song system motor nuclei.

MATERIALS AND METHODS

Animals. Adult (>125 d old) male zebra finches (*Taeniopygia guttata*) were used for experiments. One bird was purchased from a local supplier, and the rest were raised from hatching in our colony. The care and treatment of experimental animals was reviewed and approved by a university animal care and use committee at the University of California, San Francisco (UCSF). Birds were selected for recordings on the basis of size and singing frequency and were then isolated in a small cage inside a sound-attenuation chamber (Acoustic Systems, Austin, TX) and supplied with food, water, and grit *ad libitum*, with occasional supplements of boiled chicken egg. Several days after spontaneous singing occurred frequently in this chamber, a chronic recording apparatus was implanted.

Surgical procedures. Birds were anesthetized with an intramuscular injection of 30–40 μ l equithesin (0.85 gm chloral hydrate, 0.21 gm pentobarbital, 0.42 gm MgSO₄, 2.2 ml 100% ethanol, and 8.6 ml propylene glycol in H₂O to a total volume of 20 ml) and placed in a stereotaxic device stabilizing the head at the beak and ear canals. The beak was positioned at an angle of 50° from vertical. The position of the posterior border of the divergence of the central sinus at the boundary of the forebrain and the cerebellum was noted. A small hole was made in the skull at a specific position relative to this reference coordinate: for LMAN and Area X, 5 mm anterior and 1.7 mm lateral; for HVC, 0.0–0.2 mm anterior and 2.0–2.2 mm lateral; for RA, 0.2–0.7 mm anterior and 2.1 mm lateral. All recordings were made from nuclei in the right hemisphere. A previous study has reported that the pattern of singing-related activity in the right and left HVC is similar (McCasland, 1987). A lightweight (~1 gm) microdrive (UCSF Physiology Shop) carrying an insulated tungsten electrode of impedance 2–5 mOhm (either AM-Systems, Carlsborg, WA, or FHC, Bowdoinham, ME) was positioned stereotaxically such that the electrode tip was ~500 μ m above LMAN or RA or ~300 μ m above HVC. A reference ground electrode (uninsulated tungsten electrode, AM-Systems) was implanted in the contralateral (for LMAN and Area X) or ipsilateral (for HVC and RA) hemisphere such that it passed within ~2 mm of the targeted nucleus. The microdrive and its connector socket (FHC) were secured to the skull with epoxy (Dev-Con, Wood Dale, IL) and dental cement (Dentsply, Milford, DE), and a protective cap (3 M, St. Paul, MN) was fixed around it (see Fig. 1C).

Deafening. Deafening was performed in two stages, by the method of Konishi (1965). Before implantation of the microdrive and electrode, birds were anesthetized with an injection of equithesin or by inhalation of isoflurane. Feathers occluding the right external auditory meatus were plucked, and the skin flap within the ear canal was cut off to expose the tympanic membrane. The tympanic membrane was detached from the columella, and the columella was removed. A hook of tungsten wire was inserted through the oval window and withdrawn with the attached cochlea. The extracted cochlea was examined under a dissecting microscope to ensure that the entire structure had been successfully removed, and the skin incision was closed with cyanoacrylate adhesive. Several weeks later, after the skin inside the right ear canal had regrown, the bird could be positioned in the stereotaxic device using ear bars, and the microdrive-electrode was implanted. After several days of recording in LMAN, the left cochlea was removed as above, to complete the deafening.

Physiological recording. During each recording session, one end of a flexible lead terminating in a small operational amplifier circuit (TLC27L7C, Texas Instruments, Dallas TX) (Buzsaki et al., 1989) was connected to the small socket on the bird's head, and the other end was connected to a rotating commutator (H. Adams, Caltech Machine Shop, or Radio Shack, Fort Worth, TX). The neural activity signal passed through the commutator to a differential amplifier (A-M Systems), where it was filtered between 300 Hz and 5 kHz. The acoustic signal in the sound box was recorded by a small microphone (Radio Shack) adjacent to the cage. After several days of adaptation to being attached to recording leads, most birds began singing while they were isolated in the sound box. During experiments, the electrode was lowered to a position where large action potentials of single or multiple neurons could clearly be differentiated from background neural activity (larger spike amplitudes in each recording session ranged from ~300 μ V to >1 mV, peak to peak). The bird's behavior was monitored via a video camera inside the sound box, and the video, neural, and acoustic signals were archived on videotape. A computer program (developed by C. Malek, Caltech, and C. Roddey, UCSF) monitored the sound amplitude inside the sound box

and triggered the recording to computer disk of the acoustic and neural signals (both digitized at 32 kHz) for an adjustable period before and after passage of the sound amplitude over a threshold level. Previously recorded versions of the bird's own song, and in some cases songs of other zebra finches, were played from a small speaker in the sound box during periods when the bird was not singing. The volume of song playback from the speaker was set such that its intensity in the vicinity of the bird's cage and the amplitude of sound produced by the bird were approximately equal. The intensity of both signals was monitored by a microphone immediately adjacent to the cage (~10 cm from the bird).

Recordings were made at intervals of 1 d to 1 week, over periods of 2 weeks to 2 months. After most recording sessions, the electrode was retracted to a position above the nucleus. For each successive recording session, the electrode was advanced by at least 80 μ m beyond the previous recording depth. After completion of recordings, small electrolytic lesions (20 μ A for 5 sec) were made at previously recorded sites. Animals were deeply anesthetized with Metofane (Pitman-Moore, Mundelein, IL) and intracardially perfused with 0.9% saline, followed by 3.7% formalin. Lesions were localized in 40 μ m Nissl-stained brain sections. Locations of all recording sites were confirmed to be within song nuclei by position relative to depth of marker lesions.

Behavioral analysis. Zebra finches, like other songbirds, produce several distinct types of vocalizations, which have been extensively characterized in behavioral studies (for review, see Zann, 1996). Here, we have limited our analysis to “undirected song,” which birds produced while out of visual contact with other birds. In zebra finches, undirected song contains a variable number of motifs, a stereotyped series of approximately 3–10 discrete vocal elements (syllables) (see Fig. 1A for examples of song components), preceded by several simple introductory elements. Song initiation was defined by the onset of the first introductory element or song syllable with a <300 msec interval between its offset time and the onset time of the next song element (further introductory element or motif syllable). Song termination was defined by the offset time of a syllable that was followed by at least 300 msec before the next vocalization (~1000% of average intersyllable interval duration; data not shown). This interval was lengthened for one bird that had an exceptionally long intersyllable interval within its motif. For characterization of presinging activity level, only song initiations (as defined above) before which the bird had been silent for at least 3 sec were used. For characterization of postsinging activity level, only song terminations after which the bird was quiet for at least 3 sec were used. Thus, for many songs, the presong or postsong period or both were not included in the analysis of activity. Background nonsinging activity was recorded during periods in which no sounds were produced; such periods were required to follow any vocalization by at least 3 sec and to precede any vocalization by at least 3 sec.

Analysis of acoustic and neural signals. All analysis was performed on digitized acoustic and neural signals using Matlab (Mathworks, Boston, MA). For classification of vocalizations, the sound waveform was first compressed to speed analysis routines. After filtering (0.5–8 kHz band-pass), the sound waveform was rectified and smoothed with a 1 msec (SD) Gaussian function and resampled at 1 kHz (see, for example, Fig. 5C). Individual song components were readily identifiable by visual inspection of this signal. To locate productions of a specific pattern of syllables in a series of data files, a template was made consisting of the rectified song waveform representing the song elements (typically a template consisted of approximately three syllables of ~300 msec duration, including some chosen because of their complex temporal structure). For each data file, this template was iteratively subtracted from the rectified sound waveform, at steps of 5 msec. Close matches of the template to the test sound resulted in local minima in the resulting error function. The accuracy of motif detection by alignment to such minima was confirmed by visual inspection. This alignment was used for subsequent analysis of corresponding neural activity. Digitized neural activity waveforms were rectified and smoothed with a 10 msec (SD) Gaussian function and resampled at 1 kHz (see, for example, Fig. 2C). This signal was used in most analyses of LMAN activity level.

To estimate the onset time of the mean activity peak located near song initiation, a line was first fitted (by the least-squares method) to the linearly rising phase of the activity peak, beginning after the smaller slope of the slowly rising phase of activity level was replaced by the larger slope during the sharp rise in activity level, and ending before the time at which the slope of the activity level began decreasing at its peak. The onset time was then estimated by determining the intersection of this fitted line with the average background activity level.

To quantify the temporal relationship between LMAN neural activity

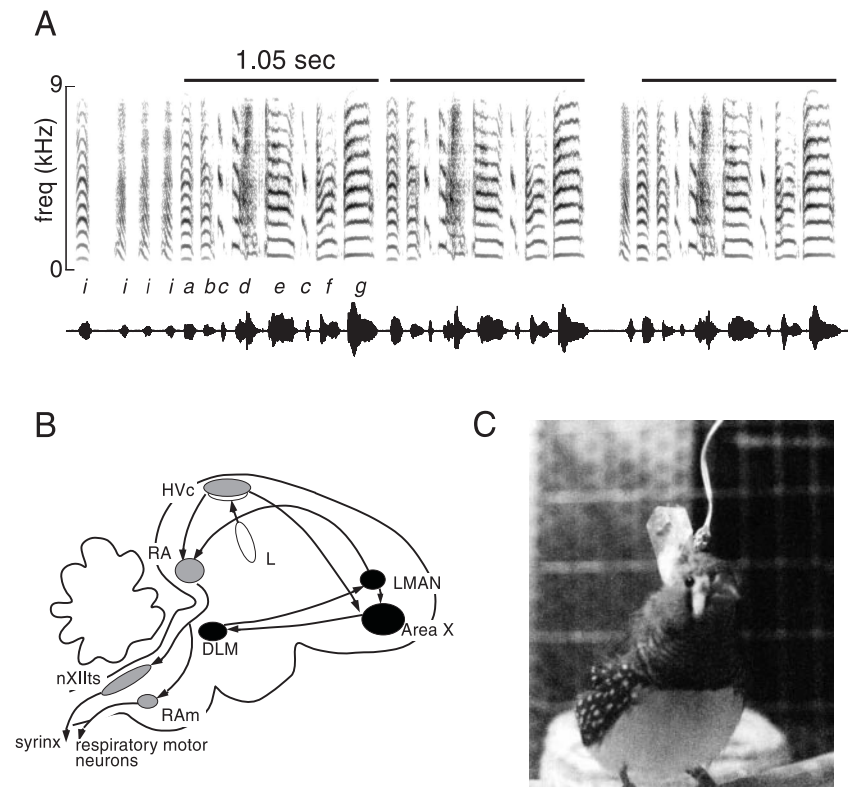


Figure 1. Singing behavior, neural circuitry, and experimental subjects. *A*, A spectrogram (plot of frequency vs time, with loudness indicated by the darkness of the signal) of a zebra finch song shows the characteristic features of song. Stereotyped sequences of syllables (*lower case letters*), called “motifs” in zebra finches (indicated by *dark bars*), may be sung from one to several times in succession, preceded by a variable number of short introductory elements (*i*). Amplitude oscillogram of song is plotted below spectrogram. Audio file of song is located at <http://www.keck.ucsf.edu/~neal/jns99/fig1.wav>. *B*, The song system is a network of discrete nuclei involved in song learning and production. Nuclei in the motor pathway (*gray*) are critically involved in song production, whereas nuclei in the AFP (*black*) are necessary for song learning but not for production of crystallized adult song. *C*, Photograph of a representative experimental bird during a recording session. A small microdrive (approximate dimensions: 15 mm high, 6 mm long, 3 mm wide) is inside a protective polyurethane cap on top of the bird’s head; flexible copper-lead and silver wires terminate at the bird’s head in a small op-amp circuit.

level and song components (see Fig. 10), for each motif rendition the rectified sound waveform was converted into a step function with syllable presence indicated by a 1 and absence by a 0. We then calculated the cross-covariance of the neural activity trace with the discretized song (i.e., cross-correlation after normalization of means of both sequences; the cross-covariance was normalized such that auto-covariance of a signal with itself at zero lag equals 1.0). The cross-covariance was calculated with the neural activity signal shifted from -100 to $+100$ msec relative to the song signal, at steps of 1 msec.

For several recordings in LMAN and all recordings in Area X, the firing of single units or clusters of several neurons was analyzed by off-line sorting of units from neural activity traces. Sorting was performed by counting waveform events that passed below a negative threshold and then above a positive threshold within 0.5 msec. Accurate isolation of single units was verified by visual inspection and by examination of interspike intervals.

Results for all analyses are presented as mean \pm SD.

RESULTS

Singing-related neural activity in LMAN

Overall relationship of LMAN firing to song production

Neural activity was recorded in LMAN of 11 adult zebra finches during both song production and nonsinging periods. In all birds, a conspicuous change in multiunit firing occurred during singing. In a representative recording, activity increased markedly when a bird produced each of three successive songs (Fig. 2*A,B*). Beneath the raw neural activity signal is plotted the smoothed rectified waveform (Fig. 2*C*), which reflects overall multiunit activity level (see Materials and Methods for details). Activity level is clearly different between singing and quiet periods. Additional characteristics of singing-related activity can be seen in an expanded view of the final 4 sec from Figure 2*A–C* (Fig. 2*D,E*). The complex spectral and temporal structure of this bird’s song, which includes two iterations of a multinote syllable (*), is evident in the spectrogram (Fig. 2*D*). The neural activity trace (Fig. 2*E*) shows condensed bursts of activity, as well as inhibition

of background neural activity; both of these features were typically associated with singing.

Singing-related LMAN activity for two additional representative birds is shown in Figure 3. As in the recording shown in Figure 2, the level of neural activity during singing both rises to higher peaks and falls to equal or lower minima than it does during nonsinging periods (mean background activity level is indicated by the *dotted lines*). To summarize these data quantitatively, we compared the distribution of neural activity levels during singing to that during background nonsinging periods (the smoothed rectified waveform was sampled at 1 msec intervals). Epochs of recording sessions were used in which the amplitude of neural activity during singing periods was relatively stationary; that is, there was no overall drift in spike amplitudes. For example, note the lack of change in response amplitude between the two songs displayed for each bird in Figure 3 that were selected from the first and last 10% of the recording session (duration of recording used for dc-12 was 3 hr and for dc-18 was 1 hr). For each of these two birds, the 10th, 50th, and 90th percentiles of the activity distributions during the entire recording session for nonsinging (*bg*: background, sampled from quiet periods) and singing (*sing*) periods are displayed immediately to the right on the same ordinate. The median and range of neural activity levels are clearly different between nonsinging and singing periods. To the right of these distribution plots, activity level distributions from additional recording sessions at different sites within LMAN of the same birds are plotted. Note the similarity in range and medians of the activity distributions for the three recording sessions for dc-12 (top) and the two sessions for dc-18 (bottom).

To quantify the effect of singing on overall activity level for all birds, the average activity level in successive 1 sec epochs of background and singing periods was calculated (the duration of the shortest songs was ~ 1 sec). In 27 of 27 recordings from 11

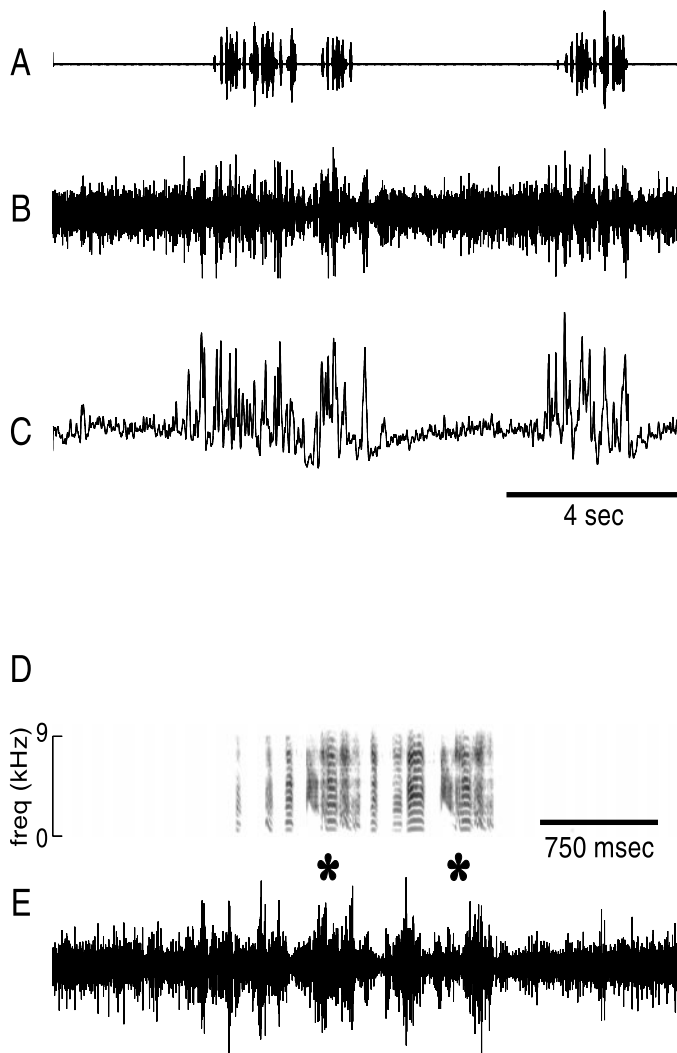


Figure 2. Neural activity in LMAN associated with song production. *A*, Amplitude oscillogram of sound recorded in a 14.5 sec period during which the bird mx-17 produced three songs. *B*, Simultaneously recorded neural activity in LMAN. *C*, Rectified waveform of the neural activity in *B* smoothed with a 10 msec Gaussian function. *D*, Spectrogram of the final 4 sec of singing from *A* (epoch indicated by scale bar in *C*). *E*, The neural activity waveform associated with the song in *D* highlights the characteristic burstiness of singing-related firing; the asterisks denote two renditions of the same complex syllable. Audio file of songs in *A* is located at <http://www.keck.ucsf.edu/~neal/jns99/fig2.wav>.

birds, the level of activity in LMAN was higher during singing than during nonsinging periods ($p < 0.01$ for all unpaired t tests for each recording session), with an average increase of $132 \pm 8\%$ (SD) (Fig. 3*B*). In all recording sessions, modulation of activity also increased during singing (Figs. 2, 3*A*). This effect was quantified by calculating the coefficient of variation (c.v.), the SD of the values of an activity distribution normalized by the mean activity, during singing and nonsinging periods. Paired comparisons within recording sessions revealed a significantly larger c.v., and thus increased modulation, of neural activity levels during singing compared with that during background periods ($p < 0.01$, sign test) (Fig. 3*C*). Across all recordings, the mean c.v. for background nonsinging periods was 0.11 ± 0.02 , and for singing periods the c.v. was 0.22 ± 0.05 .

For the recording sessions shown in Figures 2 and 3*A*, we also

quantified activity level by calculating the multiunit spike rate, using off-line window discrimination (for details, see Materials and Methods). When spike arrival times were convolved with a 10 msec Gaussian function, as was done for the rectified neural waveforms, the distributions of activity levels for background and singing periods were similar to those quantified by rectifying and smoothing activity waveforms. The magnitude of increase in activity level from background to singing periods for these three recording sessions was similar when quantified by these two methods: 1.17 versus 1.25, 1.31 versus 1.26, and 1.38 versus 1.28 for rectified waveform and sorted multiunit spikes respectively. Thus, as a measure of overall population activity level, the rectified and smoothed activity waveform is approximately equivalent to sorted multiunit spikes.

Activity level in LMAN rises before and is depressed after singing

In addition to its increase during singing, the level of activity was consistently altered before and after singing in all recordings from LMAN. This effect, which was evident in Figures 2 and 3*A*, is shown in Figure 4*A–C* in a different manner for three representative birds, by plotting LMAN activity level during multiple presong and postsong epochs. To limit the possible influence of previous or subsequent singing, we examined only song initiations before which the birds had been silent for >3 sec, and only song terminations after which birds refrained from vocalizing again for >3 sec (for details, see Materials and Methods). In the top section of each panel, the multiunit activity level during successive initiations and terminations, as quantified by rectified neural waveforms, is plotted on a gray scale, with darker shades representing high activity levels. These plots illustrate the consistent rises in activity level before song initiations and the clear diminution of activity after song terminations.

These features of the relationship of activity to song initiation and termination are evident in plots of the mean activity level across all renditions (Fig. 4*A–C*, bottom panels). In all recordings, a prominent peak of activity was located very near song onset (peri-initiation peak). Initiation latencies of these peaks for the three examples in Figure 4 are indicated by an asterisk beneath the presinging mean activity trace (latencies estimated by fitting a line to the peak's rising phase; for details, see Materials and Methods). Across all recordings, the mean onset latency of the peri-initiation peak relative to song initiation was 70 ± 24 msec, with the average peak maximum located 9 msec before song onset (Fig. 4*Db*) (24 recording sessions from 10 birds; mean 56 song initiations per session; for recordings with fewer than 10 songs in which bird was quiet for the preceding 3 sec, songs in which the bird was quiet for at least 1 sec were used; data from one bird for which there were fewer than 10 such songs were not included in this analysis).

Along with the sharp rise of activity near song onset, in many recordings a gradual rise in activity level also preceded song initiation. To compare this feature across recordings, we calculated the time at which the mean activity level rose and remained above the 70th percentile of the background activity level distribution (represented by the dotted line on the left side of the bottom panels in Fig. 4*A–C*). For these measurements, we selected song initiations that were preceded by at least 2.5–3 sec of silence: 15 recording sessions contained more than 10 such songs. The mean activity level in 10 of 15 of these recordings rose above the 70th percentile of background activity level before the immediate presong peak, at an average latency of 419 ± 137 msec (Fig. 4*Da*)

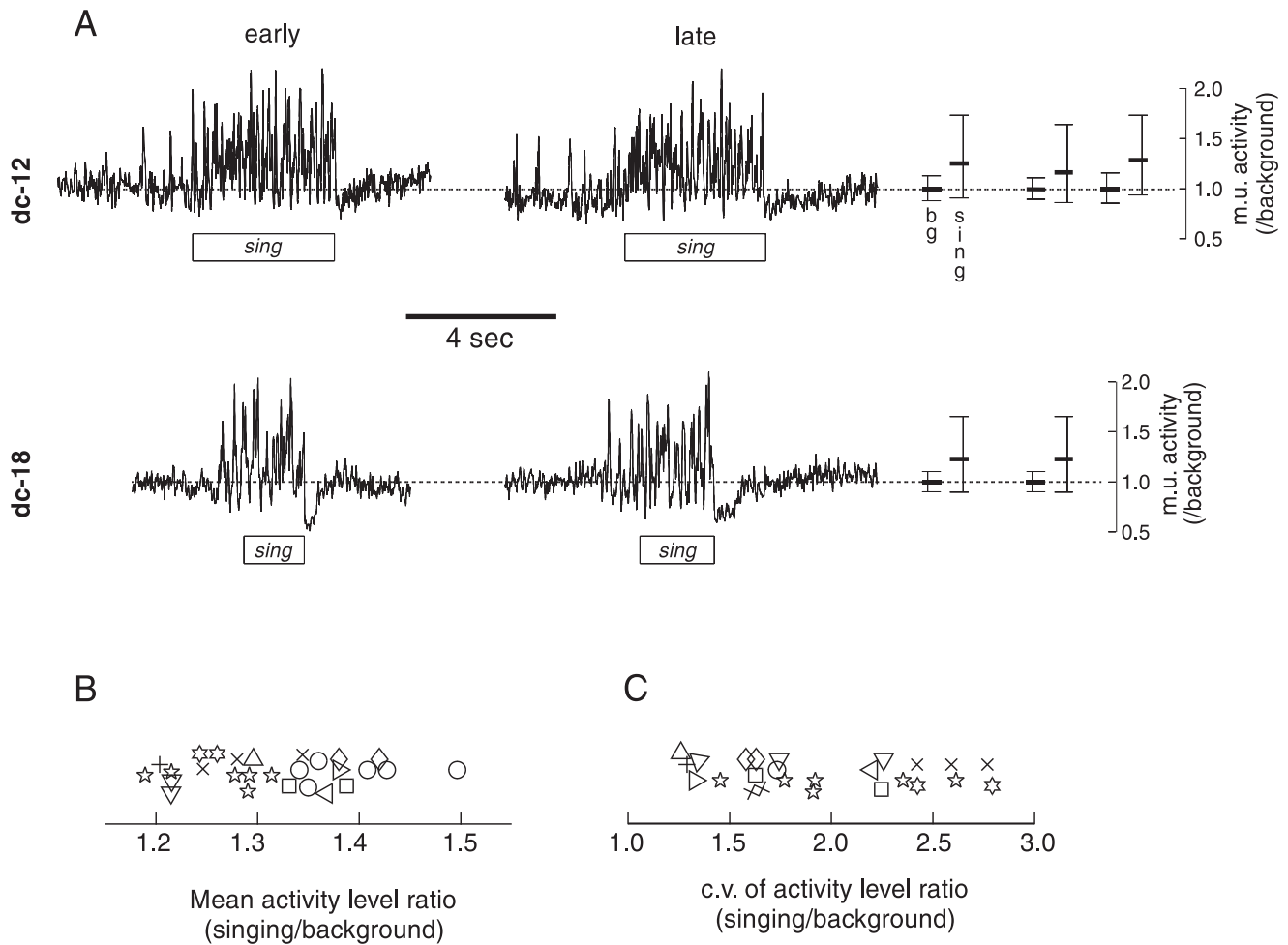


Figure 3. Singing-related activity is seen consistently in all recordings from LMAN. *A*, Rectified and smoothed (10 msec Gaussian window) waveforms of neural activity recorded during one session for bird dc-12 and one session for bird dc-18. For both sessions, in the *left panel* the pattern of activity before, during, and after a song bout from the first 10% of the recording session (*early*) is plotted; in the *right panel* activity is plotted in the same way for songs from the final 10% of the session (*late*). *Bars* to the immediate *right* depict the distribution of activity levels (sampled at 1 msec intervals) during nonsinging background (*bg*) and singing (*sing*) periods across the entire recording session. *Top* and *bottom ticks* denote the 90th and 10th percentiles of these distributions, whereas *middle thick ticks* denote medians. Values for both distributions have been normalized by the mean background level. The total duration of singing used was 108 sec for dc-12 and 270 sec for dc-18. Activity level distributions for two (*dc-12*) and one (*dc-18*) additional recording sessions at different sites in LMAN are plotted on the same abscissa to the *far right*. *B*, Mean activity level during singing divided by mean activity level during background period for all recording sessions from each bird (each *symbol* represents a different bird; some *symbols* have been offset vertically to reduce overlap). *C*, coefficient of variation (*c.v.*) of activity level during singing divided by *c.v.* of activity level during background period for each recording from every bird (each *symbol* represents a different bird). For comparisons of activity in *B* and *C*, the mean duration of singing was 158 ± 168 sec (range 16–802 sec) and of background was 136 ± 157 sec (range 30–795 sec).

(for recording sessions in Fig. 4*A,B*, latencies were 331 and 352 msec, respectively). The mean activity level in the remaining five recording sessions did not rise above background levels until the immediate peri-initiation peak, as in the example shown in Figure 4*C*.

In contrast to the presinging activity increases, firing was consistently depressed after song termination and required from 1 sec to several seconds to return to nonsinging levels (Fig. 4*A,C*, *right panels*). Across all recordings, activity level reached a minimum amplitude of 0.77 ± 0.11 relative to background at 198 ± 99 msec after song termination (Fig. 4*Dc*) (16 recordings from eight birds; data from three birds for which there were fewer than 10 quiet postsong intervals were not analyzed). Following this minimum, an average of 1172 ± 407 msec was required for activity levels to recover to over the 30th percentile of background activity level (represented by the *dotted line* in *right panels* of Fig.

4*A–C*; for recording sessions in *A–C*, latencies were 1485, 1479, and 314 msec, respectively).

These presinging and postsinging changes in activity level were extremely consistent; for all birds they occurred before and after almost every song. For example, in the recording session shown in Figure 4*A*, the mean activity level in a period from 50 msec preceding song to song onset was, for every rendition, higher than in a period from 2 to 1.5 sec before singing (Fig. 4*E*, ●). Conversely, the activity level in the immediate postsong period (averaged in a 50 msec window around the location of the minimum of the mean) was consistently lower than that from 1.5 to 2 sec after song termination (Fig. 4*E*, ○). For this and other recordings in which at least 10 presinging and post singing renditions were obtained, activity level was significantly higher just before song onset than 2 sec before singing and was significantly lower just after song termination than 2 sec later (comparisons were made

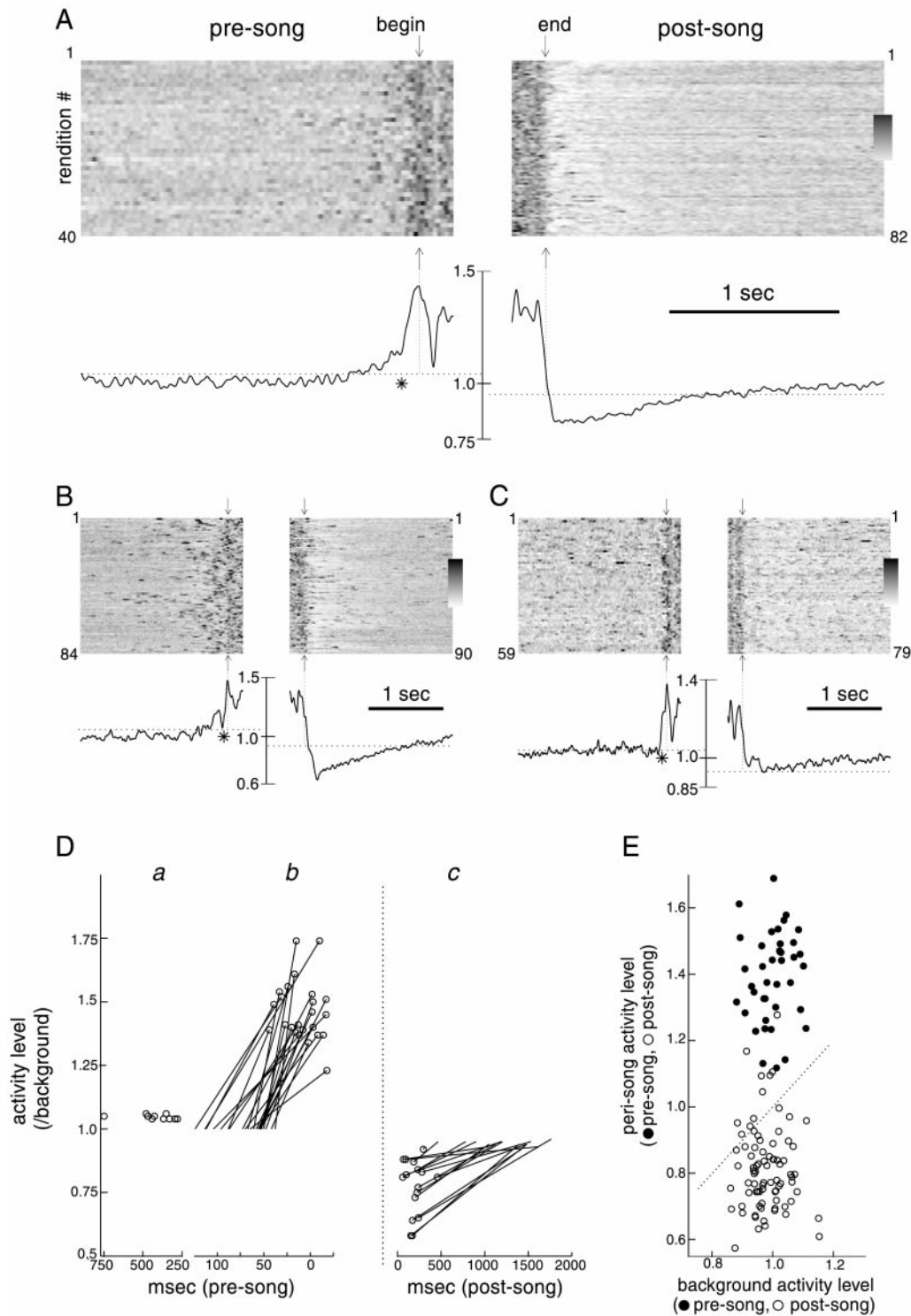


Figure 4. LMAN activity is consistently altered before song initiations and after song terminations. *A–C*, *Left* and *right* panels depict LMAN multiunit activity level during presinging and postsinging intervals for three representative birds. For each, the *top* gray scale panels plot activity level over time (*left* to *right*) in relation to successive song renditions (*horizontal* rows). Activity level has been quantified by rectifying and smoothing raw activity waveforms (see Materials and Methods). Increases in activity level are represented by darker shade, and decreases are represented by lighter shade. Times of song initiation and termination (by which activity traces for *left* and *right* panels, respectively, were aligned) are indicated by *arrows* above and below gray scale panels. Mean activity level (normalized to background nonsinging level) across all renditions is plotted below the gray scale plots, along with *dotted lines* indicating 70th (*left* panels) and 30th (*right* panels) percentiles of background activity level distributions; background activity was recorded during periods interspersed among presong and postsong recordings. *Asterisks* underneath mean activity level traces in *left* panels indicate calculated onset time of sharp activity peaks. Lower and upper limits of gray scale bars are 0.75 and 2.0, 0.25 and 2.75, and 0.75 and 1.9 for *A*, *B*, and *C*, respectively. *D*, Summary of timing and amplitudes for average presinging and postsinging parameters: *a*, time at which mean activity level rose above the 70th percentile of background activity level distribution for the 10 of 15 recording sessions in which this occurred before the peri-initiation peak. *b*, Peri-initiation peak timing and magnitude relative to background (○), connected by a *line* to the onset latency at which this peak rose from background (*Figure legend continues*)

as in Fig. 4E; $p < 0.01$ for all paired t tests comparing presong and postsong periods in each of 16 recording sessions from eight birds).

LMAN firing during repeated renditions of a stereotyped song motif

In addition to characterizing the general association of firing to song production, we examined the relationship between repeated song elements and patterns of neural activity. In zebra finch song, sequences of several syllables are produced in a stereotyped order to form a song motif, which can then be repeated a variable number of times (Fig. 1A). Here, by using a template-matching algorithm, we detected multiple renditions of stereotyped song motifs during single recording sessions and analyzed the pattern of neural activity associated with each motif (see Materials and Methods for details of detection protocol).

A representative example of data used in such an analysis is shown in Figure 5. A common song type of bird dc-18 contained a variable number of introductory elements (range 1–4 per song, mean 1.7), followed by two stereotyped five-syllable motifs that were separated by an intervening syllable (Fig. 5A–C, *s*). Note the similarity in acoustic structure between the repeated motifs. This song type was produced 159 times during a 4 hr recording session. The amplitude envelopes of all successive song renditions, aligned so that the amplitude envelopes of the final four syllables of the song were most similar, are shown in Fig. 5D (see Materials and Methods for details of alignment procedure). The temporal structure of all song renditions was extremely stereotyped. The alignment of the first motif appears more variable, primarily because of variability in the interval between syllable *e* of the first motif and syllable *s* (SD of this interval duration = 6.0 msec, the mean SD of nine other interval durations = 2.9 msec, range 1.6–4.3 msec). Across all renditions, the first motif of the songs in this recording session had a slightly shorter duration than the second motif ($p < 0.01$, paired t test, $n = 159$; first motif duration = 721 ± 8.2 msec, second motif 736 ± 7.88 msec).

During all of these song renditions, the pattern of activity in LMAN was relatively consistent in relation to song elements (Fig. 5E). High and low activity levels occurred during similar song components across the recording session, as reflected by the presence of dark and light vertical bands in Figure 5E. The pattern of neural activity averaged across all song renditions appeared very similar during production of the first and second motifs (Fig. 5F). Because of the greater variability in alignment of syllables across renditions for the first motif, the mean activity pattern during the first motif was less sharp than that during the second motif (Fig. 5F). When songs were aligned by the last four syllables of the first motif, however, the mean activity pattern for the first motif was nearly identical to that for the second motif in Figure 5F; the correlation coefficient between these activity patterns was then 0.96. In contrast to the reproducibility of the mean activity pattern when averaged across many renditions, there was a lower degree of stereotypy when activity was examined on a rendition-to-rendition basis (as could be seen previously in Fig.

2D,E). This variability, as well as the characteristic “bursty” quality of LMAN singing-related activity, is highlighted by the four successive activity waveforms, and their associated rectified waveforms, displayed in Figure 5G.

Variability in LMAN activity across multiple motif renditions

When activity level in LMAN during many renditions of a motif was examined, the overall pattern of activity could be seen to resemble the mean activity level (Fig. 5, compare *E*, *F*). When small numbers of motifs were inspected, however, it was more difficult to discern any repeated pattern of activity, because of the variability of activity between individual motifs. This is clear in a representative recording, in which activity level in LMAN during 10 successive renditions of the stereotyped song motif is plotted (Fig. 6A, *thin lines*). We quantified the degree of stereotypy of neural activity across renditions at single recording sites in two ways. As one measure, we calculated the cross-rendition c.v. of activity level across all renditions of an identical song motif (for each millisecond time point in the motif, SD of activity level across all renditions was divided by the mean activity level) (Fig. 6C, *dotted line*). The mean value of this measure across the entire song motif was similar in magnitude in all recordings from LMAN (Fig. 6D) (only recording sessions with more than 10 motif renditions were included), with an average of 0.18 ± 0.05 . As a second measure of variability across renditions, we calculated the correlation coefficients between the activity waveform for each individual rendition (Fig. 6A, *thin lines*) and the mean activity level across all renditions (Fig. 6A, *thick line*). For this recording in LMAN, the mean of these correlation coefficients across 360 motif renditions was 0.43 ± 0.17 (by comparison, for 159 renditions of the two-motif song in Fig. 5 the mean was 0.47 ± 0.12). Similar results were obtained for all other birds; across all recording sessions in LMAN the average correlation coefficient was 0.47 ± 0.09 (Fig. 6E).

In contrast to the results for LMAN recordings, when we recorded and analyzed multiunit activity in the same way for the song system motor nuclei HVC and RA, there was much higher stereotypy across motif renditions [see also Yu and Margoliash (1996); Vu et al. (1998)]. As can be seen for representative LMAN and RA recording sites, activity patterns during successive motif renditions were much less variable for multiunit responses in RA (Fig. 6A,B). Furthermore, when quantified in the two ways presented above, activity in nine of nine recording sessions from four birds in HVC (mean 305 motifs/session) and seven of seven recording sessions from five birds in RA (mean 65 motifs/session) was more stereotyped across renditions than that in all recordings from LMAN (Fig. 6D,E) (mean cross-rendition c.v. values for HVC and RA recordings were 0.08 ± 0.02 and 0.07 ± 0.01 , respectively, whereas mean correlation coefficients between activity during each rendition and mean activity levels were 0.83 ± 0.07 and 0.91 ± 0.03 , respectively).

Given the multiunit character of the recordings from these three nuclei, it is possible that some of the differences in stereotypy reflected a tendency for sites in HVC and RA recordings to

←

activity level, for 24 recording sessions from 10 birds. *c*, Timing and magnitude of postsong minimum (○), connected by a line to the time at which mean activity level rose above the 30th percentile of background activity level, for 16 recording sessions from eight birds. *E*, Reliability of presinging and postsinging activity changes for one recording. *Filled circles* represent, for each song initiation, mean activity level from 50 to 0 msec before initiation versus mean activity level from 2 to 1.5 sec before initiation. *Open circles* represent mean activity level in a 50 msec window around the average minimum versus mean activity level in a period from 1.5 to 2 sec after song termination. The *dotted line* indicates values at which activity level in these two periods would be equal. These data are from the recording session shown in *A*.

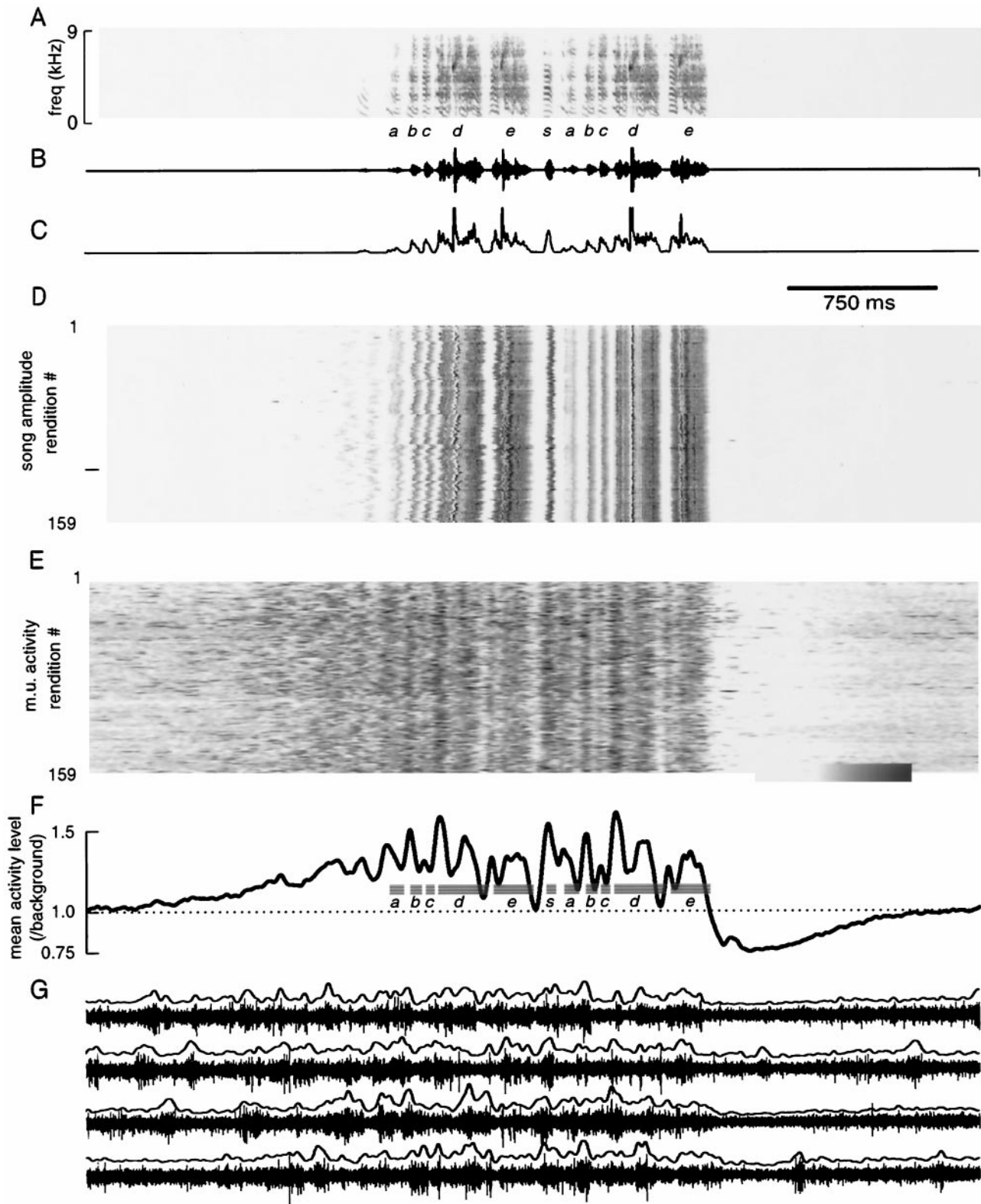


Figure 5. Neural activity in LMAN associated with repeated song motif renditions. *A–C*, Spectrogram (*A*), oscillogram (*B*), and amplitude envelope (*C*) of a song repeatedly produced by bird dc-18. *D*, Amplitude envelopes of 159 consecutive song renditions recorded over 3 hr, aligned for maximal overlap between the final four syllables. Darkness is proportional to sound intensity. *E*, Multiunit neural activity level recorded during each song rendition displayed in *D*. Activity level has been quantified by rectifying and smoothing raw activity waveforms. Lower (*left*) and upper (*right*) limits of gray scale bar are 0.75 and 2.25. *F*, Mean neural activity level (normalized by background nonsinging activity level: *dotted line*) across these 159 song renditions. Mean timing of syllable occurrence is overlaid for comparison of activity to song production. *G*, Raw waveforms of neural activity during four consecutive song renditions. The rectified integrated version of each waveform is plotted above the raw waveform. Audio files of five representative songs are located at <http://www.keck.ucsf.edu/~neal/jns99/fig5/fig5.html>.

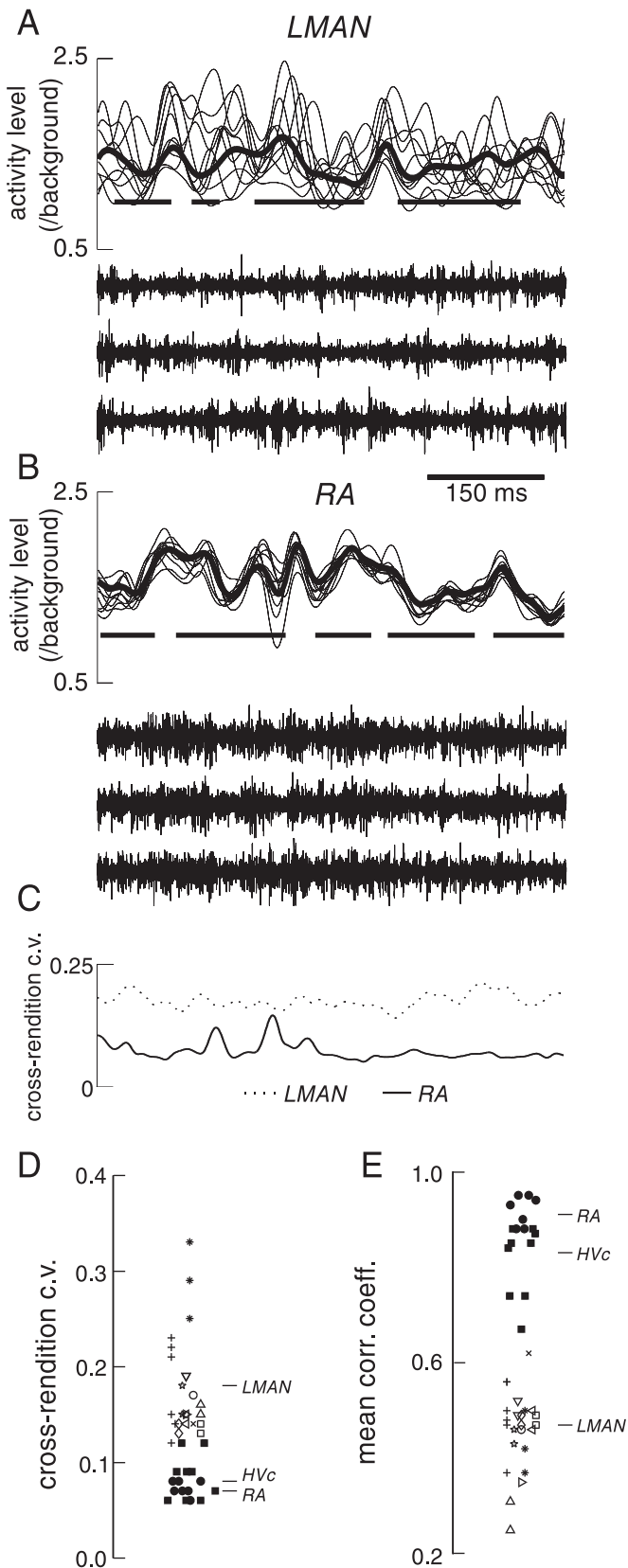


Figure 6. Quantification of response variability across multiple motif renditions. *A, B*, Multiunit activity level during 10 successive motif renditions (*thin lines*) for representative LMAN (*A*, bird dc-17) and RA (*B*, bird xr-1) recording sessions, along with the mean activity level across all renditions over a stationary recording epoch (*thick lines*). Average timing of syllables in the motif is indicated by *thick black lines* placed at $y = 1$.

include greater numbers of neurons. However, these differences existed even between LMAN and RA recording sessions that appeared to be sampling a similar number of units. Furthermore, we have recently shown that differences in stereotypy such as those between RA and LMAN can be detected while recording from a single group of neurons in a single recording session. When birds sang “directed song” aimed at a recipient, the mean cross-rendition c.v. during 13 recording sessions from six birds in LMAN was 0.10 ± 0.03 , significantly different from that during undirected singing at the same sites (mean = 0.16 ± 0.03) (Hessler and Doupe, 1999). The low variability during directed singing is also different from that seen here during undirected singing in LMAN (Fig. 6*D*) (mean 0.18 ± 0.06) and is more similar to the level of cross-rendition variability seen here during undirected singing in HVc and RA (Fig. 6*D*).

The variability in activity between single renditions in LMAN was even more striking when activity of single rather than multiple units was examined. Figure 7 shows an example of one recording session in which a single unit was well isolated. During this period, the bird produced nine renditions of a fairly stereotyped song (Fig. 7*A*) (song elements indicated by the overlying line were repeated in all renditions; there were slight variations in the timing and presence of additional introductory elements, and in one rendition several additional syllables followed the components beneath the line). This unit fired much more strongly during singing than nonsinging periods; the mean firing rate increased from 9.5 to 16.2 spikes/sec. Spikes tended to occur in clusters (the median interspike interval was 8.19 msec during singing). During multiple renditions of the same song elements, there was little reproducibility in the pattern of firing (Fig. 7*B, C*). To address this more closely, we examined the pattern of firing during repeated renditions of a complex syllable that was repeated 45 times during the session. There were some general tendencies for spikes to occur in particular points of the syllable, evident in the mean spike rate (Fig. 7*F*) as well as in the raster plots (Fig. 7*E*). However, the level of reliability of firing was quite low; during some renditions many spikes occurred at times at which none occurred during other renditions (Fig. 7*F*, *asterisk*). These preliminary data suggest that at the level of single units, firing patterns may be even less consistently linked to production of specific song elements than was indicated by multiunit recordings.

Pattern of LMAN activity at different recording sites within a bird

In contrast to such variability in activity level between individual renditions at a single site, the pattern of average multiunit activity related to the song motif was very similar at different recording sites in LMAN of individual birds. In most birds, multiple re-

Neural activity waveforms recorded during three successive renditions of the song motifs are plotted below for both nuclei. *C*, Timing of c.v. (SD/mean) across all renditions for both recordings ($n = 61$ for LMAN, $n = 95$ for RA). *D*, Average c.v. of activity level across all renditions of the stereotyped song motif (recordings from each bird are represented by different symbols). *Empty symbols* and *crossed lines* are used for 24 recording sessions from 10 birds in LMAN; *filled squares* and *circles* are used for 9 recordings from 4 birds in HVc and 7 recordings from 5 birds in RA, respectively. Some symbols have been slightly offset horizontally to reduce overlap between them. Mean c.v. values across all recordings from nuclei are denoted by *lines* to the *right* of symbols. *E*, Mean correlation coefficient of activity level waveform for each individual rendition with the mean activity level across all renditions; *symbols* are as in *D*.

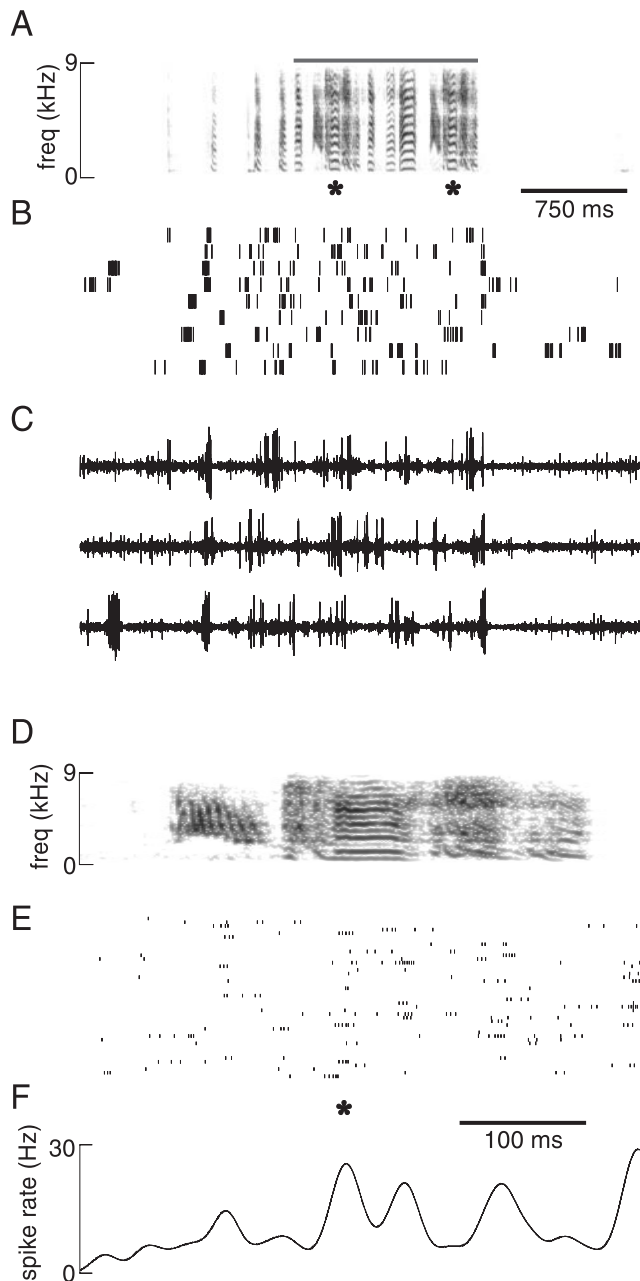


Figure 7. Firing of a single unit in LMAN during multiple song renditions. *A*, Spectrogram of a typical song produced by bird mx-17. Song components beneath *bar* were produced consistently in nine song renditions. *B*, Spike arrival times of action potentials of a single unit during production of these renditions. Songs were aligned so that the amplitude profile of the last syllable in the song matched best across renditions. Stereotypy of timing across the entire song underneath the *bar* remained precise. The SD of song termination time was 2.4 msec, whereas the SD of song initiation time (on average 1290 msec earlier) was increased only to 11.1 msec. *C*, Neural activity waveforms from which the top three spike rasters in *B* were obtained. *D*, Spectrogram of a complex syllable produced 45 times during recording session (* in *A*). Timing of production of this syllable was very stereotyped. Across all renditions, syllable onset time had an SD of 2.9 msec, as did syllable termination time. *E*, Spike raster of single-unit firing during multiple renditions of this stereotyped syllable. *F*, Mean firing rate across all renditions of this syllable; spike arrival times were smoothed with a 10 msec Gaussian window.

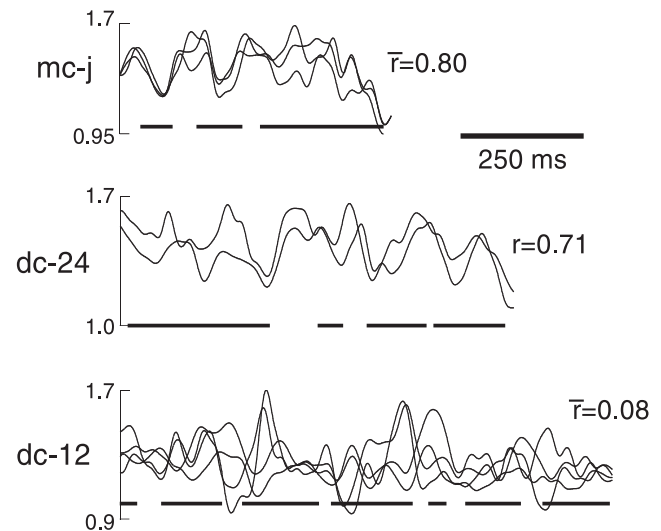


Figure 8. Pattern of multiunit activity at multiple recording sites in individual birds. Mean activity level recorded during production of stereotyped song motifs during three (*mc-j*), two (*dc-24*), and four (*dc-12*) recording sessions at different sites in LMAN. Activity level for each session has been normalized by background activity level during nonsinging periods. Average timing of syllable occurrence for each bird is indicated by *bars* at $y = 1.0$. The average number of motifs per recording session was 279 ± 216 . Correlation coefficients between mean activity waveforms (or averages of these when there are more than 2 sites) are displayed to the *right* for each bird.

Recording sessions were made through the depth of LMAN (successive recording sites were separated by $\sim 80 \mu\text{m}$, except for one pair in bird *mc-j*, which was separated by $\sim 40 \mu\text{m}$) over a period of several days to several weeks. In general, peaks and troughs of average activity profiles appeared in similar locations in all recordings from a single bird (data from three representative birds are displayed in Fig. 8). To quantify the degree of similarity between activity patterns across multiple recording sessions, we calculated correlation coefficients between the mean activity level during different recording sessions from individual birds. For six of seven birds in which recordings were made at more than one site, correlations between response profiles at different sites were high, with a mean across birds of 0.76 ± 0.06 (range 0.66 to 0.82) (Fig. 8, *mc-j*, *dc-24*). For the seventh bird, the pattern was more inconsistent between sessions (Fig. 8, *dc-12*). This inconsistency did not appear to result from the greater number of recording sessions in this bird and thus a slightly increased distance between different recording sites compared with the other birds. The average correlation coefficient for the three comparisons between recording sites separated by $\sim 80 \mu\text{m}$ was 0.12, for two comparisons at $160 \mu\text{m}$ was 0.06, and for one comparison at $240 \mu\text{m}$ was 0.53.

Temporal relationship between LMAN activity level and song elements

For all recordings from LMAN, there was a highly phasic activity profile when neural activity was averaged over multiple motif renditions (Figs. 5*F*, 8). It was difficult, however, when comparing the range of responses across all birds, to ascertain a simple relationship between the mean response profile at a recording site and the complex acoustic structure of the associated song output. This difficulty is attributable in part to differences in the timing and structure of songs of different birds. However, even for three birds that copied their songs from the same tutor and thus sang

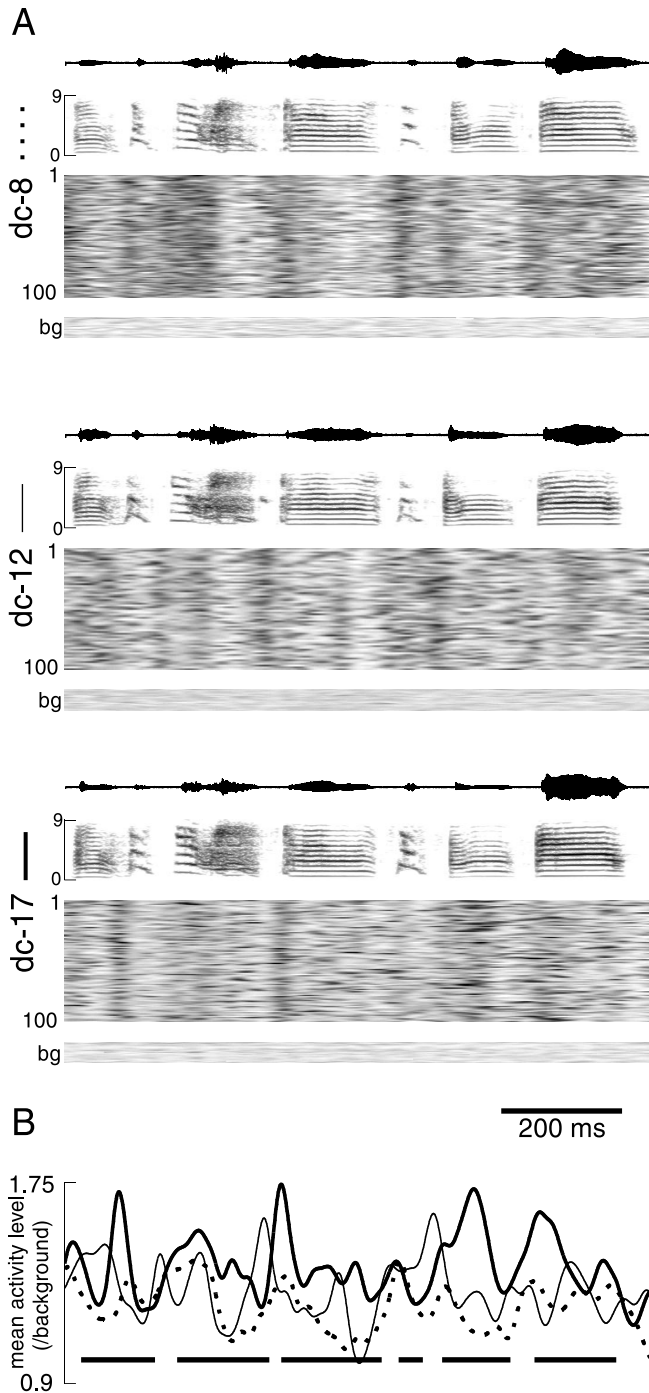


Figure 9. Multiunit activity during similar motifs produced by three different birds. *A*, Oscillograms (top panels) and spectrograms (middle and bottom panels) of stereotyped song motifs of birds dc-8, dc-12, and dc-17, with gray scale plots of multiunit activity level versus time (as in Figs. 4–5) during 100 successive renditions of these song motifs. *bg*, Background nonsinging activity level during period from which activity during motifs was recorded, plotted on the same gray scale axis as activity during motifs. *B*, Mean pattern of activity across 100 renditions of the song motif for each bird, normalized by background activity level. Individual birds are represented by line styles shown at left in *A*. Mean timing of syllable occurrence for bird dc-17 is overlaid for reference. Audio files of representative song motifs for these birds are located at <http://www.keck.ucsf.edu/~neal/jns99/fig9.html>.

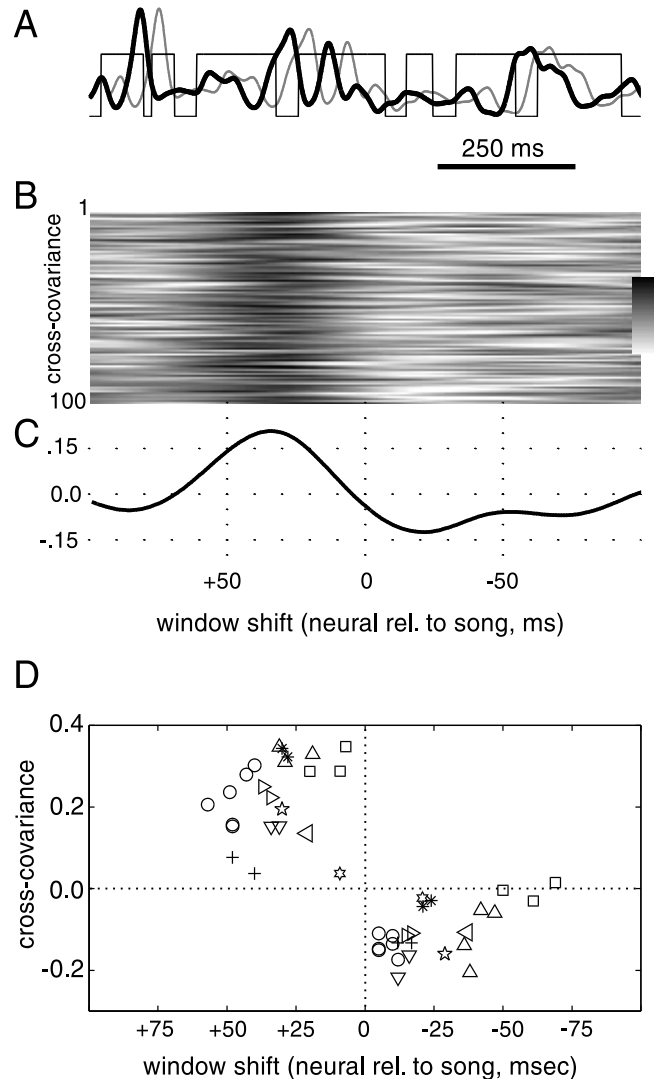


Figure 10. Quantification of relationship of activity patterns to motif elements. *A*, Binary representation of syllables (thin black line) in a representative motif of bird dc-17 (taken from recording session shown in Fig. 9C). The presence of a syllable is represented by a value of 1, the absence of a syllable by a value of 0. Multiunit activity level during one rendition of this song motif is indicated by the thick black line. The thick gray line denotes the same activity trace after it was shifted by +35 msec to the right relative to the syllable trace (i.e., shifting it by the amount that produced the maximum average cross-covariance between activity and song). *B*, Magnitude of cross-covariance for 100 renditions of the motif; for each rendition, the cross-covariance between activity and the discretized song was calculated at successive 1 msec steps, with activity waveform progressively shifted from +100 to –100 msec relative to song. Lower and upper limits of gray scale bar are –0.3 and 0.4. *C*, Mean cross-covariance across all motif renditions. *D*, Summary plot shows peak and trough locations and amplitudes for the mean cross-covariance function in each LMAN recording from every bird (each bird is represented by a different symbol).

very similar motifs (Fig. 9A), average LMAN activity patterns clearly differed from each other (Fig. 9B).

Therefore, we used a simple measure to compare the relationship of activity with song structure across birds singing different songs. We determined when activity tended to occur relative to vocalizations, without consideration of syllable identity. The way in which this was done is summarized in Figure 10A–C, using the data shown in Figure 9A for bird dc-17. For every motif rendition,

the cross-covariance between the activity level trace and a discretized version of the associated song motif (Fig. 10A) was calculated, with the neural activity trace shifted from -100 to $+100$ msec relative to the song (Fig. 10B). Averaged across all renditions (Fig. 10C), this cross-covariance function had a peak when neural activity was shifted forward in time (in the plot, to the right) relative to song by 35 msec (Fig. 10A, gray line) and a trough when activity was shifted backward relative to song by 21 msec. Across all motif renditions, the values of the peak and minimum were significantly different from zero ($p < 0.01$, t test; averaged in a 15 msec window at peak and minimum). Although the mean firing patterns of the birds in the top two panels of Figure 9A were quite different from that for the bird dc-17 analyzed in Figure 10A–C, the mean cross-covariance functions for them also had a peak value when activity was shifted forward relative to songs.

Such a peak for forward activity shifts was seen in all but one recording session from all birds, and in all cases the value of the cross-covariance function at the peak was significantly different from zero across all motif renditions (Fig. 10D) (t test, all $p < 0.01$ for each of 24 recording sessions from 10 birds in which more than 10 motifs were produced, average 325 motifs per session, range 23–750). Over all recordings, the peak location averaged across all renditions in a session was similar to that in Figure 10C, averaging 32 ± 14 msec forward relative to song. For 8 of 10 birds there was also a significant minimum across all renditions when neural activity was shifted backward in time relative to syllables (averaged in a 15 msec window at the minimum; t test, all $p < 0.01$ for each of 20 recording sessions from eight birds). Thus, as in the example in Figure 10A, shifting the neural activity pattern forward relative to the song tended to increase the overlap between peaks of activity and syllables, which resulted in the peak in the mean function. Similarly, shifting neural activity backward relative to song often placed peaks of activity within an intersyllable interval and reduced the covariance between activity and syllables. These results reflect the general tendency (Figs. 5E,F, 8, 9) for peaks of LMAN activity to occur near the beginning of syllables, usually just before them.

As a way of comparing the activity of LMAN and the motor pathway, we performed the same analysis on the relationship of HVC activity to song elements during undirected singing. The mean cross-covariance functions for these recordings were quite similar to those obtained from LMAN. On average, peak locations for HVC recordings were at a shift of 33 ± 5 msec forward (to the right) relative to song elements (range 24–38 msec, $n = 9$ recording sites from four birds). Thus, by this assay, the average temporal relationship between neural activity and song components is similar for HVC and LMAN, although given the variability between birds, it will ultimately be important to compare such analyses in recordings from the same birds.

Singing-related firing of Area X neurons

The AFP nucleus Area X receives the initial input from the song system motor pathway nucleus HVC (Fig. 1B), so if singing-related activity in LMAN originates in HVC it should be present in Area X as well. We therefore investigated whether the firing of neurons in Area X was also modulated during singing. As in LMAN, recordings were made both while birds were silent and while they were producing undirected song.

The background firing characteristics of Area X were quite different from those of LMAN. Although all recordings from within LMAN sampled a grossly similar pattern of multiunit activity, with a range of sizes of spikes, at many sites in Area X no clear neural activity could be distinguished above the recording noise level. At intervals of hundreds of micrometers, however, a fast regular firing pattern of spontaneous activity could be recorded (Fig. 11A). Because these spikes were significantly larger than others recorded by the electrode, single units or clusters of one to two units could be readily discriminated from action potentials of smaller cells, in contrast to the situation in LMAN. The average background spike rate of such regular-firing single units was 160 Hz (range 118–197 Hz, $n = 7$ units), higher than that reported for these neurons in anesthetized birds (the highest rate for adult Area X units was 75 Hz) (Doupe, 1997).

The firing of Area X neurons, despite the differences from LMAN in background level, also changed conspicuously during singing. A representative example of such a response is shown in Figure 11. From a stable fast firing rate of ~ 100 Hz, this neuron increased its firing rate to >200 Hz during singing. Note also, especially in the expanded time scale in the bottom panel, the occurrence of pauses in firing before vocal output (marked by asterisks above the neural activity waveform in Fig. 11B). Overall, increases in firing rate during singing were seen in 11 of 13 single-unit and small-cluster recordings from six birds (Fig. 12A,B). For two sites, however, the mean firing rate during singing was decreased relative to background levels (Fig. 12A,B; one in *mx-11*, one in *dc-24*). For these sites, as was true to a lesser extent for many sites with overall firing increases, pauses of firing were more frequent during singing than during quiet periods (reflected in the decreased 10th percentile of firing rate distribution compared with background).

As was true for LMAN neurons, the activity of Area X neurons was modulated before and after singing. Figure 13 illustrates an example of such modulation for a recording from a small cluster of neurons in Area X. Before song initiation, there was an increased frequency of both pauses and bursts of firing (see also Fig. 11). After song termination, there was a slow decline in firing rate. Such a gradual decrease to background level was different from the postsinging inhibition consistently seen in LMAN recordings (Fig. 4Dc) and was observed for all Area X recordings in which we obtained a sufficient number of silent postsinging epochs for analysis (>10 epochs, $n = 4$ recording sessions from four birds). The mean activity level during the initial 300 msec after song termination for these four recordings was 1.33 ± 0.17 times background activity level (see also Fig. 11; compare with Fig. 4A–C, right panels).

Although Area X neurons always modulated their firing during repeated motifs, the precision of firing relative to song elements during individual song renditions was rather low. The unit shown in Figure 14 greatly increased its firing during the song motif, from a background rate of 115 Hz. It was difficult, however, to discern any tendency for greater firing to occur at any particular time during the motif. In all other recordings from Area X, as well, neural firing appeared more weakly related to individual song elements than in LMAN, although this may in part reflect the smaller number of units contributing to each Area X recording compared with recordings in LMAN. When the cross-covariance function between activity level and song elements was calculated for Area X recordings (as in Fig. 10), there was less

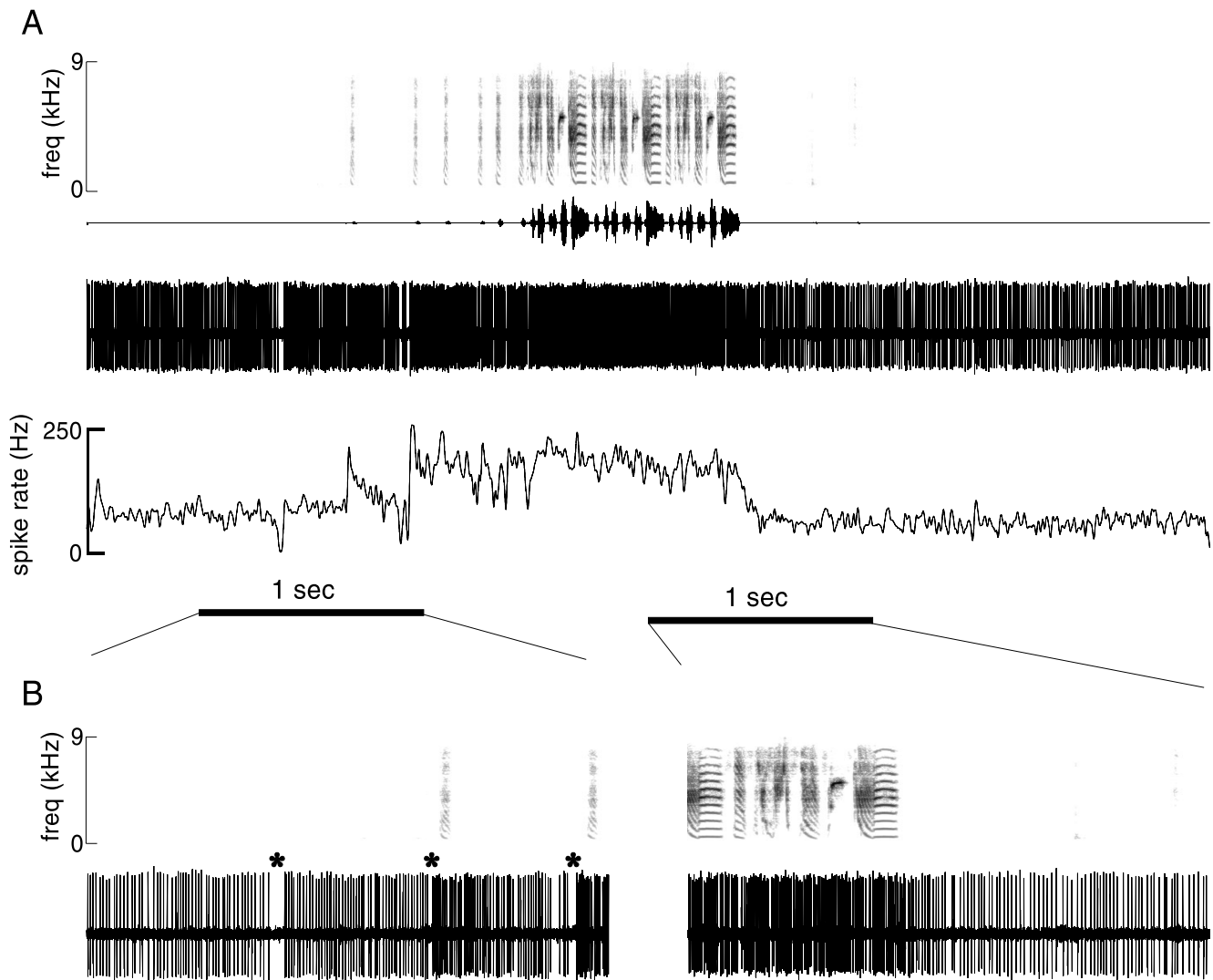


Figure 11. Firing of a single unit in Area X is associated with song production. *A*, Spontaneous singing in a 10 sec period by bird mx-13, represented in spectrogram and oscillogram formats (*top two panels*), along with the concurrently recorded neural activity of a single unit in Area X (*bottom panel*). Below the raw neural waveform is plotted the instantaneous firing rate of this neuron, calculated by convolving spike arrival times with a 10-msec-wide Gaussian window. *B*, Spectrogram of vocalization (*top panel*) and concurrently recorded neural activity (*bottom panel*) from two 1 sec epochs at the beginning (*left*) and end (*right*) of song in *A*, plotted on an expanded time scale. Pauses in neural firing are denoted by asterisks above neural activity waveform. Audio file of song in *A* is located at <http://www.keck.ucsf.edu/~neal/jns99/fig11.wav>.

consistency in locations of maxima and minima across different birds than was seen for LMAN recordings. In nine recording sessions from six birds (mean 87 motif renditions/session, range 13–267), average locations of maxima ranged from -100 to $+68$ msec, and of minima from -100 to $+100$ msec relative to song; this was distinct from the similarity across birds in locations of such peaks and minima for LMAN (from Fig. 10*D*, LMAN peak location = $+32 \pm 19$ msec, mean minima location = -27 ± 19 msec). Furthermore, the pattern of activity appeared even less stereotyped across multiple song renditions than was true for LMAN. For the nine recording sessions from Area X, the mean cross-rendition c.v. was 0.41 ± 0.22 (range 0.2–0.86; by comparison, the mean for all LMAN recording sessions was 0.18 ± 0.05) (Fig. 6*D*). Similarly, for Area X sites the mean correlation coefficient between activity level during individual renditions and mean activity level was 0.26 ± 0.08 (range 0.16–0.39; the mean for all LMAN recording sessions was 0.48 ± 0.08) (Fig. 6*E*).

Influence of auditory feedback on singing-related activity in LMAN

Various indirect evidence suggests that the AFP is involved in processing refferent auditory feedback to the song system during learning. The AFP is required during song acquisition and modification (Bottjer et al., 1984; Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Morrison and Nottebohm, 1993), and in both adult and juvenile anesthetized birds it contains neurons that respond selectively to playback of the bird's own song (Doupe and Konishi, 1991; Doupe, 1997; Solis and Doupe, 1997). To begin to investigate whether the AFP has an auditory role during singing, we used two methods of dissociating the activity associated with the motor act of singing from activity elicited as an auditory response to the bird's self-produced vocalizations. First, we deafened birds to remove the refferent sensory signal during song production, and second, we played back the sound of the

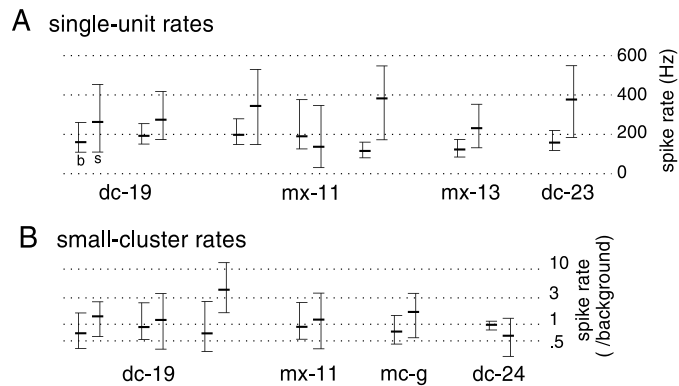


Figure 12. Singing-related modulation of neural firing was seen in all recordings from Area X. *A*, Distributions of instantaneous spike rate for seven single-unit recordings from four birds, recorded during background (*b*, left distribution for each recording) and singing (*s*, right distribution for each recording) periods. In each plot, *upper* and *lower horizontal ticks* denote the 90th and 10th percentiles of the firing rate distribution, whereas the *middle thick tick* denotes the median. *B*, Distributions of instantaneous firing rate for six small cluster recordings from four birds. Individual recordings represented as in *A*. The multiunit firing rate for each recording has been normalized by background nonsinging rate, because the number of neurons contributing to a recording site may have differed between recording sessions. The average duration of singing and background recording periods used were 67 and 38 sec, respectively.

bird's own song outside of the context of singing. Here, we determined the effect of these manipulations on the activity of neurons in LMAN.

Complete removal of auditory feedback does not strongly affect singing-related firing

Complete deafening did not immediately affect song production, as has been observed previously for zebra finches (Nordeen and Nordeen, 1992). An example of this result for a representative bird is shown in Figure 15*A*. In one recording session, this bird produced >300 renditions of a stereotyped seven-syllable song motif (Fig. 15*A*, *hear*). On the next day, when no auditory feedback was possible (after cochleae had been removed), the bird produced a very similar motif (Fig. 15*A*, *deaf*). Fundamental frequencies of syllables with prominent harmonic stacks (syllables *d*, *e*, *f*) were very similar between the two recording sessions (the largest difference in fundamental frequency was 0.4%; from 813 ± 7.9 to 816 ± 7.8 Hz for syllable *f*). Other parameters of syllable structure such as noisiness and timbre (relative weighting of harmonic frequencies within a stack) also did not appear grossly affected by deafening (little difference could be discerned by an experienced listener; for example, compare *hear* and *deaf* song audio files for bird dc-8 at <http://www.keck.ucsf.edu/~neal/jns99/fig15.html>). Moreover, there was little difference in the temporal structure of motifs: although motif duration was significantly shorter 1 d after than just before deafening ($p < 0.01$, *t* test; 345 motif renditions before deafening, 250 after deafening), the magnitude of decrease was very small: from 956 ± 15 to 953 ± 14 msec. Furthermore, this magnitude of difference is within the range of normal daily variability: motif duration 1 d before deafening was 947 ± 15 msec.

Deafening had little immediate effect on motif production in three other birds as well. The most prominent difference was seen in one bird whose motif contained several syllables with simple harmonic stacks. The fundamental frequency of one of four syllables with simple harmonic structure increased significantly

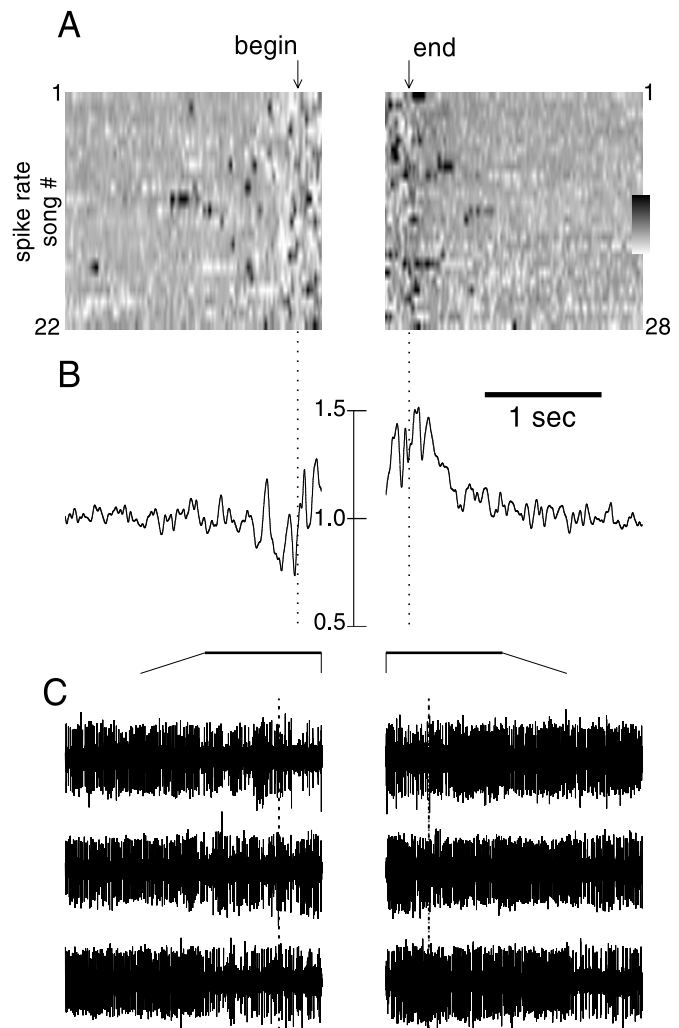


Figure 13. Firing of a small cluster of neurons in Area X is consistently modified before and after singing. *A*, Firing rate during multiple presinging and postsinging intervals of bird mc-g. Instantaneous firing rate was estimated by convolving multiunit spike arrival times with a 10-msec-wide Gaussian window. Firing rate for each song initiation and termination is plotted on successive rows as in Figure 4, with darkness proportional to spike rate. Absolute spike rate has been normalized by background nonsinging spike rate. Lower and upper limits of gray scale bar are 0.2 and 3.0. *B*, Mean firing rate across multiple song initiations and terminations, normalized by background firing rate. Song initiation and termination times are denoted by vertical dotted lines. *C*, Representative neural activity waveforms from which firing rates in *A* and *B* were calculated. Note that the time scale for this panel is expanded from that in *A*; 1 rather than 2 sec of pre-onset and post-termination activity level is displayed. Song initiation and termination times are denoted by vertical dotted lines.

1 d after deafening, from 784 ± 7.4 to 806 ± 6.2 Hz ($p < 0.01$, *t* test; the magnitude of change was 2.7%); this difference was larger than any seen between predeafening recording sessions (song audio files for bird dc-17 are located at <http://www.keck.ucsf.edu/~neal/jns99/fig15.html>). In the other two birds, no audible difference could be detected between predeafening and postdeafening motifs, although detailed analysis of syllable structure was not attempted, because the birds' songs contained only complex noisy syllables (song audio files for birds dc-18 and dc-21 are located at <http://www.keck.ucsf.edu/~neal/jns99/fig15.html>). As in the recording in Figure 15, the temporal structure of motifs for the other three birds was very similar before and after deaf-

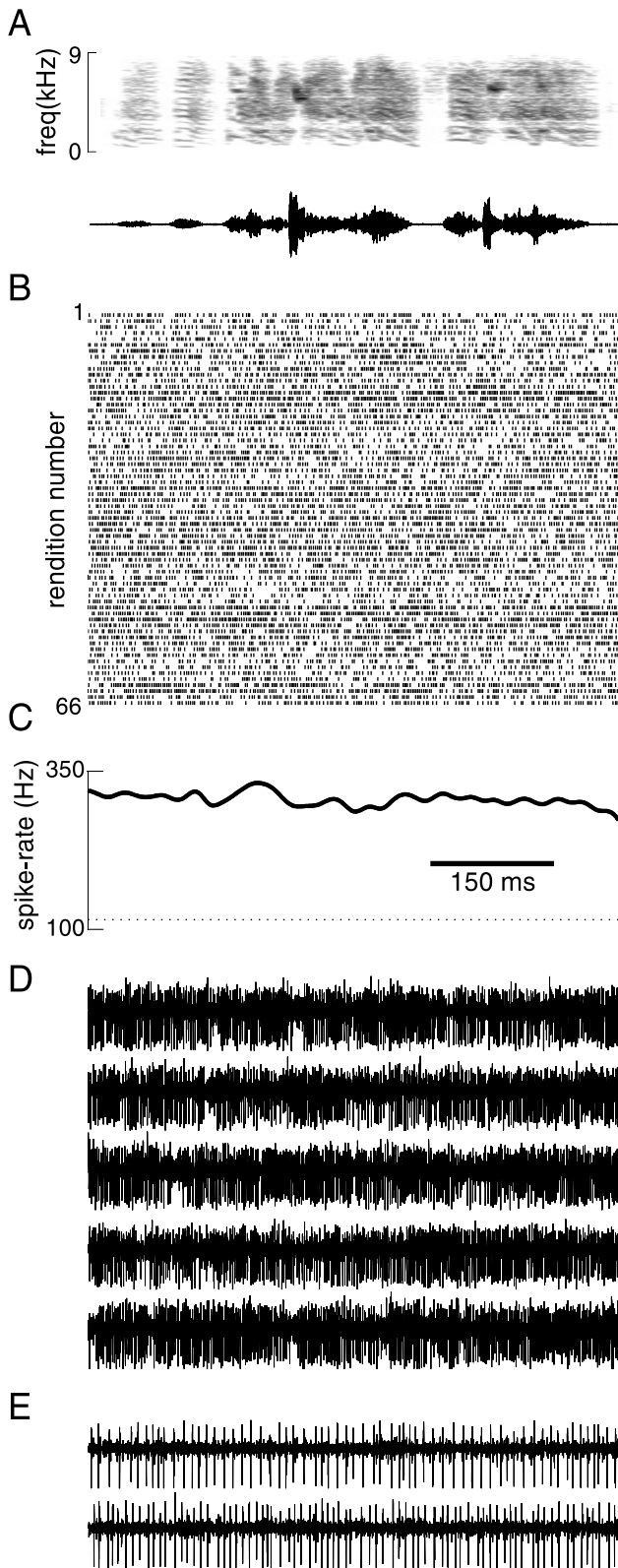


Figure 14. Single-unit firing in Area X during multiple renditions of a stereotyped song motif. *A*, Spectrogram (top panel) and oscillogram (bottom panel) representations of a typical song motif of bird mx-11. *B*, Single-unit spike arrival time raster plot of firing during multiple renditions of this song motif. *C*, Average instantaneous spike rate, made by convolving all spike arrival times with a 10-msec-wide Gaussian window. The mean background nonsinging spike rate is indicated by dotted line. *D*, Successive activity waveforms produced during this motif. *E*, Activity

ening: motif duration of one bird increased by <2% (from 947 to 962 msec), and duration changes were <1% for the remaining two birds. The magnitudes of these changes were not different from those between individual predeafening recording sessions (for four birds, the percentage difference in motif duration ranged from 0.01 to 1.6%). Thus, across all four birds, deafening had little or no immediately detectable effect on song production.

The pattern of multiunit activity in LMAN during song motifs was also not clearly affected by removal of auditory feedback. Because the electrode was retracted above LMAN during the cochlear removal surgery, recordings after deafening were made at slightly different multiunit sites than those before deafening. Nevertheless, there was a consistent pattern of activity in relation to song motifs both before and 1 d after deafening: an example of this result is shown in Figure 15*B*, demonstrating that peaks and troughs of activity occur in similar locations in the motif in both conditions (e.g., peaks before the second syllable *b* and syllable *e*). The average pattern of activity across all renditions of this bird's song (300 renditions for *hear*, 166 for *deaf*) was very similar as well, in both amplitude and time course (Fig. 15*C*).

To characterize the degree to which the pattern of neural activity was affected by deafening, we calculated the correlation coefficients between mean activity patterns before and after deafening and compared them with correlations between mean activity patterns recorded on different days at different sites before deafening. As was shown earlier, average activity patterns recorded at different sites in a given bird were generally similar (Fig. 8). Before deafening, in the recording shown in Figure 15 as well as in three other birds, the correlation coefficients between mean activity patterns at different sites on different days ranged from 0.79 to 0.83 (Fig. 15*D*, *H* vs. *H*). After deafening of these four birds, the mean activity patterns were similar to those recorded while the bird could hear. When the two hearing recordings were compared with the deaf one, the degree of matching was, for seven of eight comparisons, very similar to that between the two hearing recording sessions (Fig. 15*D*). The pattern of activity in one predeafening recording session (from bird shown in Fig. 15*A–C*, 1 d before hearing session shown) was somewhat more different from the session after deafening (Fig. 15*D*, *open circle* with *dot* in center).

These results indicate that a large component of the singing-related firing is not dependent on auditory feedback. However, more precise, reversible manipulations of feedback during singing may be necessary to reveal an acute influence of feedback on AFP activity.

AFP neurons in awake birds can respond selectively to the sound of a bird's own song

The sound of a bird's own song can strongly excite neurons in the AFP and song system motor pathway nuclei of anesthetized adult and juvenile birds (Margoliash, 1983; Doupe and Konishi, 1991; Volman, 1993; Doupe, 1997). Such responses have also been observed, although somewhat less consistently, in motor circuit nuclei of awake birds (McCasland and Konishi, 1981; Margoliash, 1986; Dave et al., 1998; Schmidt and Konishi, 1998). Therefore, while recording in LMAN, we presented the bird with a previ-

←
waveforms of equivalent duration recorded during a nonsinging period. Audio files of four representative song motifs of this bird are located at <http://www.keck.ucsf.edu/~neal/jns99/fig14.html>.

ously recorded version of its own song, during periods when the bird was awake and quiet.

In 5 of 10 birds, song presentation evoked a clear modulation of LMAN neural firing. The largest such response that we recorded is shown in Figure 16. Passive presentation of a two-motif version of the bird's own song (Fig. 16*A*, first and second motifs indicated by *mot-1*, *mot-2*) consistently modulated multiunit neural activity (Fig. 16*B,C*) during this recording session. The acoustically similar first and second motifs elicited a similar pattern of firing on average (Fig. 16*C*).

We quantified this stimulus-evoked modulation of firing by comparing, within trials, both the mean level and amount of modulation of multiunit activity for stimulus versus prestimulus periods (for the recording session displayed in Fig. 16, the period indicated by the *bar* was compared with the entire displayed period preceding it). The comparison of the mean response amplitudes, which is similar to the quantification used in previous studies of song responsiveness (Doupe, 1997; Solis and Doupe, 1997), indicates whether there is a consistent change in activity level. In addition, we quantified changes in modulation of activity level by calculating the within-trial c.v., the SD of activity level sampled at 1 msec intervals, divided by the mean activity level during the period. This measure is sensitive to localized modulations of activity level. For example, a stimulus could elicit an excitatory period followed by an equivalent inhibition, which would be reflected in a comparison of the background versus stimulus within-trial c.v., but not in a comparison of mean levels. For the recording shown in Figure 16, both the mean activity level and within-trial c.v. were significantly higher during song playback than during the period preceding it (Fig. 16*D,E*).

When analyzed within trials, there were significant changes in LMAN activity level during presentation of the bird's own song in eight of nine recording sessions of five birds (comparison performed on data as in Figs. 16*D,E*; paired *t* test, $p < 0.01$, the mean number of stimulus presentations was 18.3, range 9–42). Of these, in three birds, only the mean response level was larger; in one, only the within-trial c.v. was larger; and in one (one of the two recording sessions for this bird is shown in Fig. 16), both the mean and within-trial c.v. were larger during the stimulus presentations than during the preceding background periods. For the recordings in which mean responses during stimulus presentation were significantly larger, the average ratio of activity level during the stimulus to that preceding it was 1.05 ± 0.02 ($n = 7$ recording sites), whereas for recordings from the five birds in which no response was detected, the mean ratio was 1.02 ± 0.01 ($n = 9$ recording sites). For recordings in which the within-trial c.v. differed, the average c.v. during the stimulus was 1.3 ± 0.12 times larger than that during the prestimulus period; for others this ratio averaged 1.07 ± 0.17 .

Besides testing for responses evident within single trials, we tested whether in some cases modulation of activity could be seen only when the response profile was averaged over all presentations. Such a case could arise, for example, if a response too small to be seen on individual trials occurred consistently at one point during the stimulus. To test for this, we calculated the mean response profile for prestimulus and stimulus periods (as in Fig. 16*C*), divided each profile into 10 intervals of equal duration, and compared the distributions of mean activity level in the 10 intervals between stimulus and prestimulus periods. The binned distribution of mean activity level during the stimulus was significantly different from that preceding the stimulus for seven recording sessions in which significant responses had been de-

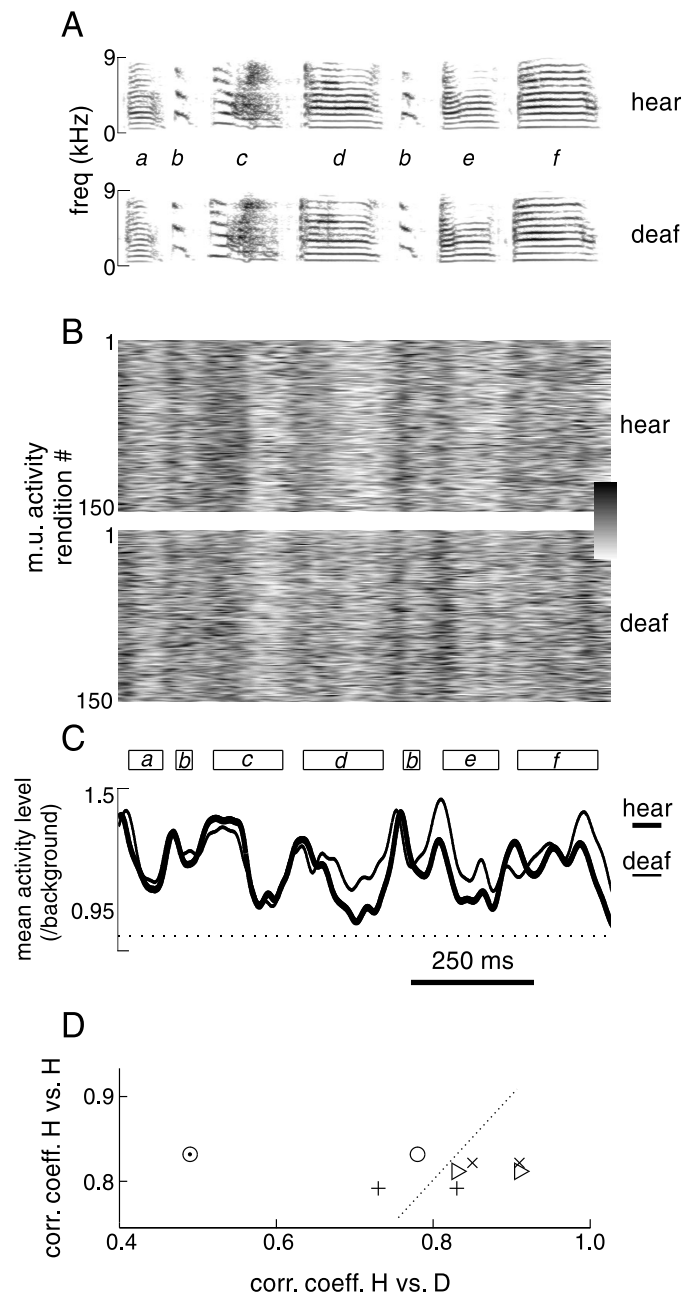


Figure 15. LMAN activity during singing before (while hearing) and after deafening. *A*, Spectrograms of song motifs produced by bird dc-8 several hours before (top panel, *hear*) and 1 d after (bottom panel, *deaf*) deafening. Syllable identity is indicated by letters between spectrograms. *B*, Multiunit activity level in LMAN during a representative 150 renditions of the song motif, during a recording session when the bird could hear (top panel), and after deafening (bottom panel). Activity level for both recordings was normalized by average background activity level. Lower and upper limits of gray scale bar are 0.6 and 2.3. *C*, Mean activity level during 300 renditions of motif before deafening (thick trace) and 161 renditions of motif after deafening (thin trace). Average timing of syllables during hearing recording session is displayed above mean traces. *D*, Each symbol represents the correlation coefficient between mean activity level profiles of two predeafening recording sessions, plotted against the correlation coefficient between each predeafening session and the postdeafening session for that bird (each bird represented by a different symbol). The dotted line indicates where symbols would lie if correlation coefficients between predeafening sessions were equal to those between predeafening and postdeafening sessions. Audio files of song motifs of the four birds before and immediately after deafening are located at <http://www.keck.ucsf.edu/~neal/jns99/fig15.html>.

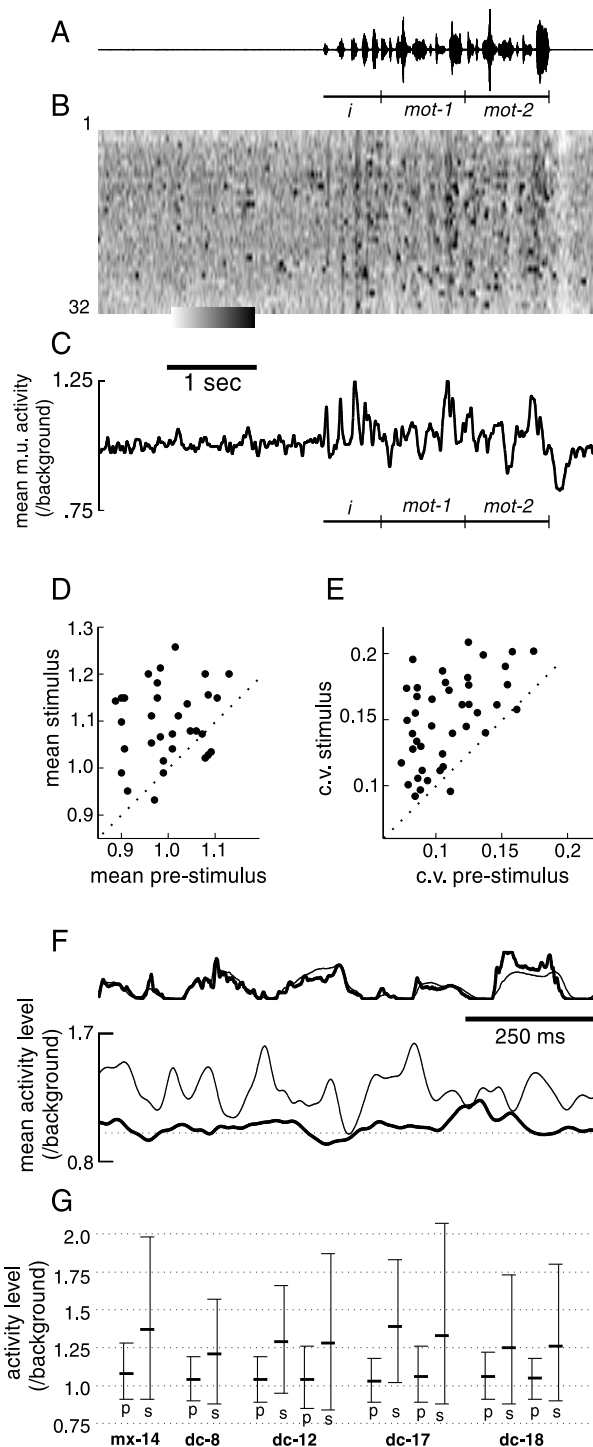


Figure 16. LMAN activity evoked by passive playback of bird's own song. *A*, Oscillogram of acoustic stimulus played from a speaker to bird dc-12. Introductory elements (*i*) and two repeated song motifs (*mot-1* and *mot-2*) are denoted by *bar* below oscillogram. *B*, Multiunit activity level in LMAN during multiple trials of song playback, normalized by background nonsinging activity level. Lower (*left*) and upper (*right*) limits of gray scale bar are 0.65 and 1.7. *C*, Mean activity level across all playback trials displayed in *B*. *D*, Mean activity level for each trial during the stimulus, plotted versus mean activity level in the period 2.5 sec preceding stimulus onset. *E*, For each trial, within-trial c.v. during the stimulus period plotted versus within-trial c.v. during the prestimulus period. Dotted lines in *D* and *E* indicate where points would lie if means and c.v. values were equal between stimulus and prestimulus periods. *F*, *Top traces*, Mean amplitude of song motifs recorded when played from the

speaker (*thick line*; both first and second motifs of 32 playback trials were averaged) and when sung during same recording session by the bird (*thin line*; 74 motif renditions interspersed with song playback trials). *Bottom traces*, Mean neural activity level recorded when song motif was played from the speaker (*thick line*, responses to first and second motifs averaged) and when the same motif was produced by the bird (*thin line*). *G*, Distribution of activity level sampled at 1 msec intervals during song playback (*p*) and singing (*s*) for the eight recording sessions from five birds in which there were significant playback responses within individual trials. Activity levels were normalized by background nonsinging activity level. *Top*, *middle*, and *bottom ticks* denote 90th, 50th, and 10th percentiles of activity level distributions, respectively.

tected on single trials, and in two additional recording sessions ($p < 0.01$, Kolmogorov–Smirnov test). One of the additional responses detected by this assay was in the one session from the five birds in which the within-trial response was not significant; the other was an additional recording session from a sixth bird. Overall, the mean ratio of c.v. during the stimulus to c.v. during the prestimulus period for these recording sessions was 2.4 ± 0.07 ($n = 8$). For other recording sessions (in which significant responses were not seen), this ratio was 1.28 ± 0.5 ($n = 8$; mean number of stimulus presentations for recording sessions with significant response = 20.5, mean number for those with no significant response = 15.4).

For the eight recording sessions in which responses were apparent to single presentations and one of two in which the response appeared after averaging multiple trials, modulation of the mean response profile was present throughout the stimulus, as for the response in Figure 16*C*. Such response profiles are similar to those reported for LMAN responses of anesthetized birds (Doupe, 1997). For the remaining recording session from the sixth bird in which responses were evident only on averaging multiple trials, however, the only modulation from background firing was a peak near the onset of the stimulus. No difference in activity level was seen when the prestimulus period was compared with the final two-thirds of the stimulus for this recording; this may indicate that the onset peak was a nonspecific arousal response.

Initial results suggest that, as is true in anesthetized birds (Doupe and Konishi, 1991; Vicario and Yohay, 1993), responses of LMAN neurons in awake birds are selective to their own song. Two of the five birds with sustained responses to the bird's own song were presented with the acoustically distinct song of an unrelated bird and responded significantly less strongly to it than to their own song ($p < 0.01$, *t* test).

Although responses to song playback were distinct from background firing, overall activity level during playback was lower than that when the bird itself produced very similar sounds. For the recording with the largest response, the mean activity level associated with production of a repeated motif was much higher than that during passive presentation of the motif (Fig. 16*F*) (*thick line* = 64 motifs for playback, *thin line* = 74 motifs for singing). It was also difficult to discern a clear relationship between the activity pattern during passive presentation of the motif with that during active production of the same motif. Note, however, that the difference in mean response level between singing and playback periods results to some degree from the large trial-to-trial variability in the magnitude of responses to song playback (Fig. 16*D*). The largest responses to single song presentations were similar in magnitude to those during singing. In all recording sessions in which significant responses to song

←

speaker (*thick line*; both first and second motifs of 32 playback trials were averaged) and when sung during same recording session by the bird (*thin line*; 74 motif renditions interspersed with song playback trials). *Bottom traces*, Mean neural activity level recorded when song motif was played from the speaker (*thick line*, responses to first and second motifs averaged) and when the same motif was produced by the bird (*thin line*). *G*, Distribution of activity level sampled at 1 msec intervals during song playback (*p*) and singing (*s*) for the eight recording sessions from five birds in which there were significant playback responses within individual trials. Activity levels were normalized by background nonsinging activity level. *Top*, *middle*, and *bottom ticks* denote 90th, 50th, and 10th percentiles of activity level distributions, respectively.

playback were seen, mean activity during song playback was of much lower magnitude than singing-related activity. Across all sessions, the mean activity level during singing was 1.23 times higher than that during song playback (Fig. 16G).

DISCUSSION

This study provides the first detailed description of singing-related activity in the anterior forebrain nuclei LMAN and Area X. Several features of this activity indicate that it may derive in part from singing-related firing of the song system motor pathway nucleus HVC and provide evidence for a functional similarity of the AFP to mammalian cortical–striatal motor systems. Initial experiments suggest that auditory feedback has little acute influence on the singing-related AFP activity. These results support the idea that the AFP has some function related to adult song production, perhaps serving as an efference copy of motor command signals from HVC (Troyer et al., 1995).

Singing-related activity of anterior forebrain nuclei

Although complete removal of AFP nuclei does not affect song production in adult zebra finches (Nottebohm et al., 1976; Bottjer et al., 1984; Nordeen and Nordeen, 1993), we show here that this pathway is electrophysiologically engaged during singing. At least some of this activity must be independent of auditory feedback, because it begins before vocal output. Singing-related AFP activity may originate in the motor pathway nucleus HVC, in which similar presong activity is seen, and which provides the major input to the AFP (Fig. 1). Direct confirmation of this idea, however, will require further studies, for instance an analysis of AFP singing-related activity before and after severing of the output from HVC to Area X. Our demonstration of neurophysiological activity of the AFP during singing is consistent with a previous report describing induction of an immediate early gene in LMAN and Area X during singing (Jarvis and Nottebohm, 1997; Jin and Clayton, 1997).

Although the overall timing of AFP activity relative to singing is similar to that of the motor pathway, the pattern of AFP firing across repeated song renditions is less consistent than that of motor pathway nuclei [for single-unit recordings, Yu and Margoliash (1996); for multiple-unit recordings, Vu et al. (1998); and this report]. The differences in variability between the two pathways may reflect their different contributions to the control of adult song production. Motor pathway nuclei are critical for song production throughout life (Nottebohm et al., 1976; Vu et al., 1994) and are presumably involved in generating the motor commands that give rise to precisely repeated song elements. In contrast, lesions of AFP nuclei disrupt motor output only in young birds and not in normal adult finches (Bottjer et al., 1984; Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Nordeen and Nordeen, 1993). Thus, the relationship of AFP activity to adult song output must be more subtle than is true for motor pathway nuclei.

Although firing in the AFP is more variable than that of the motor pathway, in some behavioral contexts it can approach the stereotypy of the motor nuclei. In the current report, the high level of rendition-to-rendition variability in the pattern of AFP activity occurred while birds sang in isolation, during production of undirected song. In our previous study (Hessler and Doupe, 1999), when birds sang directed song, which is addressed to another bird, the level of rendition-to-rendition variability was much closer to that shown here for activity in HVC and RA. The differences in AFP activity in the two social contexts may also

explain a discrepancy between the results presented here and those of an earlier study that reported no modulation of AFP activity during singing (McCasland, 1987). Unlike the clear changes in singing-related firing relative to background activity levels during undirected singing, firing during directed singing, which was the sole type of behavior monitored in the previous paper (McCasland, 1987; J. S. McCasland, personal communication), can be much less different from background activity levels (Hessler and Doupe, 1999).

In a given bird, the average pattern of activity related to the song motif tends to be similar across multiple LMAN recording sites, unlike the markedly distinct firing patterns across RA (Yu and Margoliash, 1996). If this similarity holds for neurons throughout LMAN, the entire nucleus would send a uniform singing-related signal to the motor pathway. Alternatively, because multiple recording sites from single birds were always in the same medial–lateral plane, we may have sampled activity from single topographic groups in LMAN (Johnson et al., 1995).

Relationship of the anterior forebrain pathway to cortical–basal ganglia loops in mammals

The song system is a specialization of avian circuits that are homologous in various ways to mammalian higher motor control systems. The songbird motor pathway nuclei may be comparable to mammalian cortical areas that are directly involved in precise motor control. For instance, RA, like primary motor cortex, projects directly onto brainstem motor pattern generators as well as onto cranial nerve motor nuclei (Nottebohm et al., 1976; Vicario, 1991; Wild, 1997). Anatomical evidence suggests that, in contrast, the AFP is more similar to mammalian “indirect” motor circuits. Specifically, the HVC/LMAN→Area X→DLM→LMAN loop may be related to the mammalian cortex→basal ganglia→thalamus→cortex loops (Bottjer and Alexander, 1995; Luo and Perkel, 1999a,b) (for review, see Alexander et al., 1986; Parent and Hazrati, 1995; Veenman, 1997).

The results presented here provide physiological support for this model. As is true for many mammalian cortical and striatal regions (for review, see DeLong and Georgopoulos, 1981; Georgopoulos, 1995), neurons in LMAN and Area X are strongly active preceding and during specific movements. Supporting the idea that Area X contains a pallidal component (Luo and Perkel, 1999b), the activity of the Area X cells that we recorded resembles that of regular-firing mammalian pallidal/nigral projection neurons, in both its high tonic level and its relationship to movement (DeLong, 1971; Hikosaka and Wurtz, 1985). If these Area X neurons are indeed pallidal, their singing-related pauses in firing could be caused by inhibitory input from other, more striatal-like cells within Area X and could thus result in intermittent disinhibition of thalamic neurons and subsequent excitation of LMAN neurons (Luo and Perkel, 1999b) [see Chevalier and Deniau (1990) for discussion of the analogous mammalian circuit]. It is not surprising that this avian dorsal forebrain–basal ganglia circuitry is involved in the complex process of song learning and production, because its mammalian counterpart has been implicated in various movement sequencing, motor learning, and reward processes (for review, see Kimura, 1995; Mink, 1996; Schultz et al., 1998).

Possible functions of singing-related activity in the AFP of adults

Motor or sensorimotor control: efference copy

Overall, the characteristics of the AFP activity that we observed raise the possibility that it is a copy of the motor commands for

song that are also sent from HVC to RA. Such efference copy signals are common in sensorimotor control systems and can be useful for providing information about intended motor activity to multiple areas of the brain. Efference copies can be used to modulate sensory information, compensating for reafferent signals that arise as a consequence of the animal's own behavior (von Holst and Mittelstaedt, 1950; Bell, 1989; Bridgeman, 1995), or used during motor learning as internal predictions that are compared with a target reference (Jordan, 1995). These ideas suggest that monitoring singing-related AFP activity during systematic alteration of sensory feedback could be very informative. The subtle effects of deafening described here, however, suggest that AFP activity, unlike that in cerebellar systems (Bell et al., 1997), does not include a large component that is a negative image of expected sensory feedback.

Adult auditory feedback

After the process of learning culminates in crystallized song, song production can remain stable for years thereafter. This has been thought to reflect the presence of a stable central motor program for song, i.e., a complex anatomical/physiological network that can continue producing the same vocal output, independent of external guidance (Konishi, 1965; Nottebohm, 1968). However, it has recently become clear that the production of stable song in adult zebra finches is not completely "hard-wired," but instead is actively maintained. When auditory feedback is removed or altered in adult zebra finches, song output degrades gradually, as does human speech after adult deafening [for finches, see Nordeen and Nordeen (1992); Leonardo and Konishi (1999); for humans, see Waldstein (1989); Cowie and Douglas-Cowie (1992)]. Active maintenance of song by auditory feedback may be required for birds to adapt to normally occurring peripheral and central changes in the network responsible for song production (Alvarez-Buylla et al., 1988).

The AFP in adults could process auditory feedback of song and transmit it to motor nuclei, much as this circuit has been proposed to do in young birds (Bottjer and Arnold, 1986; Mooney, 1992; Doupe, 1993). The lack of effect of LMAN lesions on adult song production, in contrast to the effects of deafening, has been interpreted as evidence against such a function for adult LMAN (Nordeen and Nordeen, 1993). The results described here, however, serve as a reminder of the important functional difference between LMAN lesions and deafening. The neural output from the AFP to RA is necessarily eliminated in birds with LMAN lesions, whereas the AFP is strongly active during singing and thus may interact with the motor pathway in both hearing and deafened birds. Indeed, several recent studies suggest that changes in adult song caused by deafening (M. Brainard, unpublished results) or motor nerve damage (Williams and Mehta, 1999) can be blocked by lesions of LMAN.

In this context, the effects of deafening on LMAN activity constrain hypotheses about the function of the AFP in adult song maintenance. Acute effects of deafening on LMAN activity are not dramatic. If the modulation caused by sensory feedback from the bird's own voice is of the same order of magnitude as the small responses to auditory playback of song, however, it may be obscured by the prominent singing-related activity. Moreover, because song deteriorates very slowly in deafened adult finches, it may be that although auditory feedback must change immediately after deafening, effects on signals in the AFP and subsequently on song emerge only gradually.

Sensory or perceptual function of adult LMAN

We show here that as is true in anesthetized birds (Doupe and Konishi, 1991; Vicario and Yohay, 1993), playback of the sound of a bird's own song can selectively excite LMAN neurons in birds that are awake. These responses likely originate in similar responses in HVC of awake birds (McCasland and Konishi, 1981; McCasland, 1987; Dave et al., 1998). Because selective auditory responses in HVC and RA are dependent on the precise behavioral state of the animal (Dave et al., 1998; Schmidt and Konishi, 1998), the lack of detectable responses in LMAN of 50% of birds tested may have been caused by some subtle differences in arousal or attentional state between birds.

Several general classes of function have been proposed for neurons that respond selectively to the sound of a bird's own song. Such responses could play a role in the perception of songs of other birds, allowing an individual to classify songs of other birds on a metric of similarity to its own song (Margoliash, 1986). Such a function is supported by the recent report that lesions of AFP nuclei can disrupt performance on a song discrimination task, especially when a bird must distinguish its own song from that of a conspecific (Scharff et al., 1998). Alternatively, or in addition, selectivity for a bird's own song could play a role in the processing of auditory feedback during song production. In this context, because sensory input can be "gated" by behavior in many motor systems (Suga and Schlegel, 1972; Creutzfeldt et al., 1989) (for review, see Pearson, 1993; Nelson, 1996), it is possible that responses to passive presentation of song do not accurately reflect the influence that auditory feedback of a bird's own voice has on AFP activity during singing.

Modulation of AFP activity and behavior by social context

In LMAN and Area X, both singing-related neural activity and immediate early gene expression differ greatly depending on the social context in which the bird sings (Jarvis et al., 1998; Hessler and Doupe, 1999). This raises the possibility that social cues act via the AFP to modulate song, perhaps enhancing plasticity in the motor system during undirected compared with directed singing (Jarvis et al., 1998). Consistent with this idea, we show here that the output of LMAN onto RA is quite variable across renditions during undirected singing; such "noise" might facilitate plasticity in the HVC–RA circuitry. Alternatively, differences in activity and gene expression could reflect differences in communicative function between the two types of singing (Dunn and Zann, 1996a,b).

In conclusion, the AFP is strongly active during singing in adult birds, suggesting that it maintains some function during adult singing. Because this circuit is also crucially involved in song learning, its role will be further elucidated by studies in young birds. Moreover, because the AFP contains a basal ganglia component that has apparently been specialized for one particular and discrete motor behavior, singing, this system could have marked advantages for providing insight into the very general functions performed by basal ganglia circuits in many vertebrates.

REFERENCES

- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381.
- Alvarez-Buylla A, Theelen M, Nottebohm F (1988) Birth of projection neurons in the higher vocal center of the canary forebrain before, during, and after song learning. *Proc Natl Acad Sci USA* 85:8722–8726.

- Basham ME, Nordeen EJ, Nordeen KW (1996) Blockade of NMDA receptors in the anterior forebrain impairs sensory acquisition in the zebra finch. *Neurobiol Learn Mem* 66:295–304.
- Bell C (1989) Sensory coding and corollary discharge effects in mormyrid electric fish. *J Exp Biol* 146:229–253.
- Bell C, Bodznick D, Montgomery J, Bastian J (1997) The generation and subtraction of sensory expectations within cerebellum-like structures. *Brain Behav Evol* 50 [Suppl 1]:17–31.
- Bottjer SW, Alexander G (1995) Localization of met-enkephalin and vasoactive intestinal polypeptide in the brains of male zebra finches. *Brain Behav Evol* 45:153–177.
- Bottjer SW, Arnold AP (1986) The ontogeny of vocal learning in songbirds. In: *Handbook of behavioral neurobiology*, Vol 8 (Blass EM, ed), pp 129–161. New York: Plenum.
- Bottjer SW, Meisner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224:901–903.
- Bottjer SW, Halsema KA, Brown SA, Meisner EA (1989) Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *J Comp Neurol* 279:312–326.
- Brenowitz EA (1997) Comparative approaches to the avian song system. *J Neurobiol* 33:517–531.
- Bridgeman B (1995) A review of the role of efference copy in sensory and oculomotor control systems. *Ann Biomed Eng* 23:409–422.
- Buzsaki G, Bickford RG, Ryan LJ, Young S, Prohaska O, Mandel RJ, Gage FH (1989) Multisite recording of brain field potentials and unit activity in freely moving rats. *J Neurosci Methods* 28:209–217.
- Catchpole CK, Slater PJB (1995) *Bird song: biological themes and variations*. Cambridge, UK: Cambridge UP.
- Chevalier G, Deniau J (1990) Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci* 13:277–280.
- Cowie R, Douglas-Cowie E (1992) Postlingually acquired deafness: speech deterioration and the wider consequences. Berlin: Mouton de Gruyter.
- Creutzfeldt O, Ojemann G, Lettich E (1989) Neuronal activity in the human lateral temporal lobe. II. Responses to the subject's own voice. *Exp Brain Res* 77:476–489.
- Dave A, Yu AC, Margoliash D (1998) Behavioral state modulation of auditory activity in a vocal motor system. *Science* 282:2250–2254.
- DeLong MR (1971) Activity of pallidal neurons during movement. *J Neurophysiol* 34:414–427.
- DeLong MR, Georgopoulos AP (1981) Motor functions of the basal ganglia. In: *Handbook of physiology*. (Brooks VB, ed), pp 1017–1061. Bethesda, MD: American Physiological Society.
- Doupe AJ (1993) A neural circuit specialized for vocal learning. *Curr Opin Neurobiol* 3:104–111.
- 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, Konishi M (1991) Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad Sci USA* 88:11339–11343.
- Doupe AJ, Kuhl P (1999) Birdsong and human speech. *Annu Rev Neurosci* 22:567–631.
- Dunn AMZ, Zann RA (1996a) Undirected song in wild zebra finch flocks: contexts and effects of mate removal. *Ethology* 102:529–539.
- Dunn AMZ, Zann RA (1996b) Undirected song encourages the breeding female zebra finch to remain in the nest. *Ethology* 102:540–548.
- Georgopoulos A (1995) Current issues in directional motor control. *Trends Neurosci* 18:506–510.
- Hessler NA, Doupe AJ (1999) Social context modulates singing-related neural activity in the songbird forebrain. *Nat Neurosci* 2:209–211.
- Hikosaka O, Wurtz RH (1985) Modification of saccadic eye movements by GABA-related substances. I. Effect of muscimol and bicuculline in monkey superior colliculus. *J Neurophysiol* 53:266–291.
- Jarvis ED, Nottebohm F (1997) Motor-driven gene expression. *Proc Natl Acad Sci USA* 94:4097–4102.
- Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F (1998) For whom the bird sings: context-dependent gene expression. *Neuron* 21:775–788.
- Jin H, Clayton DF (1997) Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. *Neuron* 19:1049–1059.
- Johnson F, Sablan M, Bottjer SW (1995) Topographic organization of a pathway involved with vocal learning in zebra finches. *J Comp Neurol* 358:268–275.
- Jordan MI (1995) Computational motor control. In: *The cognitive neurosciences* (Gazzaniga M, ed), pp 567–610. Cambridge, MA: MIT.
- Kimura M (1995) Role of basal ganglia in behavioral learning. *Neurosci Res* 22:353–358.
- Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z Tierpsychol* 22:770–783.
- Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* 399:466–470.
- Luo M, Perkel DJ (1999a) Long-range GABAergic projection in a circuit essential for vocal learning. *J Comp Neurol* 403:68–84.
- Luo M, Perkel DJ (1999b) A GABAergic, strongly inhibitory projection to a thalamic nucleus in the zebra finch song system. *J Neurosci* 19:6700–6711.
- 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.
- McCasland JS (1987) Neuronal control of bird song production. *J Neurosci* 7:23–39.
- McCasland JS, Konishi M (1981) Interactions between auditory and motor activities in an avian song control nucleus. *Proc Natl Acad Sci USA* 78:7815–7819.
- Mink J (1996) The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 50:381–425.
- Mooney R (1992) Synaptic basis for developmental plasticity in a bird-song nucleus. *J Neurosci* 12:2464–2477.
- Morrison RG, Nottebohm F (1993) Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *J Neurobiol* 24:1045–1064.
- Nelson R (1996) Interactions between motor commands and somatic perception in sensorimotor cortex. *Curr Opin Neurobiol* 6:801–810.
- Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol* 57:58–66.
- 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 (1968) Auditory experience and song development in the chaffinch (*Fringilla coelebs*). *Ibis* 11:549–568.
- Nottebohm F, Stokes TF, Leonard CM (1976) Central control of song in the canary, *Serinus canarius*. *J Comp Neurol* 165:457–486.
- Okuhata S, Saito N (1987) Synaptic connections of thalamo-cerebral vocal control nuclei of the canary. *Brain Res Bull* 18:35–44.
- Parent A, Hazrati L-N (1995) Functional anatomy of the basal ganglia. In: *The cortico-basal ganglia-thalamo-cortical loop*. *Brain Res Brain Res Rev* 20:91–127.
- Pearson K (1993) Common principles of motor control in vertebrates and invertebrates. *Annu Rev Neurosci* 16:265–297.
- Scharff C, Nottebohm F (1991) Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *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.
- Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. *Nat Neurosci* 1:513–518.
- Schultz W, Tremblay L, Hollerman JR (1998) Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology* 37:421–429.
- 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, Doupe AJ (1997) Anterior forebrain neurons develop selectivity by an intermediate stage of song learning. *J Neurosci* 17:6447–6462.
- Solis MM, Doupe AJ (1999) Contributions of tutor and bird's own song experience to neural selectivity in the songbird anterior forebrain. *J Neurosci* 19:4559–4584.
- Suga N, Schlegel P (1972) Neural attenuation of responses to emitted sounds in echo-locating bats. *Science* 177:82–84.

- Troyer TW, Doupe AJ, Miller KD (1995) An associational hypothesis for sensorimotor learning of birdsong. In: Computational neuroscience: trends in research (Bower J, ed), pp 409–414. New York: Academic.
- Veenman C (1997) Pigeon basal ganglia: insights into the neuroanatomy underlying telencephalic sensorimotor processes in birds. *Eur J Morphol* 35:220–233.
- Vicario DS (1991) Organization of the zebra finch song control system. II. Functional organization of outputs from nucleus Robustus archistriatalis. *J Comp Neurol* 309:486–494.
- 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.
- von Holst E, Mittelstaedt H (1950) Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften* 37:464–476.
- Vu ET, Mazurek ME, Kuo YC (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14:6924–6934.
- Vu ET, Schmidt MF, Mazurek ME (1998) Interhemispheric coordination of premotor neural activity during singing in adult zebra finches. *J Neurosci* 18:9088–9098.
- Waldstein R (1989) Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *J Acoust Soc Am* 88:2099–2114.
- Wild JM (1997) Neural pathways for the control of birdsong. *J Neurobiol* 33:653–670.
- 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.
- Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. *Science* 273:1871–1875.
- Zann R (1996) *The zebra finch*. Oxford, UK: Oxford UP.