

# Contributions of Tutor and Bird's Own Song Experience to Neural Selectivity in the Songbird Anterior Forebrain

Michele M. Solis and Allison J. Doupe

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

Auditory neurons of the anterior forebrain (AF) of zebra finches become selective for song during song learning. In adults, these neurons respond more to the bird's own song (BOS) than to the songs of other zebra finches (conspecifics) or BOS played in reverse. In contrast, AF neurons from young birds (30 d) respond equally well to all song stimuli. AF selectivity develops rapidly during song learning, appearing in 60-d-old birds. At this age, many neurons also respond equally well to BOS and tutor song. These similar neural responses to BOS and tutor song might reflect contributions from both song experiences to selectivity, because auditory experiences of both BOS and tutor song are essential for normal song learning. Alternatively, they may simply result from acoustic similarities between BOS and tutor song. Understanding which experience shapes selectivity could elucidate the function of song-selective AF neurons.

To minimize acoustic similarity between BOS and tutor song, we induced juvenile birds to produce abnormal song by denervating the syrinx, the avian vocal organ, before song onset. We

recorded single neurons extracellularly in the AF at 60 d, after birds had had substantial experience of both the abnormal BOS (tsBOS) and tutor song. Some neurons preferred the unique tsBOS over the tutor song, clearly indicating a role for BOS experience in shaping neural selectivity. In addition, a sizable proportion of neurons responded equally well to tsBOS and tutor song, despite their acoustic dissimilarity. These neurons were not simply immature, because they were selective for tsBOS and tutor song relative to conspecific and reverse song. Furthermore, their similar responses to tsBOS and tutor song could not be attributed to residual acoustic similarities between the two stimuli, as measured by several song analyses. The neural sensitivity to two very different songs suggests that single AF neurons may be shaped by both BOS and tutor song experience.

**Key words:** auditory selectivity; song selectivity; experience-dependent plasticity; *NIIs* transections; LMAN; Area X; zebra finch

Songbirds, much like humans, depend on auditory experience during early life to learn their vocal behavior. This learning occurs in two stages, called the sensory and sensorimotor phases (Fig. 1A). During the sensory phase, a young bird listens to and memorizes the song of its tutor; this memory is called the "template." The sensorimotor phase begins with the onset of singing; using auditory feedback, the juvenile compares its immature vocalizations with the tutor song template and gradually modifies the plastic song until it produces a mature "crystallized song," which is highly stereotyped and resembles the tutor song. Thus, experience of both the tutor song and the bird's own song (BOS) is required for normal song learning (Konishi, 1965; Price, 1979).

Likely candidates for circuits involved in processing BOS and tutor song experience during learning lie within the song system, a group of nuclei dedicated to song learning and production (Fig. 1B). The motor pathway, which is necessary for normal song production throughout life, includes HVC, the robust nucleus of the archistriatum (RA), and the tracheosyringeal portion of the

hypoglossal nucleus (nXII<sub>ts</sub>). The nXII<sub>ts</sub> contains the motor neurons innervating the muscles of the syrinx, the avian vocal organ. RA also projects to a group of nuclei associated with respiration, such as nucleus retroambiguus (RAm) and nucleus paraambiguus (PAm) (Wild, 1993, 1997; Reinke and Wild, 1998); these participate in vocalization by controlling the respiratory musculature involved in airflow through the syrinx. In contrast to the motor pathway, nuclei of the anterior forebrain (AF) pathway are not required for singing in adulthood, but play a critical, unknown role during song learning (Bottjer et al., 1984; Sohrabji et al., 1990; Scharff and Nottebohm, 1991; Basham et al., 1996). The AF pathway comprises Area X (X), the medial nucleus of the dorsolateral thalamus (DLM), and the lateral magnocellular nucleus of the anterior neostriatum (LMAN), and indirectly connects HVC to RA. Thus, the AF might process auditory information essential for learning and might use it to modulate motor pathway activity.

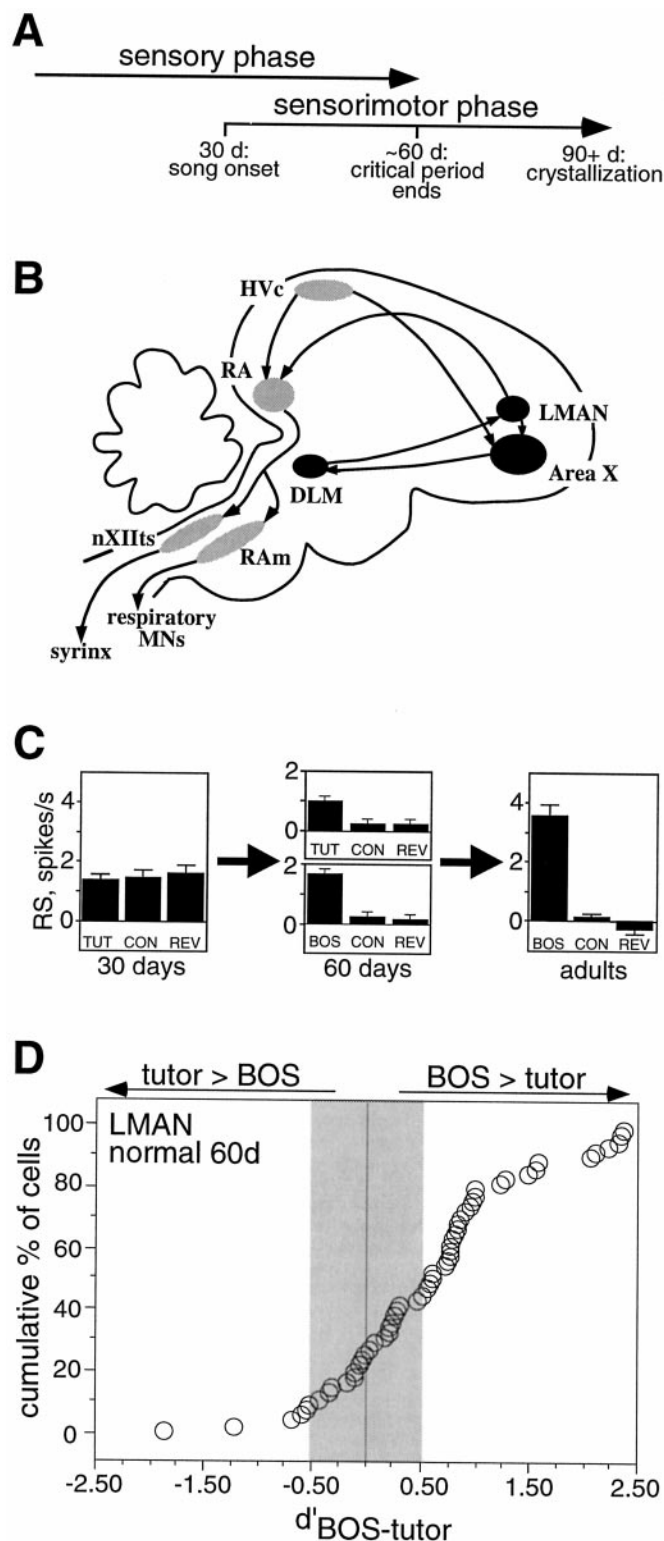
Consistent with an auditory role for the AF during learning, AF neurons in adult, anesthetized birds are auditory and respond selectively to BOS (Doupe and Konishi, 1991). Neurons selective for BOS prefer it to the songs of other zebra finches (conspecific song) and to BOS played in reverse. These song-selective neurons resemble those found in HVC (Margoliash, 1983), as well as neurons tuned to species-specific vocalizations found in bats (Suga et al., 1978; Esser et al., 1997), rhesus monkeys (Rauschecker et al., 1995), and marmosets (Wang et al., 1995). AF neurons from young juvenile birds lack selectivity, however, re-

Received Nov. 30, 1998; revised Feb. 10, 1999; accepted March 10, 1999.

This work was supported by National Institutes of Health Grants MH55987 and NS34835, the Merck Fund, and the EJLB Foundation (to A.J.D.), and a National Institute of General Medical Sciences fellowship (to M.M.S.). We thank Michael Brainard and Frédéric Theunissen for excellent technical help and Michael Brainard and Charlotte Boettger for insightful comments on this manuscript.

Correspondence should be addressed to Allison Doupe, Department of Physiology, Box 0444, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143-0444.

Copyright © 1999 Society for Neuroscience 0270-6474/99/194559-26\$05.00/0



**Figure 1.** *A*, Zebra finches learn to sing in two overlapping phases. The sensory phase ends at ~60 d; the sensorimotor phase begins at ~30 d and continues until ≥90 d. *B*, Anatomy of the song system. Motor pathway nuclei are gray, and the AF nuclei are black. *C*, AF neurons develop selectivity for song during development. At 30 d, LMAN neurons have equal RS to tutor song (TUT), conspecific song (CON), and reverse tutor song (REV). At 60 d, these neurons respond significantly more to TUT than to CON or to REV. In addition, BOS elicits a stronger RS than CON or reverse BOS (REV). In adults, LMAN neurons are extremely selective for BOS. *D*, At 60 d, there is a range of BOS versus tutor song preferences

responding equally well to all song stimuli at 30 d of age (Fig. 1*C*). Song selectivity develops rapidly, because it is found in 60-d-old juveniles (Solis and Doupe, 1997).

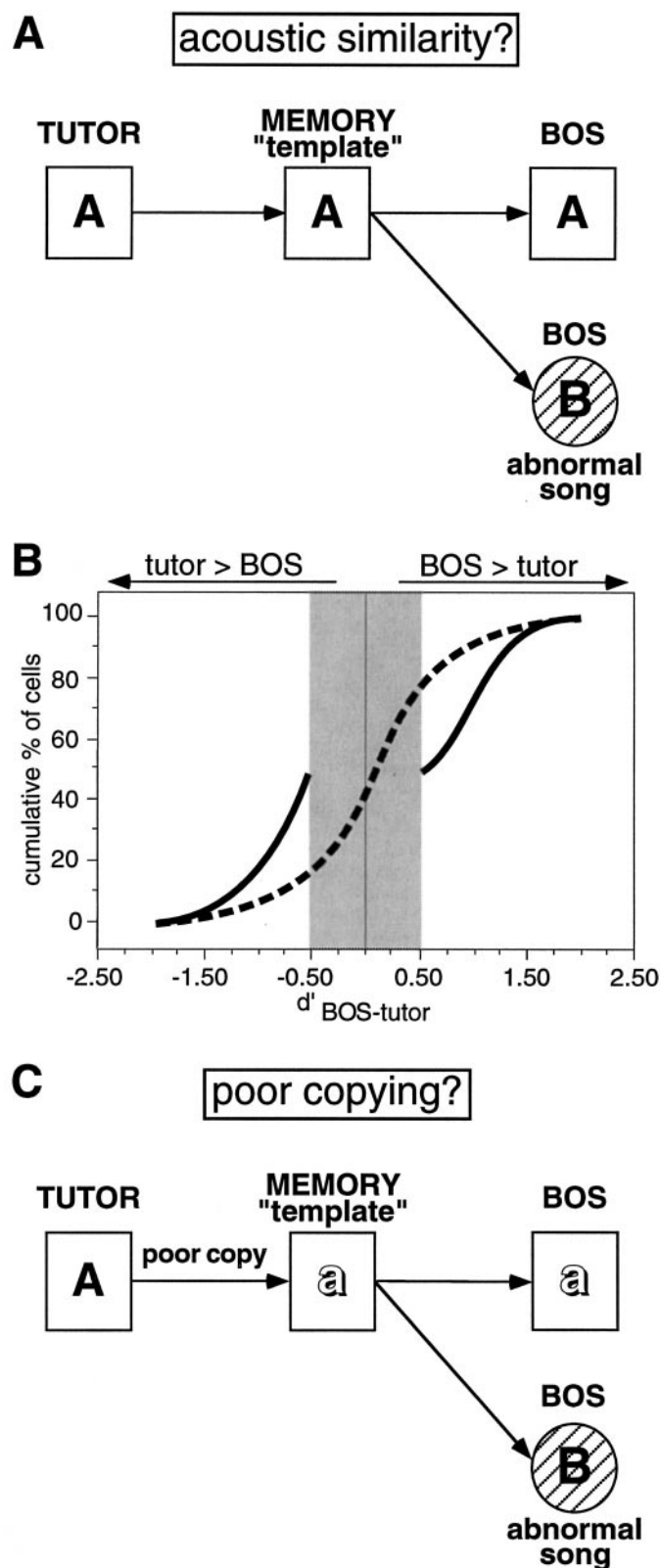
Determining the experience responsible for AF neuron selectivity could elucidate AF function during song learning. For example, neurons tuned by BOS experience could provide feedback about the current state of BOS, whereas those tuned by tutor song experience could store tutor song information. When neural responses to BOS and tutor song are compared at 60 d, a range of preferences for one song over another is evident (Fig. 1*D*; adapted from Solis and Doupe, 1997). Many neurons prefer BOS over tutor song, suggesting a role for BOS experience in shaping selectivity. A few neurons prefer tutor song over BOS, suggesting that they were tuned by tutor song experience. Finally, many neurons respond equally well to both BOS and tutor song. These neurons are clearly selective, because they do not simply respond to any song stimulus. Such neurons could have been shaped by both BOS and tutor song experience. Alternatively, these neurons might indicate acoustic similarities between the two songs; by 60 d some juveniles' plastic songs clearly resemble their tutor song.

If neurons with similar responses to BOS and tutor song result from acoustic similarities between the two songs, then it is unclear which song experience is responsible for neural selectivity. Inducing a juvenile bird to produce an abnormal song could resolve this issue, because it would reduce similarity between BOS and tutor song (Fig. 2*A*). If neurons with equivalent responses to BOS and tutor song result from the similarities between these two songs, then such neurons should not exist in birds with songs very different from their tutor song (Fig. 2*B*, solid line). Alternatively, if such neurons reflect the contributions of both song experiences, then neurons with similar responses to the abnormal song and tutor song should persist (Fig. 2*B*, dashed line).

Birds producing abnormal songs could also clarify the experience responsible for neurons that prefer BOS over tutor song in normal 60 d birds. The simplest interpretation is that these neurons are shaped by BOS experience. If, however, a bird has poorly copied the tutor song during the sensory phase, then these neurons might instead represent the template. This possibility is schematized in Figure 2*C*; if a bird stores a poor copy of the tutor (*A*) as its template (*a*) and models its own song accurately after the template (*a*), then BOS itself is a better representation of the template than the tutor song. This issue could be resolved with birds induced to produce very abnormal songs; if neurons preferring BOS over tutor song persist in such birds, then it is likely that they result from experience of the song unique to that bird.

In this study, we minimized the similarity between the songs of juvenile birds and their tutors by transecting the tracheosyringeal portion of the hypoglossal nerve [NXII(ts)], which innervates the syringeal muscles, before song onset. Extracellular recordings of single LMAN and X neurons in these birds at 60 d showed that, although the BOS and tutor song were now acoustically very different, many neurons still responded equally well to both stimuli. This result is similar to that found in normal 60 d birds and

among LMAN neurons. The cumulative distribution of preferences is shown, as quantified with a  $d'_{\text{BOS-tutor}}$  value for each neuron (see Materials and Methods). Neurons with values ≥0.5 are considered to prefer BOS over tutor song, and neurons with values of −0.5 or less are considered to prefer tutor over BOS. Gray shading highlights those values for which there was no strong preference for one song over the other ( $-0.5 < d'_{\text{BOS-tutor}} < 0.5$ ).



**Figure 2.** Consequences of decreasing similarity between BOS and tutor song. *A*, When a juvenile stores a good copy of the tutor song (*A*) as its template (*A*) and accurately models its own song after the template, the resulting BOS (*A*) will highly resemble the tutor song. Thus, if a neuron is tuned by BOS experience only, it could also respond well to tutor song when the two songs are similar enough. This ambiguity could be resolved by making the BOS very different (*B*) from the tutor song. *B*, Decreasing

suggests a role for both song experiences in shaping AF selectivity.

## MATERIALS AND METHODS

**Animals.** Experiments used male juvenile zebra finches (*Taeniopygia guttata*). The care and treatment of experimental animals was reviewed and approved by an university animal care and use committee at University of California, San Francisco (UCSF). Birds were raised in individual cages, with their parents and siblings from the same clutch. Opaque dividers between cages visually isolated birds from other conspecifics in the colony. Because juvenile birds shared a cage with a single adult male tutor and were visually isolated from other conspecifics within earshot, their learning should have been restricted to the tutor in their cage (Immelmann, 1969; Eales, 1987; Eales, 1989; Williams, 1990).

**Surgery.** When birds were 26–33 d old (mean  $\pm$  SD,  $28 \pm 2$  d), the tracheosyringeal portion of the hypoglossal nerve (NXIIIts) was transected bilaterally under isoflurane anesthesia [0.5–1.5% (v/v); Abbott Laboratories, North Chicago, IL]. The nerves were exposed by an incision along the skin of the neck, where lidocaine had been injected subcutaneously (2% solution; Elkins-Sinn, Cherry Hill, NJ). The NXIIIts nerve was dissected away from the trachea at the proximal end of the incision and cut; dissection then continued along the length of the neck, and the nerve was pulled to remove the distal end. This removed  $\sim 1$  cm of nerve. After bilateral transections, the skin was closed with skin adhesive (Krazy Glue; Borden, Columbus, OH). The ts cut birds were returned to their home cages until they were 60-d-old.

Two days before the experiment, we prepared birds for recording by affixing a head post to the skull and marking the location of the song nuclei on the skull (for details, see Solis and Doupe, 1997). On the day of the experiment, the bird was anesthetized with a 20% solution of urethane (5 ml/kg, i.m.; Sigma, St. Louis, MO; delivered in three injections at 30 min intervals), placed in the stereotaxic apparatus, and immobilized via its head post. Body temperature was regulated with a temperature controller (FHC, Brunswick, ME). A craniotomy was performed above LMAN and X, the dura was opened, and the electrode was lowered into the brain with a microdrive (Fine Science Tools, Foster City, CA).

**Stimuli.** One to 2 d before the experiment, the songs of the ts cut bird and its tutor were recorded. Each bird was placed in a sound-attenuated chamber (Acoustic Systems, Austin, TX) connected to an automatically triggered audio system. Approximately 90 min of bird sounds were recorded and then scanned for song. A typical plastic song rendition was usually chosen after listening to at least 25 songs and looking at several song spectrograms; a typical song was considered to be the song most frequently sung. A typical tutor song was chosen after listening to 10 songs. Songs were digitized at 32 kHz and stored on a SPARC (Sun Microsystems, Palo Alto, CA) IPX computer at similar peak intensity levels (range, 64–73 dB; software by Michael Lewicki and Larry Proctor, California Institute of Technology, Pasadena, CA). In 15 experiments, three different plastic song renditions from a bird were stored for presentation during the experiment. The durations of tsBOS and tutor songs ranged from 602 to 2461 msec.

During electrophysiological recording, acoustic stimuli were presented by a speaker 25 cm away from the bird, inside a double-walled anechoic sound-attenuated chamber (Acoustic Systems, Austin, TX). The frequency response measured at the bird's location inside the chamber was

the similarity between BOS and tutor song has two predicted outcomes on the distribution of  $d'$  BOS-tutor values. If BOS experience shapes some neurons, and tutor song experience shapes others, then the distribution should be split in two, with some neurons preferring tsBOS over tutor song and others preferring tutor over tsBOS but none responding equally well to both (*solid line*). Alternatively, if both tsBOS and tutor song influence the neural properties of single neurons, then neurons with equivalent responses should persist (*dotted line*). *C*, If a poor copy of the tutor song (*A*) is stored as the template (*a*) and then a good copy of the template is produced, then the resulting BOS (*a*) is a better model of the template than the tutor song itself. In this case, neurons preferring BOS would nonetheless reflect tutor song experience. Inducing an abnormal BOS by disrupting sensorimotor learning (*B*) should decrease the similarity between BOS and a song resulting from poor memorization of the tutor song.



flat ( $\pm 5.0$  dB) between 500 Hz and 8 kHz. The stimuli included songs of the ts cut juvenile (tsBOS), its tutor song, reverse manipulations of tsBOS and tutor song, the songs of other zebra finches (conspecific), the acoustically similar songs of other species of estrildid finches (heterospecific), broad-band noise bursts, and tone bursts. Stimuli were presented in a random, interleaved manner. An effort was made to present each neuron with 15–20 trials of each stimulus type: tsBOS, reverse tsBOS, reverse order tsBOS, tutor, reverse tutor, reverse order tutor, at least two adult conspecific songs, at least two heterospecific songs, at least two juvenile conspecific songs, at least two ts cut juvenile conspecific songs, broad-band noise bursts, and tone bursts; however, some neurons were lost before characterization was completed.

**Electrophysiology.** Extracellular neuronal signals were amplified and filtered between 300 Hz and 10 kHz (A-M Systems, Everett, WA). To locate auditory neurons, search stimuli included tsBOS, tutor song, adult conspecific song, heterospecific song, broad-band noise bursts, and tone bursts. Most neurons were isolated with a window discriminator (UCSF Physiology Shop). Twelve units were isolated offline using spike-sorting software (Lewicki, 1994). To do this, waveforms were recorded during stimulus presentation during the experiment. Later, spike models were constructed from waveforms recorded at an intermediate time during stimulus presentation. These spike models were then used to classify spikes within the rest of the waveforms. Both spike model construction and template-matching algorithms were based on Bayesian probability theory. Neuronal responses were collected and analyzed by a SPARC IPX computer using software developed by Mike Lewicki and Larry Proctor (California Institute of Technology) and Frédéric Theunissen (UCSF). Electrolytic lesions were made at selected locations for reconstructing recording sites.

**Anatomy.** At the end of an experiment, the bird was deeply anesthetized with Metofane (Pitman-Moore, Mundelein, IL) and transcardially perfused with 0.9% saline, followed by 3.7% formalin in 0.025 M phosphate buffer. Brains were post-fixed and cut in 40  $\mu$ m sections with a freezing microtome. Sections were stained with cresyl violet, and electrode tracks and lesions were identified. Only neurons histologically confirmed to be in LMAN or X were used; their specific location within each nucleus was also documented.

RA volumes were measured for each ts cut bird and for normal 60 d birds, recorded in a previous study. Measurements were made blind to the experimental condition. The Nissl-defined boundaries of RA were traced at 80  $\mu$ m intervals, and the resulting area was calculated using an image analysis program (NIH Image). The total area was multiplied by section thickness and then by the total number of sections to give a final volume. Because of individual differences in post-fixation time, each RA volume was normalized by the volume of the nucleus prefrontalis (PT), which is unrelated to the song system. Final RA/PT ratios were compared between ts cut and normal birds. When measurements from both hemispheres were available, the mean RA volume and mean PT volume were used. For nine ts cut birds, PT volume was not available. Thus, RA volumes alone were also compared within all ts cut birds for which post-fixation times were equivalent.

The syrinx of each ts cut bird was also dissected after perfusion. Each syrinx was cut 1 mm distal and 4 mm proximal of the bifurcation of the bronchi and then weighed to assess relative muscle mass, a marker of denervation success.

**Data analysis.** We quantified responses to an acoustic stimulus during the period of stimulus presentation, offset by an estimate of the latency. The latency of each neuron was measured by examining its responses to a broad-band or tone burst stimulus with a peristimulus time histogram (PSTH) divided into 5 or 10 msec bins. The latency was defined as the onset of the first of two consecutive bins during the stimulus that had at least twice as many spikes as the mean number of spikes per bin during the background. LMAN neurons often did not respond to broad-band noise or tone bursts. For these cases, the latency of another neuron from the same bird was used; if there was none, then the neuron was assigned a latency characteristic of neurons from normal 60 d birds (65 msec; from Solis and Doupe, 1997).

To be considered auditory and included for analysis, a neuron had to have an average firing rate during one of the stimuli that was significantly different from the background rate (two-tailed paired *t* test,  $p < 0.05$ ). The firing rate during a stimulus was obtained by normalizing the number of spikes elicited during the stimulus by the duration of the stimulus. The background rate was calculated by averaging the firing rate of the neuron from two different periods: 2 sec preceding stimulus onset and 2–3 sec beginning 1 sec after the end of the stimulus. The response

strength (RS) of a neuron to a stimulus was the difference between the firing rate during the stimulus (offset by the latency) and the background rate. The RS was measured for each stimulus trial and then averaged across trials to get the neuron's RS to that stimulus, expressed in spikes per second. Data for different stimuli but of the same stimulus type were also averaged in this way to get an RS for a stimulus type; e.g., to obtain the RS for adult conspecific song, the RS values for each trial of two different adult conspecific song stimuli were averaged together.

The selectivity of an individual neuron for one stimulus (A) over another (B) was quantified using the  $d'_{A-B}$  measure (Green and Swets, 1966), where:

$$d'_{A-B} = \frac{2(\overline{RS_A} - \overline{RS_B})}{\sqrt{\sigma_A^2 + \sigma_B^2}}.$$

In this equation,  $\overline{RS_A}$  and  $\overline{RS_B}$  are the mean RS to stimulus A and B, respectively, and  $\sigma^2$  is the variance of each RS. If  $d'_{A-B}$  is positive, then stimulus A elicited a greater response; if it is negative, then stimulus B elicited a greater response. Values of  $d'_{A-B}$  close to 0 indicate no difference in the RS elicited by the two stimuli. A particular  $d'$  value was calculated only for neurons that had a significant response to at least one of the two stimuli compared. A neuron was considered selective for stimulus A over stimulus B if it had a  $d'_{A-B}$  value  $\geq 0.5$ . This criterion was based on the observation that neurons with a  $d'_{A-B}$  value  $\geq 0.5$  usually had an RS to stimulus A that was at least twice as great as that to stimulus B (Solis and Doupe, 1997). Also, a  $d'_{A-B}$  value of 0.5 corresponds to a significantly greater response to stimulus A than to stimulus B, based on a paired *t* test with 20 presentations of each stimulus ( $p = 0.031$ ).

To convey the magnitude of the difference between the RS elicited by two different stimuli, the selectivity index (SI) was also calculated (Volman, 1996; Doupe, 1997). The SI compared the mean RS with each stimulus in ratio form:

$$\frac{\overline{RS_A}}{(\overline{RS_A} + \overline{RS_B})}.$$

When comparing RS to two stimuli with large differences in song duration, normalizing spike counts elicited by the two stimuli by stimulus duration may bias comparisons of the RS. For example, if two stimuli, one short and one long, elicit a similar response in which the neuron initially fires strongly and then fatigues, then normalizing by song duration will give a substantially decreased RS for the long stimulus relative to the shorter stimulus; this in turn will result in a  $d'$  value that prefers short stimuli over long stimuli. Because large differences in song duration occurred in several experiments, a peak RS was also calculated to remove bias attributable to varying song durations in the comparisons of a neural response. First, a maximum firing rate during the stimulus was found using a 500 msec sliding window, which moved across a response in 1 msec increments. Second, the maximum background rate was also found using a 500 msec window. Third, the peak RS was calculated by taking the difference between the maximum firing rate during the stimulus and the maximum background rate; this peak measurement removes duration bias because it normalizes every spike count by 500 msec, regardless of the stimulus duration. Finally, peak  $d'$  values were also calculated using the peak RS obtained from the 500 msec window. A 500 msec window was chosen for two reasons. First, it was shorter than the shortest song stimulus (602 msec). Second, for a subset of neurons (five from LMAN and five from X), a series of sliding windows (10–2000 msec) were used to calculate the peak RS and resulting peak  $d'$  values. Among those windows  $< 600$  msec, the 500 msec window gave the largest peak  $d'$  values between two stimuli of similar durations. For some cells, windows  $> 500$  msec resulted in  $d'$  values higher than those for short windows (our unpublished data); this indicates that peak  $d'$  measures can underestimate the selectivity of a cell.

**Cluster analysis.** We tested whether the  $d'_{\text{tsBOS-tutor}}$  values of neurons recorded from each bird were more similar than expected by chance. To do this, the variance of the  $d'_{\text{tsBOS-tutor}}$  values obtained experimentally from each bird was compared with a simulated distribution of variances created from the data from all birds. This distribution was determined from 1000 Monte Carlo simulations; each simulation randomly selected  $n$   $d'_{\text{tsBOS-tutor}}$  values from the pool of all experimental  $d'_{\text{tsBOS-tutor}}$  values (includes all cells from all birds) and calculated their variance ( $n$  equals the number of cells recorded in each bird). The median of the resulting distribution of simulated variances was compared with each bird's ex-

perimental variance. If the experimental variance was significantly less than the median of the simulated distribution (one-sample sign test,  $p < 0.05$ ), the  $d'_{\text{tsBOS-tutor}}$  values from that bird were considered clustered. A sign test determined whether the frequency of clustering in the group of birds studied was greater than expected by chance. This procedure was completed for  $d'_{\text{tsBOS-tutor}}$  values from LMAN neurons alone, X neurons alone, and both neuron types together.

**Song analysis: similarity.** Once electrophysiology experiments were completed, we analyzed the tsBOS and tutor songs themselves using several methods. Song is composed of syllables, which are continuous acoustical signals, 10–200 msec in duration. Syllables are separated from other syllables by a sudden fall in amplitude to near zero or by brief silent intervals. Syllables are composed of smaller continuous signals called “notes.” A repeated sequence of syllables is a “motif.” A song “bout” consists of introductory notes followed by one or more motifs (for detailed song descriptions, see Price, 1979; Sossinka and Bohner, 1980).

The first song analysis was a matching task, completed by nine human observers familiar with zebra finch song but blind to the neural properties of each bird. Observers tried to match each experimental song with that of its tutor, which was present among a group of six potential tutors. The observers listened to and looked at sonograms and oscillograms of the songs before selecting the tutor song that best matched the experimental song. Thus, the percentage of observers that correctly matched the experimental song to its tutor song indicated the overall similarity between tsBOS and tutor song; this measure was called the “percent correctly matched.” After selecting a “best match” tutor song, observers scored the song pair on spectral similarity and on temporal similarity using a scale from 1 to 5. For spectral similarity, observers only considered syllable morphology and sequence. A score of 1 referred to a song pair for which no elements in the experimental song resembled anything in the best match tutor song; 2 was given to a song pair when some notes in the experimental song resembled notes present in the best match song; 3 designated a song pair in which one or more syllables of the experimental song resembled distinctive syllables of the best match song; 4 referred to a song pair for which several experimental song syllables resembled those of the best match song, and the syllable sequences were somewhat similar; and 5 was given to a song pair when the experimental song resembled the best match song in both syllable morphology and sequence, making it a good copy of the best match song.

To judge temporal similarity, observers disregarded the spectral features of song and considered only the durations of syllables and intervals and their patterns, or rhythm, within the songs. Each song pair was scored on a scale of 1 to 5. A score of 1 referred to a song pair for which a timing similarity between the experimental song and the best match song could not be detected; 2 indicated a song pair for which the relative durations of at least two syllables and the interval between them in the experimental song resembled timing in the best match song (e.g., doublets or triplets were heard in both songs); 3 was given to a song pair when combinations of doublets or triplets in the experimental song resembled the timing structures of the best match song; 4 was given to a song pair when many syllables and intervals of the experimental song had relatively similar duration and patterning as those in the best match song; and 5 indicated a song pair for which the timing of the experimental song was highly similar to that of the best match song, although differences in speed may have been apparent.

Songs of non-ts cut birds were also included among the experimental songs for analysis, and their respective tutor songs were also present among the possible tutor choices; this provided references against which ts cut song similarity scores could be compared. Normal 60 d song ( $n = 16$ ), normal adult song ( $n = 9$ ), and randomly matched song (songs for which the correct tutor was not present among the possible tutor choices;  $n = 6$ ) were also matched to a tutor song and scored for spectral and temporal similarity. Scores given to normal adult songs provided an upper bound of similarity between songs from normal adults and their tutors, whereas scores given to randomly matched songs provided a lower bound of similarity. Randomly matched songs included those from two normal adult, two normal 60 d, and two ts cut 60 d birds.

To control for slight scoring differences between observers, we normalized each observer's score for a song by the observer's mean score for all songs. Thus, if an observer scored the spectral similarity of a song pair as a 5, but the observer's mean score was a 3, then the score for this particular song pair was  $5/3 = 1.7$ . The normalized scores for birds ranged from 0.30 to 2.43. The final score for each song was the average of each observer's normalized score. This final score included scores given to incorrect experimental–tutor song matches. Scores calculated

with incorrect matches excluded were not significantly different (paired  $t$  test,  $p < 0.05$ ); this indicates that incorrectly chosen tutor songs were as dissimilar from the experimental song as the tutor song itself. The mean score for song type (i.e., ts cut, normal 60 d, adult control, and randomly matched) was calculated from the final scores for each song belonging to the song type.

The similarity between each experimental song (ts cut, normal 60 d, normal adult, and randomly matched songs) and its tutor song was also measured with a cross-correlation algorithm (Theunissen and Doupe, 1998). One song waveform was moved relative to another in 1 msec increments, and an  $r^2$  value was calculated for each time delay. The maximum was used as the “cross-correlation measure.” Unlike the spectral and temporal similarity scoring in the matching task, cross-correlations were done between an experimental song and the correct tutor song (except for randomly matched songs; these were cross-correlated to the tutor song most often chosen by observers in the matching test).

To measure overall similarity, the entire spectrogram of an experimental song was cross-correlated to the entire spectrogram of the tutor song. To measure spectral similarity, the “syllables-only” cross-correlation measure was calculated for each song pair. For this, each isolated syllable of the experimental song was compared with each isolated syllable of the tutor song. The cross-correlation measure was calculated for each comparison, and the maximum was taken as the best match for the syllable. The resulting maxima were then averaged to produce the syllables-only cross-correlation measure. To measure temporal similarity, each song waveform was rectified and low-pass filtered at 62.5 Hz. The filtered versions of experimental song and tutor song were then cross-correlated to give the “temporal envelope” cross-correlation measure.

To further compare temporal features of song, overlap values were calculated between these song pairs (program by Michael Brainard, UCSF). For this, the syllables of each song were replaced with square pulses of equal amplitude. The resulting square pulse strings preserved syllable and interval durations and their patterns found in the original songs. The square pulse string of an entire experimental song was then compared with that of the entire tutor song by calculating the percent overlap between syllables and intervals. The proportion of overlap between experimental syllables and tutor song syllables was calculated separately from the proportion of overlap between experimental intervals and tutor song intervals. The mean of the syllable and interval overlap values was the “song–song overlap” value.

In addition, a “motif–song overlap” value was calculated, which maximized the chance of overlap. The song–song overlap measure described above could miss timing similarities between motifs of two songs, if there were different intervals between multiple motifs within a song. To avoid this, the motif–song overlap value compared a string based on a single motif of the experimental song with a string based on the entire tutor song. In addition, the song–song overlap measure could miss timing similarities if there were differences in song speed; thus, the motif–song overlap calculations allowed the motif string to stretch proportionately 80–120% of its original length, in 2% increments. The percent overlap between each stretched version of the motif string and the tutor song string was calculated, and the maximum was taken as the “maximum overlap” value. Finally, overlap values are sensitive to the complexity of the motif string of the experimental bird. For example, a simple motif comprising only two syllables is likely to give a high maximum overlap value for both the tutor song and a random song. To correct for this, the maximum overlap value was normalized by how well the motif overlapped with random songs. To obtain a measure for random overlap, the maximum overlap value was determined between the motif string and 20 randomly chosen, normal adult song strings. The mean of the 20 maximum overlap values gave the “random overlap” value. This random overlap value was used to normalize the maximum overlap value obtained from the comparison of the motif and tutor song strings, such that:

$$\text{motif–song overlap} = \frac{(\text{maximum overlap} - \text{random overlap})}{(1 - \text{random overlap})}.$$

**Song analysis: stereotypy.** We measured song stereotypy of each bird in three ways: human subjective scoring, syllables-only cross-correlations, and motif–song overlap analysis. For reference, songs of normal adult and normal 60 d birds were included in the stereotypy test. For each bird, 10 song bouts were randomly selected for analysis (except in four cases: two normal 60 d birds had two songs each; one normal 60 d bird had five songs; and one ts cut 60 d bird had only three songs).

Three observers rated how consistently a particular motif was present in each song sample from a single bird on a scale from 1 to 5. They listened to each song sample and looked at their accompanying sonograms and oscillograms before deciding on the score. Both spectral and temporal pattern repeats contributed to the score. A score of 1 referred to a group of songs that were not at all stereotyped: short syllable sequences and small temporal patterns were rarely, if at all, repeated in the song samples. A score of 2 indicated that a particular syllable sequence or brief temporal pattern was repeated in half or fewer of the song samples. A 3 was given when a short syllable sequence or temporal pattern was repeated in almost all or all song samples. Alternatively, a 3 was given if an entire motif structure was repeated in only half of the song samples. The syllables outside of the repeated structures could vary in identity and ordering. A score of 4 was given when an entire motif structure was apparent in most or all of the song samples; however, some variability remained, with syllables added or dropped from the motif in different renditions. A score of 5 was given when identical motifs were found in every song sample. Each score was normalized by the observer's mean score, as described for the similarity scoring in the matching task. Normalized stereotypy scores ranged from 0.26 to 1.40.

To isolate spectral stereotypy, we used cross-correlations to measure how consistently the syllables in one song were present in the other song samples. Spectral stereotypy was calculated in the same manner as the syllables-only cross-correlation measure of similarity described above, except that the cross-correlations were done between syllables from songs of the same bird. The mean of the resulting syllables-only cross-correlation measures (usually nine coefficients) gave the spectral stereotypy measure for the bird.

Motif–song overlap analysis was used to measure temporal stereotypy. To measure how consistently temporal patterns were repeated in song samples from a bird, a motif–song overlap value was obtained for an experimental motif string and each song sample string (usually nine sample strings), as described above. These were also normalized by a random overlap value, which was obtained by calculating the maximum overlap between the motif string and nine randomly chosen songs from all the experimental groups (adult control, 60 d control, and 60 d ts cut). The normalized motif–song overlap values for each comparison between songs from the same bird were then averaged to give an overlap stereotypy measure.

## RESULTS

### Songs of ts cut birds at 60 d

Bilateral NXIIIs (ts) transections do not disturb the respiratory outputs involved in song production; thus, birds receiving ts cuts at ~30 d of age readily sang, but because they could not control their syringeal musculature, they produced extremely abnormal songs by 60 d. These birds sang a series of simple syllables consisting of harmonically related notes. These “harmonic stack” syllables had little amplitude modulation (Fig. 3*A*), and the frequencies of the stacks often fluctuated, giving the song a wavery quality. The song of this ts cut bird (tsBOS) was very different from its tutor song (Fig. 3*B*). Although the syllables of the tsBOS shown were longer than normal, the average syllable and interval durations in ts cut song were not significantly different from those of normal adult or 60 d song ( $p > 0.635$  for all comparisons, unpaired  $t$  tests). The song of a normal 60 d sibling of the bird in Figure 3*A* is shown for comparison (Fig. 3*C*). Although this normal 60 d song had immature features such as noisier syllables and a longer song duration than the tutor, it had clear similarity in syllable morphology and timing to the tutor song. Thus, the ts cut manipulation produced songs that were considerably simpler than normal plastic song, and dramatically reduced the similarity between BOS and tutor song that can occur by 60 d.

We quantified the decrease in similarity between tsBOS and tutor songs using multiple methods of song analysis. In a matching task, observers tried to match an experimental song (ts cut 60 d, normal 60 d, normal adult, or random) with that of its tutor, which was present in a group of possible tutors (see Materials and Methods). Songs from ts cut birds were correctly matched to their

tutor song significantly less frequently than were songs from normal 60 d birds (Fig. 3*D*) (unpaired  $t$  test,  $p < 0.002$ ). Because NXIIIs transections in adult birds are known to preserve the overall timing of song but to eliminate normal spectral features (Simpson and Vicario, 1990; Williams and McKibben, 1992), observers also scored separately the spectral and the temporal similarity between each experimental song and the chosen tutor song (see Materials and Methods). The resulting mean spectral similarity score and the mean temporal similarity score for ts cut songs were significantly lower than those for normal 60 d songs (Fig. 3*D*) (unpaired  $t$  test,  $p < 0.0001$  for spectral similarity;  $p < 0.001$  for temporal similarity). For reference, the mean spectral and temporal similarity scores for randomly matched songs (songs whose actual tutor song was not present among the choices; for details, see Materials and Methods) and for normal adult songs are also shown. Note that ts cut songs had significantly lower spectral similarity scores than did the randomly matched songs (unpaired  $t$  test,  $p < 0.002$ ).

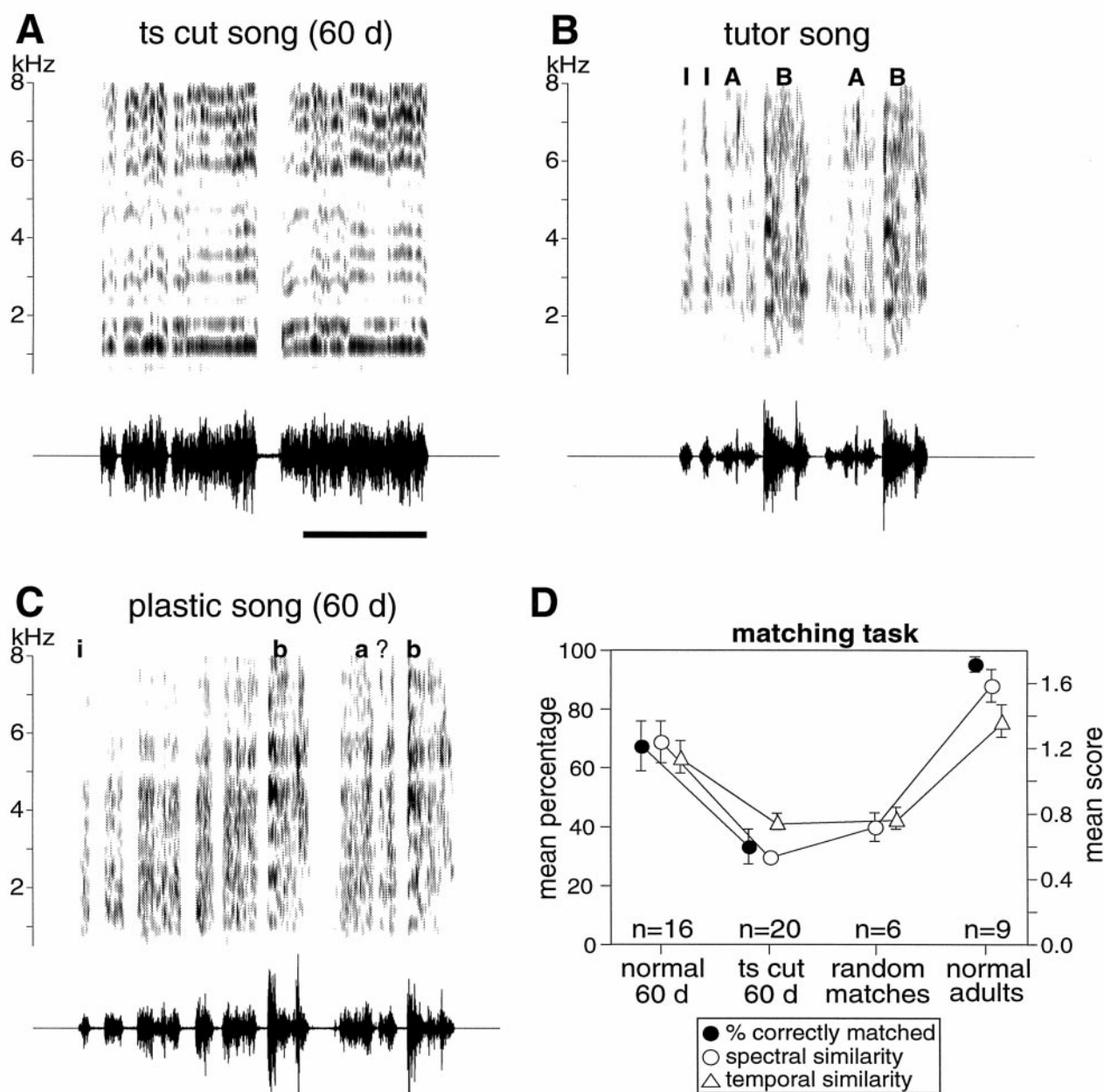
### Comparison of LMAN neural responses with tsBOS and tutor song

Extracellular recordings of 52 LMAN neurons from 16 ts cut birds revealed selectivity for tsBOS. Figure 4*A* shows a neuron that responded substantially more to tsBOS than to tutor song, adult conspecific song, and reverse tsBOS (a “mirror image” reversed song in which both entire syllables and syllable sequence are reversed). Thus, this neuron was sensitive to the spectral and temporal properties of tsBOS, despite its simple structure. Many other neurons showed a strong preference for tsBOS over tutor song. We quantified the preference for tsBOS over tutor song for each neuron with a  $d'_{\text{tsBOS-tutor}}$  value (see Materials and Methods); neurons with  $d'_{\text{tsBOS-tutor}}$  values  $\geq 0.5$  were considered to prefer tsBOS over tutor song, and neurons with  $d'_{\text{tsBOS-tutor}}$  values of  $-0.5$  or less were considered to prefer tutor over tsBOS. Classified in this way, 28% of LMAN neurons preferred tsBOS over tutor song, and only 5% of neurons preferred tutor over tsBOS (Fig. 4*B*). These strong preferences for the abnormal tsBOS over tutor song demonstrate the ability of BOS experience to shape LMAN neuron properties.

Unexpectedly, many LMAN neurons responded equally well to tsBOS and tutor song, despite the large acoustic differences between these two songs. Figure 5*A* shows an example of such a neuron, which came from a ts cut bird whose song was matched to the correct tutor song by only one of nine observers. This type of neuron represented a substantial proportion of LMAN neurons recorded (Fig. 4*B*): 67% of the neurons had  $d'_{\text{tsBOS-tutor}}$  values between  $-0.5$  and  $0.5$ , thus classifying them as neurons with equivalent responses to both tsBOS and tutor song. Overall, the mean of  $d'_{\text{tsBOS-tutor}}$  values of neurons from ts cut birds was not significantly different from that obtained from normal 60 d birds (Fig. 4*B*) (unpaired  $t$  test,  $p = 0.089$ ; normal 60 d data from Solis and Doupe, 1997). On average, tsBOS elicited a greater response than tutor song, as was true for LMAN neurons from normal 60 d birds (Fig. 4*B*, inset; paired  $t$  test,  $p < 0.004$  for neurons from ts cut birds;  $n = 46$ ).

Neurons with equivalent responses to acoustically dissimilar tsBOS and tutor songs might indicate that both song experiences shape the selectivity of single neurons. There are alternative explanations for such neurons, however. First, these neurons might not have exhibited a stronger preference for tsBOS because they were tested with a version of tsBOS that was not optimal for eliciting responses; the variability of plastic song at 60 d makes





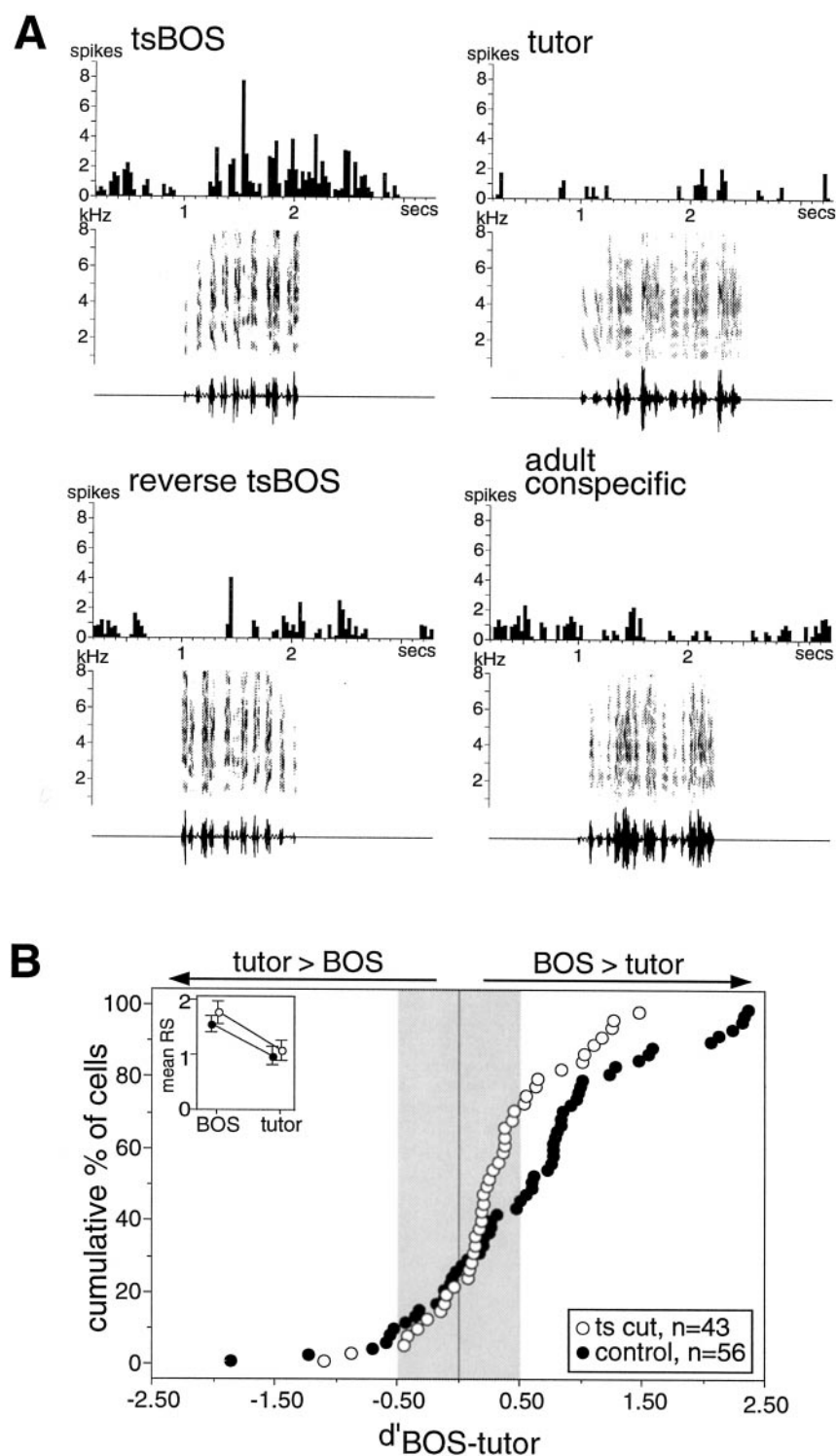
**Figure 3.** NXII's nerve transections minimized the similarity between BOS and tutor song at 60 d. *A*, Sonogram and oscillogram of the song of a ts cut bird at 60 d, which underwent nerve transections at 29 d. Sonograms plot frequency versus time, and the energy of each frequency band is indicated by its darkness; oscillograms plot the amplitude of the song waveform versus time. Calibration: *A*–*C*, 500 msec. *B*, Tutor song of the ts cut bird in *A*. Introductory notes (*I*) and syllables (*A*, *B*) are labeled. *C*, Song of a 60 d juvenile whose tutor song is also shown in *B*. Syllables that resemble those in the tutor song are labeled (i.e., syllable *b* in the juvenile song is similar to syllable *B* in the tutor song). *D*, Measures of similarity to tutor song from the matching task shown for different bird groups. Black circles show the mean percentage of observers that matched a song to the correct tutor song (left ordinate). This mean averages the frequency of matching across all songs in each song type. There is no percentage for random matches, because their correct tutor was never present among the tutor song choices. The mean spectral (white circles) and temporal (white triangles) similarity scores are plotted along the right ordinate. Error bars indicate SEM.

this possible. Second, it is possible that neurons with similar responses to tsBOS and tutor song are simply immature: younger neurons from 30 d birds respond equally well to all song stimuli (Doupe, 1997). Third, the equivalent responses to tsBOS and tutor song could be attributable to residual similarities between the two songs. Although song analysis revealed that, on average, tsBOS songs share little similarity with tutor song, it is important to compare each bird's neural properties with the similarity between its tsBOS and its tutor song. The first two alternative

explanations are discussed immediately below; the third possibility will be examined in the last section of Results, using detailed song analysis.

#### Plastic song renditions elicited equivalent neural responses

Because of plastic song variability normally present at 60 d, it seemed possible that neurons without a strong tsBOS preference had been presented with a version of plastic song to which



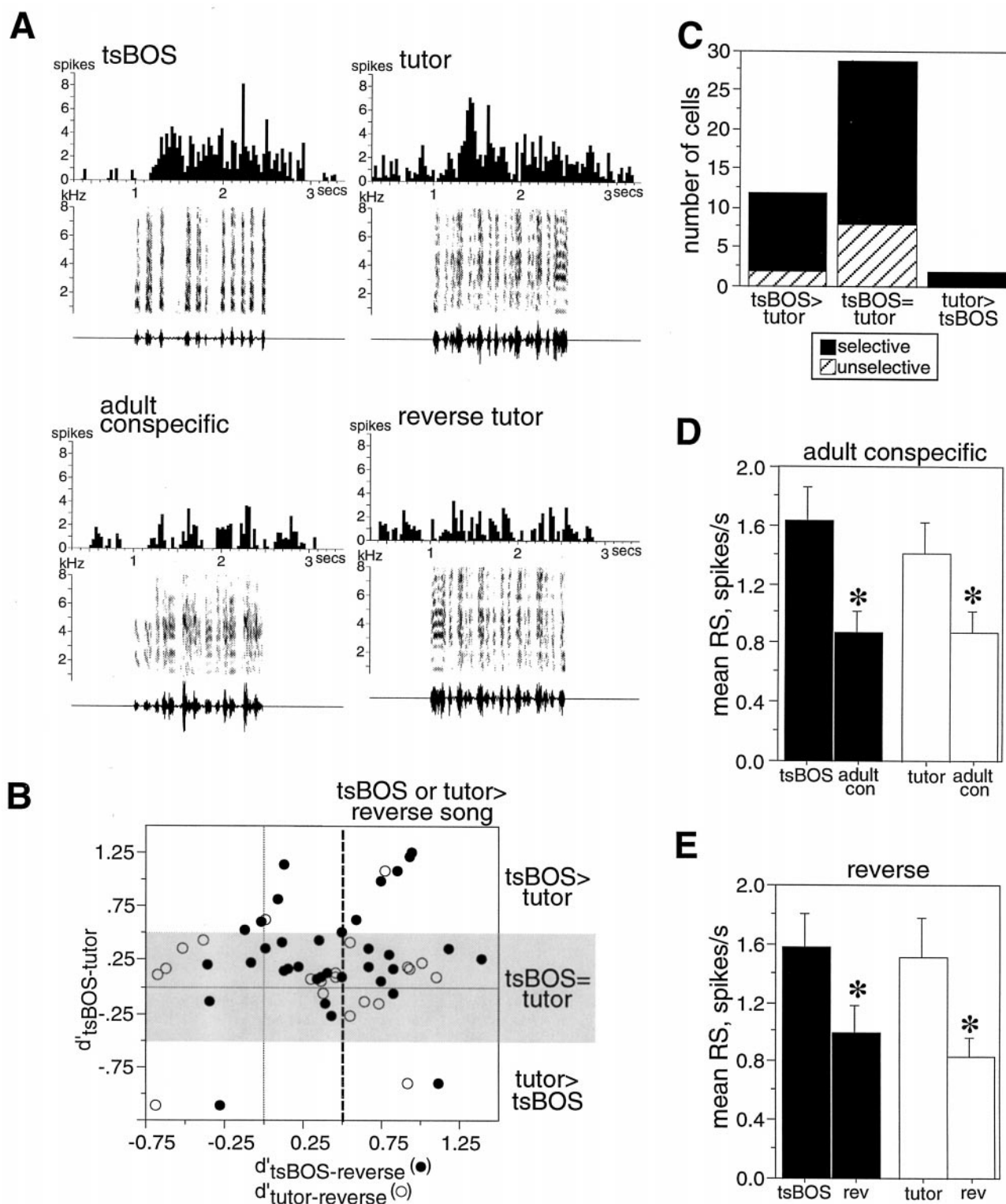
**Figure 4.** LMAN selectivity for tsBOS at 60 d. *A*, PSTHs show the greater response of a single LMAN neuron to tsBOS than to tutor song, reverse tsBOS, and adult conspecific song; 20 trials of each song were presented. For this neuron,  $d'_{\text{tsBOS-tutor}} = 1.50$ ;  $d'_{\text{tsBOS-rev}} = 1.11$ ; and  $d'_{\text{tsBOS-adult con}} = 1.41$ . *B*, The cumulative distribution of tsBOS versus tutor song preferences for all LMAN neurons recorded, as quantified with  $d'_{\text{tsBOS-tutor}}$  values, is shown with white circles. For comparison, the distribution of  $d'_{\text{BOS-tutor}}$  values from normal 60 d birds is shown with black circles. Gray shading highlights those cells considered to respond equally well to both songs. *Inset*, Mean RS of all LMAN neurons recorded to BOS and tutor song for both ts cut (white circles) and normal (black circles) 60 d birds. Error bars indicate SEM.

neurons were less responsive. To assess this, LMAN neurons were presented with three different renditions of tsBOS in eight experiments. Many neurons responded equally well to all three renditions, whereas others responded more to the tsBOS version most frequently produced by the bird. This version was always used as the primary tsBOS, to which all other songs were compared when measuring selectivity. Overall, there was no significant difference in the responses elicited by the three versions of tsBOS (ANOVA,  $p = 0.954$ ;  $n = 21$ ). Thus, it is unlikely that selectivity measurements were biased by inappropriate tsBOS presentation.

#### LMAN neurons with equivalent responses to tsBOS and tutor song were not simply immature

Because AF neuron selectivity increases between 30 d and adulthood (Doupe, 1997), selectivity can be used to assay neuronal maturity. Two types of selectivity were analyzed to determine whether neurons were immature. First, neural responses to tsBOS and tutor song were compared with those to adult conspecific songs. Second, neural responses to tsBOS and tutor song were compared with those to reversed versions of these songs; for such





**Figure 5.** Equivalent responses to tsBOS and tutor song. *A*, PSTHs show the responses of a single LMAN neuron to 13 presentations of each song. Although this neuron responded equally well to tsBOS and tutor song ( $d'_{\text{tsBOS-tutor}} = 0.14$ ), it did not respond well to adult conspecific song ( $d'_{\text{tsBOS-adult con}} = 1.13$ ;  $d'_{\text{tutor-adult con}} = 1.08$ ) or to reverse tutor song ( $d'_{\text{tutor-rev}} = 1.08$ ). *B*, The tsBOS versus tutor song preference of each LMAN neuron is compared with its selectivity by plotting  $d'_{\text{tsBOS-tutor}}$  values against  $d'_{\text{tsBOS-rev}}$  (black circles) and  $d'_{\text{tutor-rev}}$  (open circles) values. Gray shading indicates those neurons that responded equally well to tsBOS and tutor song. The dashed vertical line marks the criterion for selectivity ( $d' = 0.5$ ). *C*, This histogram shows the number of LMAN neurons classified as selective (black bars) and unselective (hatched bars) in the three different tsBOS versus tutor song preference categories. *D*, For those neurons responding equally well to both tsBOS and tutor song, histograms show paired comparisons between the mean RS to tsBOS (black bars) or tutor song (white bars) and the mean RS to adult conspecific song. *E*, For those neurons responding equally well to tsBOS and tutor song, histograms show paired comparisons between the mean RS to tsBOS (black bars) or tutor song (white bars) and the mean RS to their corresponding reverse songs. In *D* and *E*, error bars indicate SEM, and asterisks denote significant differences between each pair of stimuli.

reversed stimuli, both the entire syllables and the sequence of syllables within the song were reversed. Immature neurons would respond equally well to all of these stimuli (Fig. 1C). When we analyzed the selectivity of individual neurons with similar responses to tsBOS and tutor song, however, it was clear that these neurons were not simply immature. For example, although the neuron in Figure 5A responded strongly to both tsBOS and tutor song, it did not respond well to adult conspecific or reverse tutor song. Figure 5B further illustrates this selectivity by plotting the tsBOS versus tutor song preference of each neuron (indicated by its  $d'_{\text{tsBOS-tutor}}$  value) against a measure of selectivity ( $d'_{\text{tsBOS-rev}}$  and  $d'_{\text{tutor-rev}}$ ). Many neurons responding equally well to tsBOS and tutor song had  $d'$  values exceeding 0.5 for these measures of selectivity, indicating that they responded substantially more to tsBOS and tutor song than to reverse songs (Fig. 5B, points that lie within the *gray zone* and to the *right* of the *dashed vertical line*). Similarly, neurons with equivalent responses to tsBOS and tutor songs still discriminated between these songs and adult conspecific song (data not shown). Figure 5C shows the result of classifying neurons as selective or unselective. We considered a neuron to be selective if it had a  $d'$  value  $\geq 0.5$  for any one of four selectivity categories: tsBOS–adult conspecific, tutor–adult conspecific, tsBOS–reverse, and tutor–reverse. Classified in this way, 66% of neurons responding equally well to tsBOS and tutor song were selective. In comparison, 68% of this neuron type were classified as selective in normal 60 d birds (Solis and Doupe, 1997). Only 8 of 52 LMAN cells in the ts cut birds resembled 30 d neurons, with similar responses to every song stimulus, and seven of these came from the same animal.

Another measure of maturity is to consider the selectivity of a population of neurons by averaging their responses to different song stimuli. It is possible for individual neurons that do not themselves meet the  $d'$  criterion for selectivity but whose responses are slightly biased toward selectivity to contribute to the selectivity of an entire population of cells. As a population, LMAN neurons with similar responses to tsBOS and tutor song had greater RS on average to tsBOS and tutor song than to adult conspecific (Fig. 5D) and reverse songs (Fig. 5E) (paired  $t$  test, for tsBOS–adult conspecific,  $p < 0.0001$ ;  $n = 27$ ; for tutor–adult conspecific,  $p < 0.004$ ;  $n = 27$ ; for tsBOS–reverse,  $p < 0.0001$ ;  $n = 26$ ; for tutor–reverse,  $p < 0.011$ ;  $n = 21$ ). Thus, using both individual neuron and population measures, neurons with equivalent responses to tsBOS and tutor song exhibited selectivity, unlike immature neurons.

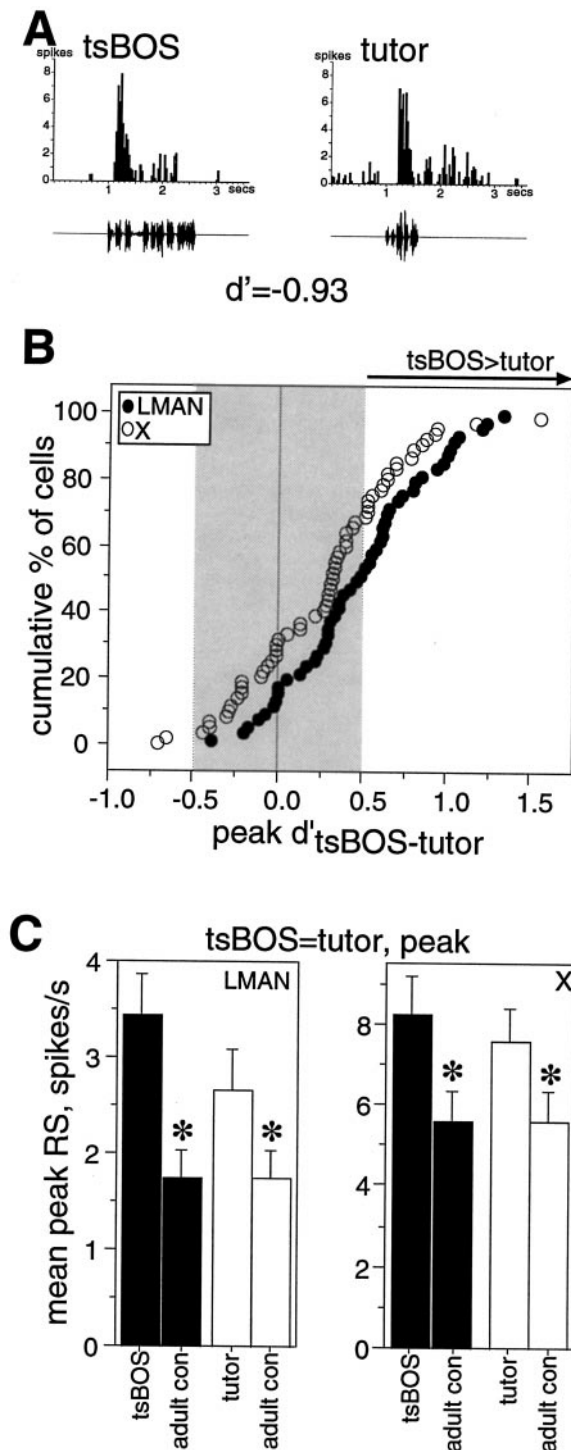
#### Alternative methods of measuring neural selectivity

In the previous analyses, comparisons of neural responses to different stimuli can be affected by stimulus duration. A neuron's RS to a stimulus was calculated by normalizing the number of spikes fired during the stimulus by the stimulus duration. If neural responses fatigue during presentation of a long stimulus, then this method will result in an RS that is less than the neuron's initial firing rate to the stimulus. This phenomenon can complicate comparisons between responses to two songs when the song durations differ substantially. For example, if two songs with large duration differences elicit the same number of spikes from a cell, then the RS to the longer song will be much less than that to the shorter song;  $d'$  measures, which compare RS to two stimuli, would tend to favor the shorter of the two stimuli. In this study, 7 of 19 experiments had substantial differences between tsBOS and tutor song duration, in which one song was at least twice as long as the other song. An example of the effect of normalizing by

song duration is shown in Figure 6A; the  $d'_{\text{tsBOS-tutor}}$  value obtained indicates a strong preference for the shorter tutor song, yet the PSTHs show qualitatively similar responses of an LMAN neuron to tsBOS and tutor song. When the  $d'_{\text{tsBOS-tutor}}$  values of individual neurons were compared with the relative difference in duration between tsBOS and tutor song, as expressed by the ratio  $(\text{duration}_{\text{tsBOS}} - \text{duration}_{\text{tutor}})/(\text{duration}_{\text{tsBOS}} + \text{duration}_{\text{tutor}})$ , a strong correlation resulted ( $r^2 = 0.584$ ;  $p < 0.0001$ );  $d'_{\text{tsBOS-tutor}}$  values reflected a preference for the shorter of the two songs.

To investigate the impact of this duration effect on the results so far described, data from experiments in which tsBOS and tutor song had similar durations (difference in duration was less than twice the shorter song) were analyzed separately (30 LMAN neurons from 10 experiments). Within this data subset, the properties described for the whole population persisted; some LMAN neurons preferred tsBOS over tutor song, whereas others responded equally well to these two songs. Among LMAN neurons responding similarly to tsBOS and tutor song, 93% (13 of 14) were classified as selective, and on average this type of neuron responded more to tsBOS and tutor song than to adult conspecific and reverse songs (data not shown) (paired  $t$  test, for tsBOS–adult conspecific song,  $p < 0.001$ ;  $n = 14$ ; for tutor–adult conspecific song,  $p < 0.008$ ;  $n = 14$ ; for tsBOS–reverse,  $p < 0.009$ ;  $n = 13$ ; for tutor–reverse,  $p < 0.031$ ;  $n = 11$ ). Thus, the neuronal properties present for the whole data set also described the subset of data collected from experiments with similar tsBOS and tutor song durations.

Another method of removing stimulus duration bias from selectivity measures is to obtain a peak firing rate for each stimulus. Peak firing rate assesses a neuron's maximum response during a stimulus, regardless of where it occurs in time. For each LMAN neuron, the maximum firing rate occurring within a sliding 500 msec window was used to calculate a peak RS to each stimulus (see Materials and Methods); thus, every response was normalized by 500 msec, regardless of stimulus duration. Peak  $d'$  values were then calculated using the peak RS to different stimuli. The resulting peak  $d'_{\text{tsBOS-tutor}}$  values indicated that there were still neurons that responded equally well to tsBOS and tutor song (53%), and neurons that preferred tsBOS over tutor song (47%) (Fig. 6B). Of the neurons responding equally well to tsBOS and tutor song, 63% were selective, as determined from their peak  $d'$  values in the four selectivity categories. In addition, the population of neurons with similar responses to tsBOS and tutor song were also selective when their responses were measured using peak RS; neurons responded on average significantly more to tsBOS and tutor song than to adult conspecific (Fig. 6C) and reverse (data not shown) songs (paired  $t$  tests: for tsBOS–adult conspecific,  $p < 0.0001$ ;  $n = 19$ ; for tutor–adult conspecific,  $p < 0.006$ ;  $n = 19$ ; for tsBOS–reverse,  $p < 0.0004$ ;  $n = 18$ ; for tutor–reverse,  $p < 0.017$ ;  $n = 16$ ). Thus, using peak RS and peak  $d'$  values, neurons that responded similarly to tsBOS and tutor song were still selective. Although peak  $d'_{\text{tsBOS-tutor}}$  values reclassified 39% of LMAN neurons in terms of their tsBOS and tutor song preferences, the overall distribution was only slightly shifted toward tsBOS preference relative to the original  $d'_{\text{tsBOS-tutor}}$  values (mean difference in  $d'_{\text{tsBOS-tutor}} = 0.22$ ; paired  $t$  test,  $p < 0.002$ ;  $n = 43$ ). Because there is no duration difference between forward and reverse versions of the same song, the maintenance of significant response differences between forward and reverse versions of song with peak RS also indicates that the 500 msec time window chosen was not too small to detect differences between responses to different stimuli.



**Figure 6.** Stimulus duration can influence the quantification of tsBOS versus tutor song preference. *A*, PSTHs show the responses of a single LMAN neuron to 12 presentations of tsBOS and a short tutor song (602 msec). Although the responses appear equivalent, the  $d'_{\text{tsBOS-tutor}}$  value indicates a preference for the shorter tutor song. *B*, The distributions of peak  $d'_{\text{tsBOS-tutor}}$  values are shown for all LMAN (black) and X (open) cells, regardless of stimulus duration. *C*, For those cells responding equally well to tsBOS and tutor song (according to their peak  $d'_{\text{tsBOS-tutor}}$  values), histograms show the mean peak RS to different stimuli for LMAN (left panel) and X (right panel) cells. Paired comparisons show that tsBOS (black bars) and tutor song (white bars) elicited greater average responses than did adult conspecific song; asterisks denote significant differences, and error bars indicate SEM.

Thus, LMAN properties in ts cut birds were the same when (1) the measurement of RS originally used in this and other studies was applied to the whole data set, (2) the original RS was used for a data subset comprising neurons collected from experiments without large duration differences between tsBOS and tutor song, and (3) peak RS was used to measure responses of the whole data set. For all three analyses, neurons that preferred tsBOS over tutor song and neurons that responded equally well to tsBOS and tutor song were apparent. The latter neurons were also selective. The original measurement of RS will be used to describe further LMAN properties in this study, because it has been used in previous studies of AF neurons.

### Selectivity of the entire population of LMAN neurons

#### Song and order selectivity

We also examined in detail the song and order selectivity of the entire population of LMAN neurons, regardless of their tsBOS versus tutor song preferences. By definition, song-selective neurons respond more to tsBOS or tutor song than to other song stimuli, such as adult conspecific and heterospecific songs. For the entire population of LMAN cells recorded, song selectivity was apparent for both tsBOS and tutor song. On average, both tsBOS and tutor song produced significantly stronger responses than adult conspecific (Fig. 7*A*) and heterospecific songs (Fig. 7*B*) (paired *t* tests:  $p < 0.0001$  for tsBOS–adult conspecific;  $n = 45$ ; and tsBOS–heterospecific;  $n = 47$ ;  $p < 0.004$  for tutor–adult conspecific;  $n = 43$ ;  $p < 0.010$  for tutor–heterospecific;  $n = 47$ ). The song selectivity of individual LMAN neurons is illustrated with scatterplots comparing each neuron's RS to tsBOS (Fig. 7*D*) or tutor song (Fig. 7*E*) with its RS to adult conspecific song. In both plots, the majority of cells lie below the diagonal line, indicating their stronger responses to tsBOS or tutor song than to adult conspecific song. The percentages of selective LMAN cells in each song selectivity category are listed in Table 1.

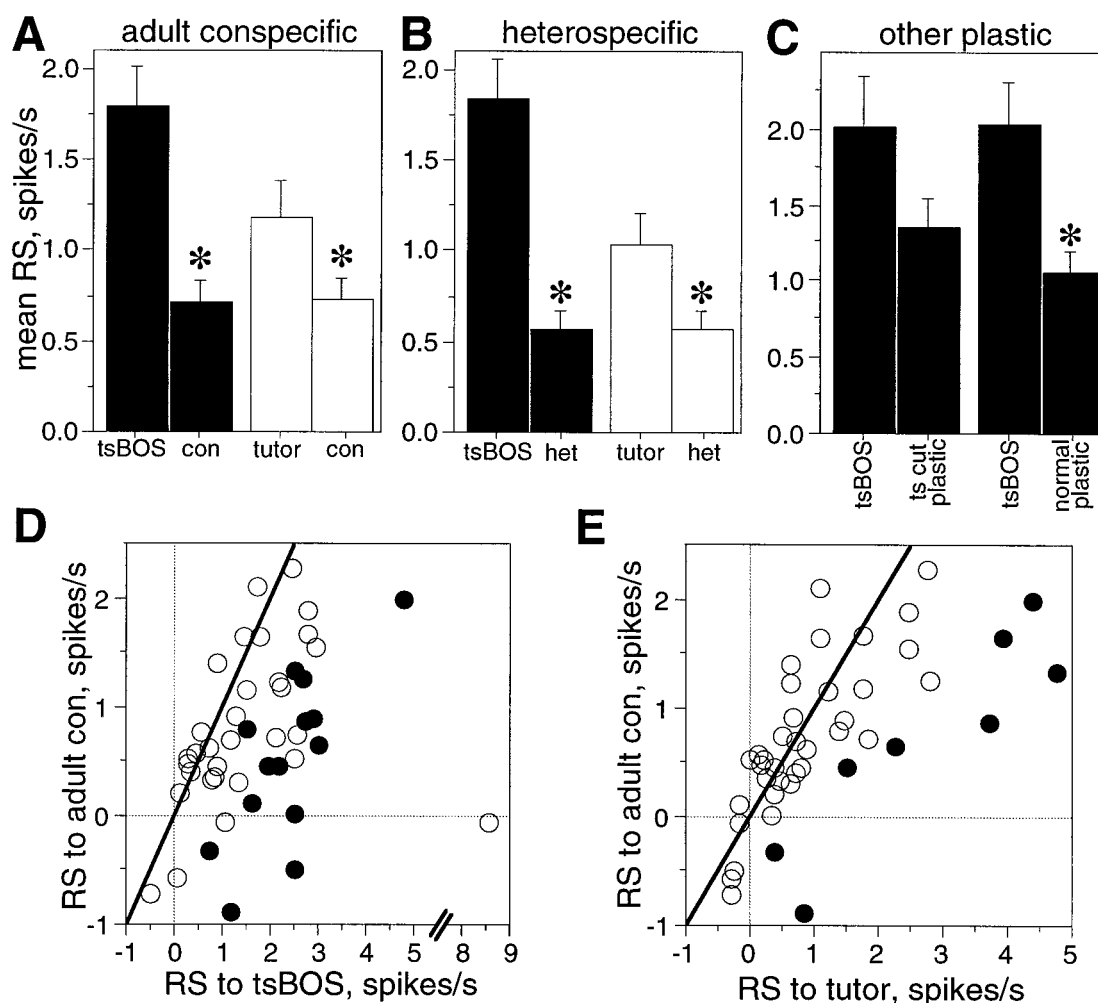
To test whether neurons were tuned specifically to tsBOS, rather than to the noisy, immature features common to all plastic songs, other plastic songs of ts cut and normal 60 d birds were presented. On average, neurons responded more to tsBOS than to other plastic songs; however, this reached statistical significance for only the tsBOS–normal plastic song comparison (Fig. 7*C*) (paired *t* test,  $p < 0.0001$  for tsBOS–normal plastic;  $n = 32$ ;  $p = 0.055$  for tsBOS–ts cut plastic;  $n = 28$ ). Thus, LMAN neurons were tuned to features specific to tsBOS.

As a population, LMAN neurons from ts cut birds were also order-selective. A neuron is considered order-selective when it responds significantly more to forward song than to a song that is completely reversed (see labels in Fig. 8*A*). On average, LMAN neurons responded significantly more to tsBOS and tutor song than to reversed versions of these songs (Fig. 8*A*) (paired *t* test,  $p < 0.002$  for tsBOS–reverse;  $n = 42$ ;  $p < 0.013$  for tutor–reverse;  $n = 30$ ). The order selectivity of individual LMAN neurons is shown by plotting each neuron's RS to tsBOS (Fig. 8*D*) or tutor song (Fig. 8*E*) against its RS to the corresponding reverse song. In these scatterplots, many cells lie below the diagonal line, indicating their stronger responses to tsBOS or tutor song than to the corresponding reverse song stimuli.

#### Features important to order selectivity

To test the importance of syllable sequence within a song for order selectivity, “reverse order” stimuli were presented. Reverse order songs maintain the temporal order within individual syllables but reverse the syllable sequence within a song (see labels in





**Figure 7.** Song selectivity of the entire population of LMAN neurons recorded in ts cut birds. Paired comparisons of mean RS show that neurons responded more to tsBOS and tutor song than to adult conspecific (*A*) and heterospecific song (*B*). *C*, Paired comparisons also show greater responses to tsBOS than to ts cut and normal 60 d songs. In *A–C*, error bars indicate SEM, and asterisks mark significant differences between song pairs. *D*, The mean RS to tsBOS of each neuron is plotted against its mean RS to adult conspecific song (*adult con*). The diagonal line marks where cells lie if the RS to the two stimuli were equal. Black circles indicate those neurons with significantly greater responses to the stimulus on the abscissa ( $p < 0.05$ , unpaired *t* test between *abscessa* stimulus trials and all adult conspecific trials). *E*, The mean RS to tutor song of each neuron is plotted against the mean RS to adult conspecific song. Conventions are as in *D*.

Fig. 8*B*). On average, cells responded significantly more to forward tsBOS and tutor song than to reverse order versions of these songs (Fig. 8*B*) (paired *t* test, for tsBOS–reverse order,  $p < 0.010$ ;  $n = 29$ ; for tutor–reverse order,  $p < 0.050$ ;  $n = 33$ ). Thus, cells were sensitive to the syllable sequences within tsBOS and tutor song.

Because of the simple harmonic stack structure of syllables in many tsBOS, it seemed possible that neurons would be insensitive to reversal of the temporal structure within syllables from tsBOS. To test the contribution of individual syllable structure to order selectivity for tsBOS, we also presented “syllable reverse” stimuli. Syllable reverse stimuli maintain the correct syllable sequence within a song but reverse the individual syllables (see labels in Fig. 8*C*). On average, cells responded significantly more to forward tsBOS than to syllable reverse tsBOS (Fig. 8*C*) (paired *t* test,  $p < 0.003$ ;  $n = 13$ ). Thus, cells were also sensitive to the temporal structure within the simpler tsBOS syllables. The percentage of selective neurons in each order selectivity category is listed in Table 1.

### Comparison of X neural responses to tsBOS and tutor song

X is the first nucleus in the AF pathway; it receives inputs from HVC and itself projects to DLM, which in turn goes to LMAN. In addition, X receives feedback via projections from LMAN. To understand the circuitry underlying AF selectivity and potential interactions between LMAN and X, 64 single X neurons were also recorded from 19 ts cut birds.

As in LMAN, some X neurons responded more to tsBOS than to tutor song. The neuron in Figure 9*A* not only strongly preferred tsBOS over tutor song, but it also preferred tsBOS over adult conspecific song and reverse tsBOS. In addition, many X neurons responded equally well to tsBOS and tutor song, despite the acoustic dissimilarity of these songs; an example of such a neuron is illustrated in Figure 10*A*. The distribution of  $d'_{\text{tsBOS-tutor}}$  values from individual X neurons is shown in Figure 9*B*: 37% of X neurons recorded preferred tsBOS over tutor song; 35% responded equally well to tsBOS and tutor song; and 28%

**Table 1. Frequencies of selective cells**

Selectivity category	LMAN	X
tsBOS > tutor	12/43 (28%)†	22/60 (37%)
tsBOS > adult conspecific	22/39 (56%)	31/63 (49%)
tsBOS > juvenile conspecific	14/30 (47%)	28/51 (55%)
tsBOS > ts cut juvenile*	7/28 (25%)	28/48 (58%)
tsBOS > heterospecific	27/38 (71%)	36/62 (58%)
tsBOS > reverse tsBOS	15/37 (41%)**	27/54 (50%)**
tsBOS > reverse order tsBOS	4/27 (15%)	10/48 (21%)**
tsBOS > syllable reverse tsBOS	4/15 (27%)	6/15 (40%)
Tutor > tsBOS*	2/43 (5%)	17/60 (28%)
Tutor > adult conspecific	7/29 (24%)	27/61 (44%)
Tutor > heterospecific	16/28 (57%)	31/60 (52%)
Tutor > reverse tutor	10/21 (48%)	12/44 (27%)**
Tutor > reverse order tutor	6/22 (27%)	13/44 (30%)

The ratio of selective cells over the total tested in each category is shown for LMAN and X cells, followed by the percentages. A cell was considered selective if it had a  $d'$  value of at least 0.5 for the two stimuli compared.

\*Comparisons for which the frequency of selective cells was significantly different between LMAN and X ( $\chi^2$  tests,  $p < 0.05$ ).

\*\* Comparisons for which the frequency of selective cells was significantly lower in ts cut than in normal 60 d birds ( $\chi^2$  tests,  $p < 0.05$ ; normal 60 d data from Solis and Doupe, 1997).

preferred tutor song over tsBOS. This distribution did not differ significantly from that obtained from X neurons in normal 60 d birds (unpaired  $t$  test,  $p = 0.711$ ; normal 60 d data from Solis and Doupe, 1997). On average, in ts cut birds, neural responses to tsBOS were not significantly different from those to tutor song (Fig. 9*B*, inset) (paired  $t$  test,  $p = 0.862$ ;  $n = 63$ ).

#### Plastic song renditions and neuronal maturity

Neurons that did not strongly prefer tsBOS were unlikely to have resulted from inappropriate tsBOS choice; X neurons responded equally well to three different renditions of tsBOS (ANOVA,  $p = 0.079$ ;  $n = 38$ ). Furthermore, neurons with similar responses to tsBOS and tutor song were also selective, indicating that they were not immature. For example, the neuron in Figure 10*A* responded strongly to both tsBOS and tutor song but substantially less to conspecific song and reverse tsBOS. The song selectivity of neurons that responded equivalently to tsBOS and tutor song was examined by plotting the  $d'_{\text{tsBOS-tutor}}$  value of each neuron against its corresponding  $d'_{\text{tsBOS-adult con}}$  value and  $d'_{\text{tutor-adult con}}$  value (Fig. 10*B*). In this scatterplot, many neurons responding equally well to tsBOS and tutor song discriminated between these songs and adult conspecific song, as demonstrated by their  $d'_{\text{tsBOS-adult con}}$  or  $d'_{\text{tutor-adult con}}$  values of at least 0.5. Similar order selectivity was apparent for these neurons when  $d'_{\text{tsBOS-tutor}}$  values were plotted against  $d'_{\text{tsBOS-rev}}$  and  $d'_{\text{tutor-rev}}$  values (data not shown). Figure 10*C* shows that 86% of X neurons responding equally well to tsBOS and tutor song were classified as selective; in comparison, 63% of this neuron type were considered selective in normal 60 d birds (Solis and Doupe, 1997). In ts cut birds, only 2 X cells (each from a different animal) resembled those from 30 d birds. When neurons with equivalent responses to tsBOS and tutor song were analyzed as a population, they were also selective. On average, these X neurons responded significantly more to tsBOS than to adult conspecific (Fig. 10*D*) and reverse songs (Fig. 10*E*) (paired  $t$  test, for tsBOS–adult conspecific,  $p < 0.001$ ;  $n = 21$ ; for tsBOS–reverse,  $p < 0.0004$ ;  $n = 18$ ). They also responded more on average to tutor song than to adult

conspecific and reverse songs, but this was only significant for the tutor–adult conspecific comparison (paired  $t$  test,  $p < 0.002$ ;  $n = 21$ ; for tutor–reverse,  $p = 0.124$ ,  $n = 18$ ). Thus, X neurons with similar responses to tsBOS and tutor song were not immature; their selectivity was apparent both in individual neurons and in most population measures.

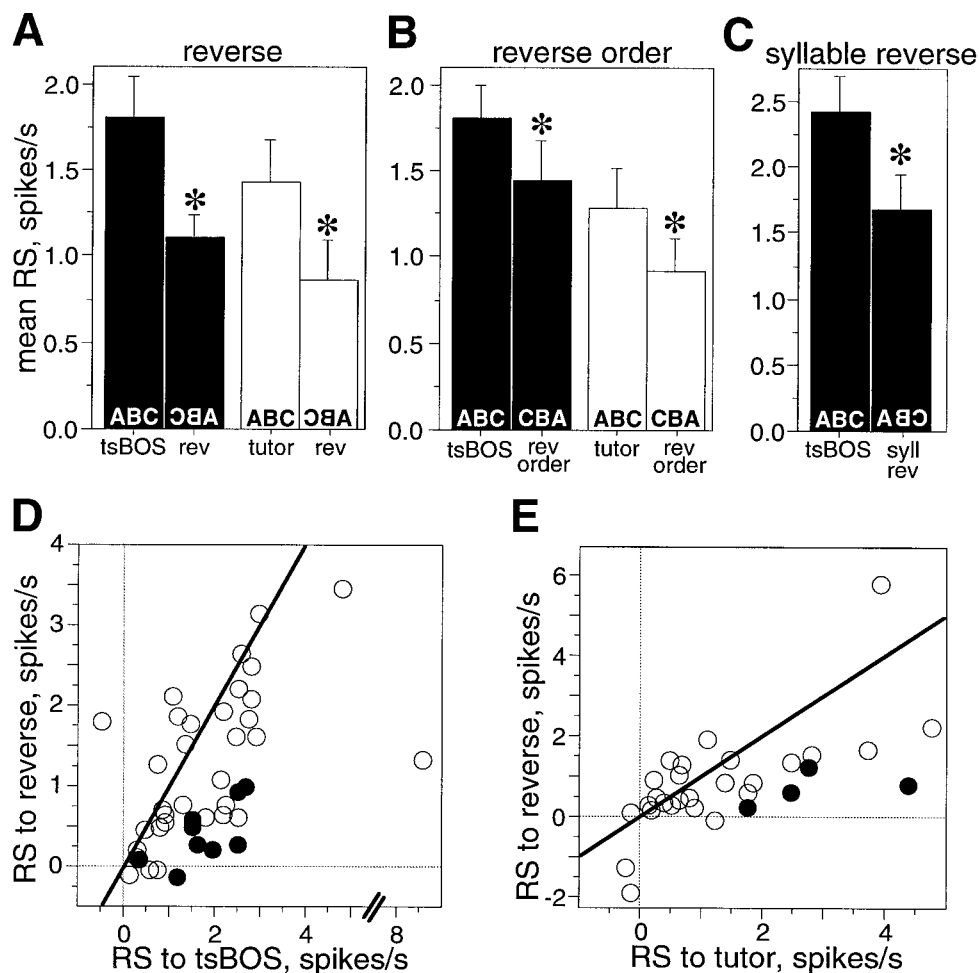
#### Alternative methods of measuring neural selectivity

The effect of differences between tsBOS and tutor song duration was also analyzed for X neurons; as for LMAN cells, the basic selectivity described above persisted. When RS was calculated by normalizing by song duration, the resulting  $d'_{\text{tsBOS-tutor}}$  values of individual X neurons correlated well with the relative difference in duration between tsBOS and tutor song ( $r^2 = 0.639$ ;  $p < 0.0003$ ). When data from experiments with similar tsBOS and tutor song durations were analyzed separately (33 X neurons from 12 experiments), only four neurons (13%) preferring tutor song over tsBOS remained (compare with 28% in the entire data set). Also, in this data subset, 40% of the neurons preferred tsBOS over tutor song, and 47% responded equally well to both. For the latter type of neuron, 79% (11 of 14) were classified as selective, and this population responded on average more to tsBOS and tutor song than to adult conspecific or reverse songs (data not shown) (paired  $t$  test, for tsBOS–adult conspecific,  $p < 0.005$ ;  $n = 14$ ; for tsBOS–reverse,  $p < 0.013$ ;  $n = 11$ ; for tutor–adult conspecific,  $p < 0.018$ ;  $n = 14$ ); however, the response to tutor song was also not significantly different from that to reverse (paired  $t$  test,  $p = 0.078$ ;  $n = 11$ ).

Peak RS and peak  $d'_{\text{tsBOS-tutor}}$  values were also calculated for all X neurons. When classifying preferences using peak  $d'_{\text{tsBOS-tutor}}$  values, only 3% of neurons preferred tutor song over tsBOS. In addition, 32% of X neurons preferred tsBOS over tutor song, and 65% responded equally well to both (Fig. 6*B*). Importantly, 64% of this latter type of neuron were considered selective given their peak  $d'$  values in the four selectivity categories. These neurons were also selective when analyzed as a population; their averaged peak RS to tsBOS and tutor song were greater than those to adult conspecific (Fig. 6*C*) and reverse songs (data not shown) (paired  $t$  tests: for tsBOS–adult conspecific,  $p < 0.0001$ ;  $n = 39$ ; for tutor–adult conspecific,  $p < 0.0002$ ;  $n = 39$ ; for tsBOS–reverse,  $p < 0.0001$ ;  $n = 31$ ). As with the original RS measurement, the greater peak response to tutor song relative to reverse was not statistically significant (paired  $t$  test,  $p = 0.097$ ;  $n = 31$ ). Thus, peak measurements found that neurons with similar responses to tsBOS and tutor song were still selective according to individual neuron and most population measurements of selectivity. Although peak  $d'_{\text{tsBOS-tutor}}$  values reclassified 50% of X neurons in terms of their tsBOS versus tutor song preference, the resulting distribution of peak  $d'_{\text{tsBOS-tutor}}$  values was similar to that of the original  $d'_{\text{tsBOS-tutor}}$  values (paired  $t$  test,  $p = 0.114$ ;  $n = 60$ ).

#### Song and order selectivity of the entire population of X neurons

When the entire population of X neurons recorded was considered, regardless of their tsBOS versus tutor song preferences, they were song selective for both tsBOS and tutor song. Using the original RS values normalized by stimulus duration, X neurons responded significantly more to tsBOS and tutor song than to adult conspecific (Fig. 11*A*) and heterospecific song (Fig. 11*B*) (paired  $t$  test,  $p < 0.0001$  for tsBOS–conspecific;  $n = 63$ ; tsBOS–heterospecific;  $n = 64$ ; and tutor–heterospecific;  $n = 63$ ;  $p <$



**Figure 8.** Order selectivity of the population of LMAN neurons recorded from ts cut birds. *A*, Paired comparisons of mean RS show that neurons responded more to tsBOS and tutor song in the forward direction than to their respective reverse songs. *B*, The mean RS to tsBOS and tutor song were greater than those to reverse order versions of these songs. *C*, The mean RS to tsBOS was greater than to the syllable reverse version of tsBOS. In *A–C*, error bars indicate SEM, and asterisks mark significant differences between song pairs. *D*, The mean RS to tsBOS of each neuron is plotted against its mean RS to reverse tsBOS. The diagonal line shows where cells lie when they respond equally to the two stimuli compared. Black circles indicate those cells that had significantly greater RS to the stimulus on the abscissa than to the reverse manipulation ( $p < 0.05$ , unpaired  $t$  test between forward song trials and corresponding reverse song trials). *E*, The mean RS to tutor song of each neuron is plotted against its mean RS to reverse tutor song. Conventions are as in *D*.

0.0004 for tutor-conspecific;  $n = 62$ ). The song selectivity of individual X neurons is illustrated with scatterplots, which compare the RS to adult conspecific song of each neuron with its RS to tsBOS (Fig. 11*D*) and to tutor song (Fig. 11*E*).

Further tests of tsBOS song selectivity indicated that neurons were tuned to tsBOS specifically, rather than to features common to plastic songs. Responses to tsBOS were significantly greater than responses to plastic songs of ts cut and normal 60 d birds (Fig. 11*C*) (paired  $t$  test,  $p < 0.0005$  for tsBOS–ts cut;  $n = 49$ ;  $p < 0.0002$  for tsBOS–normal;  $n = 51$ ). The percentages of selective X cells in each song selectivity category are listed in Table 1.

The X neurons recorded were also order-selective for tsBOS (Fig. 12*A*). On average, X neurons responded significantly more to tsBOS than to reverse tsBOS (paired  $t$  test,  $p < 0.0001$ ;  $n = 54$ ). Although the average response to tutor song was slightly more than to that to reverse tutor song, this difference was not statistically significant (paired  $t$  test,  $p = 0.144$ ;  $n = 45$ ). The order selectivity of individual neurons is shown by plotting each neuron's mean RS to tsBOS (Fig. 12*D*) and to tutor song (Fig. 12*E*) against its mean RS to the corresponding reverse stimulus.

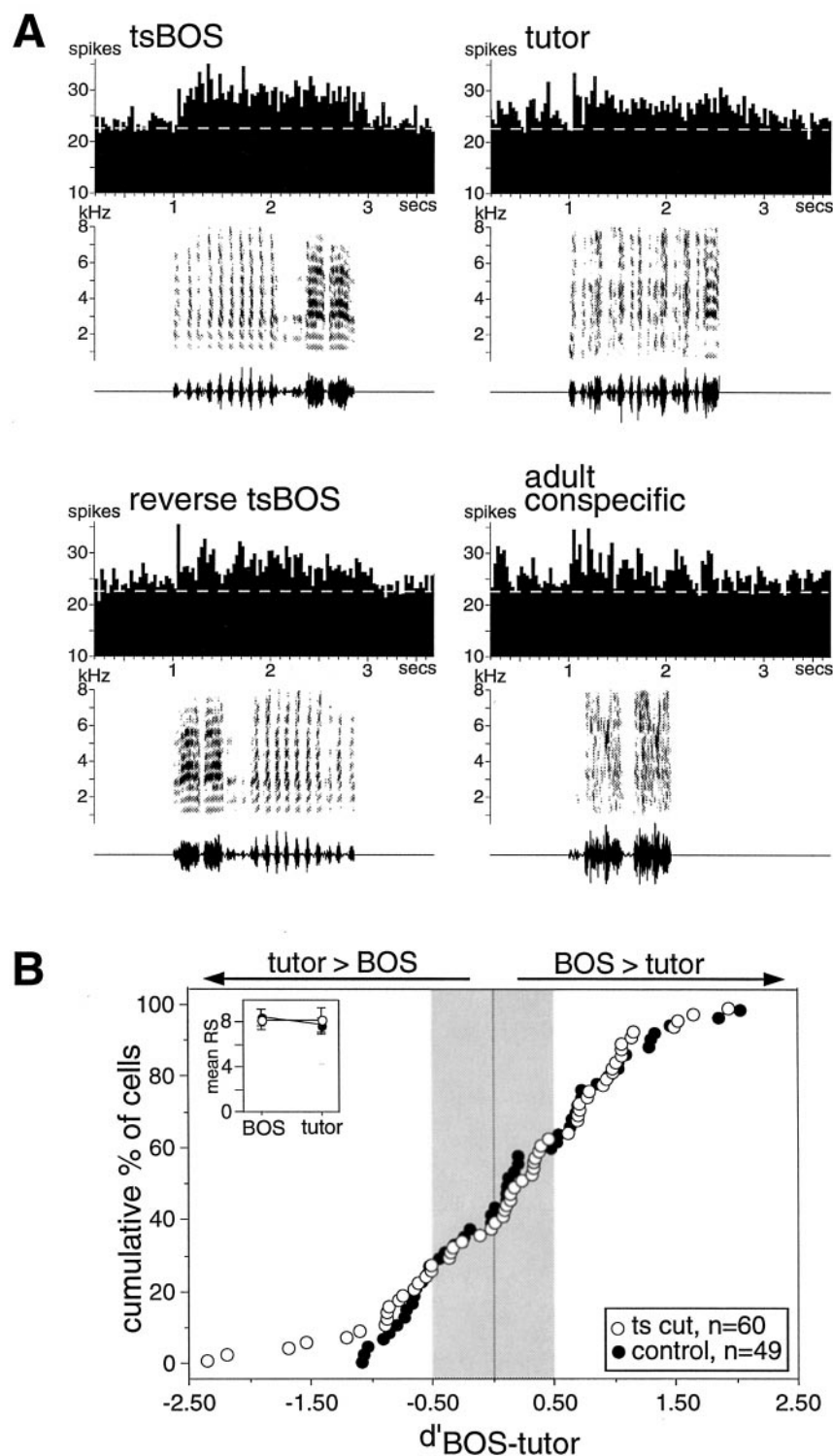
To test the importance of syllable sequence on X neuron order selectivity, reverse order stimuli were presented (Fig. 12*B*). In contrast to LMAN, X neurons did not discriminate between tsBOS and reverse order tsBOS (paired  $t$  test,  $p = 0.411$ ;  $n = 48$ ). This suggests that these neurons become selective for syllable identity first and then later for the syllable sequence within song. However, X neurons did respond significantly more to forward

than to reverse order tutor song ( $p < 0.001$ ;  $n = 45$ ), indicating that they were sensitive to the syllable sequences within tutor song. This difference between forward and reverse order tutor song was small, however, and similar to that between forward and reverse tutor song. This suggests that discrimination between the two reverse manipulation types was not really that different for X cells; consistent with this, the percentage of selective neurons in the tutor–reverse category was similar to neurons in the tutor–reverse order category (Table 1). Finally, to assay the contribution of temporal features within tsBOS syllables to X responses, syllable reverse stimuli were presented. Neurons responded significantly more to forward tsBOS than to syllable reverse tsBOS, indicating a sensitivity to the temporal structure within a syllable (Fig. 12*C*) (paired  $t$  test,  $p < 0.029$ ;  $n = 17$ ). The percentages of selective X cells in each order selectivity category are listed in Table 1.

### Comparisons of LMAN and X neurons

X neurons were more broadly responsive than LMAN neurons; they responded readily to broad-band noise bursts, tone bursts, and “nonpreferred” song stimuli (i.e., those not eliciting the largest RS). In X, 53% (34 of 64) of cells recorded responded significantly to all nonpreferred stimuli presented, whereas in LMAN, only 4% (2 of 52) of neurons did. When both nuclei were sampled in an individual bird (13 experiments), mean selectivity values for each nucleus were calculated. When selectivity was



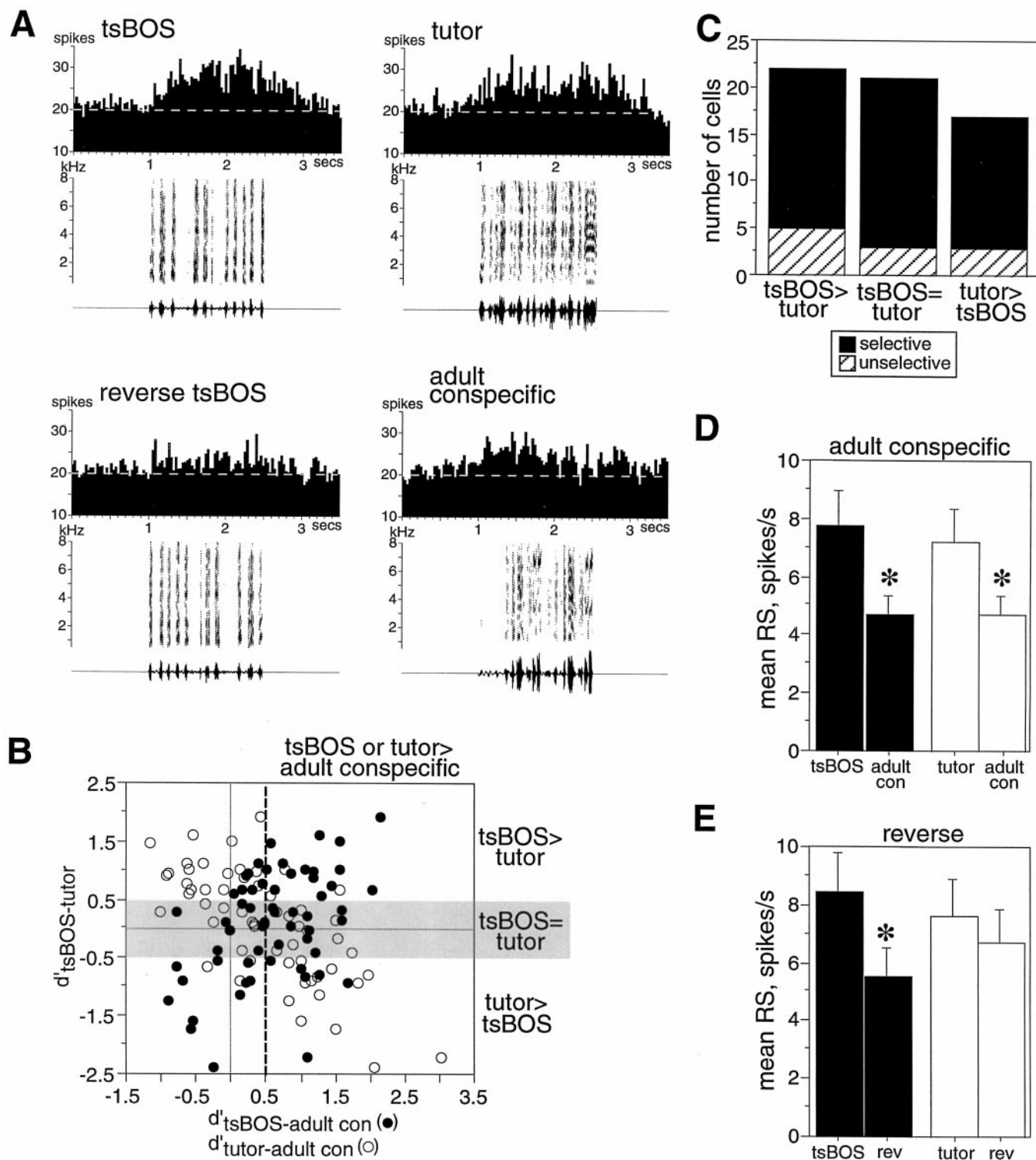


**Figure 9.** Selectivity for tsBOS in X. *A*, PSTHs show the responses of a single X neuron to 20 presentations of each stimulus. This neuron responded more to tsBOS than to tutor song, reverse tsBOS, and an adult conspecific song ( $d'_{\text{tsBOS-tutor}} = 1.52$ ;  $d'_{\text{tsBOS-rev}} = 0.80$ ; and  $d'_{\text{tsBOS-adult con}} = 2.35$ ). The dashed white line indicates the neuron's average spontaneous firing rate. Note that the ordinate of the PSTHs begins at 10 spikes/sec. *B*, Cumulative distributions of  $d'_{\text{tsBOS-tutor}}$  values of individual X neurons from ts cut birds (white circles) and normal 60 d birds (black circles). Inset, Mean RS to BOS and tutor song of the population of X neurons recorded from ts cut (white circles) and normal (black circles) 60 d birds. Error bars indicate SEM.

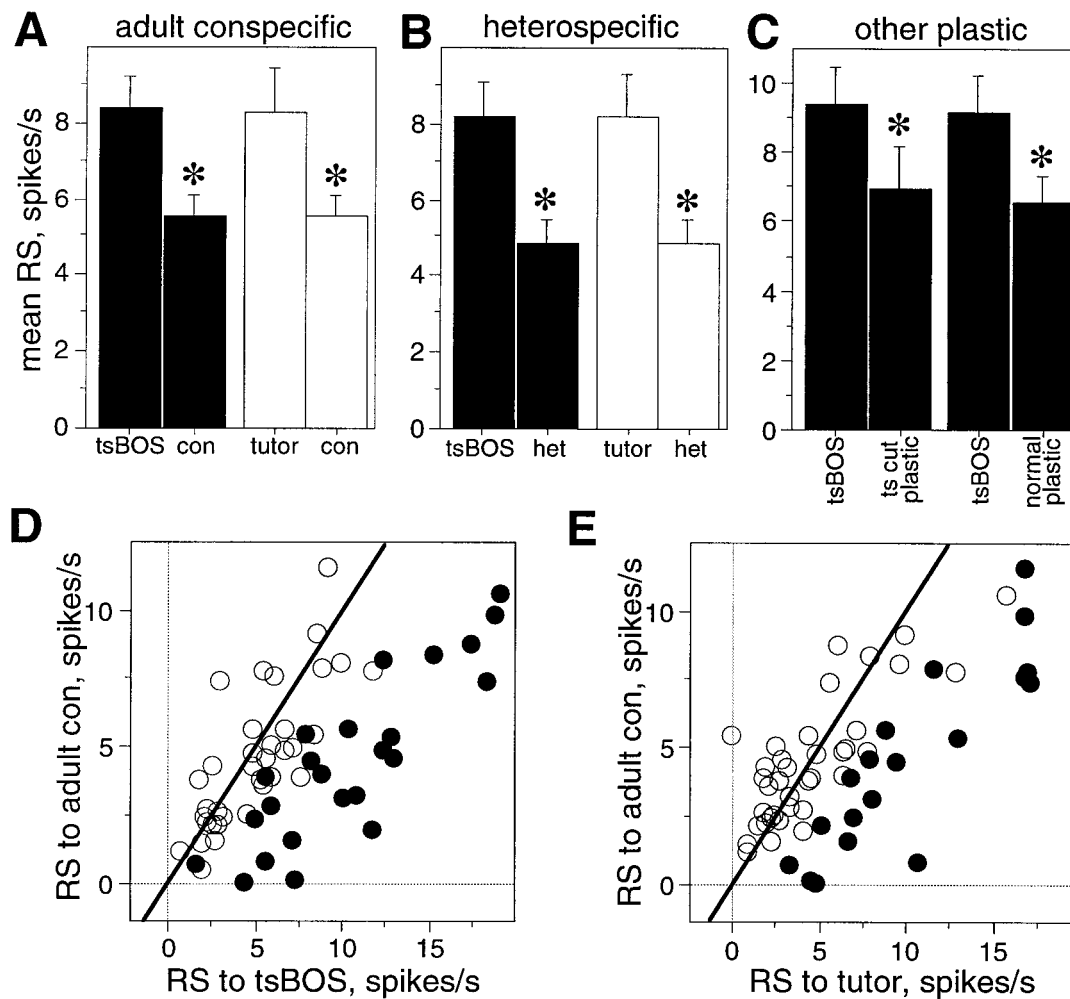
measured by the ratio of response magnitudes to two stimuli (SI; see Materials and Methods), LMAN was significantly more selective than X for only two categories of selectivity: tsBOS–tutor song and tsBOS–adult conspecific song comparisons (data not shown; paired  $t$  test,  $p < 0.006$  and  $p < 0.002$ , respectively). In contrast, comparisons of mean  $d'$  values across nuclei yielded no significant differences in selectivity. Thus, in general, LMAN and X shared similar degrees of selectivity to song stimuli.

#### Comparisons between neural properties of ts cut birds and those of normal juvenile birds

Although the distribution of their responses to tsBOS and tutor song were similar to those obtained from normal 60 d birds (Figs. 4*B*, 9*B*), AF neurons from ts cut birds were less selective than neurons from normal 60 d birds. The mean  $d'$  values for different selectivity categories are compared between ts cut and normal 60 d birds in Figure 13 (normal 60 d data from Solis and Doupe,



**Figure 10.** Some X neurons responded equally well to tsBOS and tutor song. *A*, PSTHs made from the responses of a single X neuron to 10 presentations of each stimulus. This neuron responded more to tsBOS and tutor song than to reverse tsBOS or adult conspecific. For the responses shown,  $d'_{\text{tsBOS-tutor}} = 0.43$ ;  $d'_{\text{tsBOS-rev}} = 1.43$ ;  $d'_{\text{tsBOS-adult con}} = 1.06$ ; and  $d'_{\text{tutor-adult con}} = 0.96$ . The white dashed line indicates the neuron's average spontaneous firing rate. Note that the ordinate of the PSTHs begins at 10 spikes/sec. This particular tsBOS was matched to the correct tutor song by only one of nine observers. *B*, The  $d'_{\text{tsBOS-tutor}}$  value of each X neuron is plotted against two measures of selectivity:  $d'_{\text{tsBOS-adult con}}$  (black circles) and  $d'_{\text{tutor-adult con}}$  (open circles). The gray region highlights those neurons considered to have responded equally well to tsBOS and tutor song. The dashed vertical line marks the criterion for selectivity ( $d' = 0.5$ ). *C*, Number of X neurons classified as selective (black bars) and unselective (hatched bars) in the three different tsBOS versus tutor song preference categories. *D*, For those neurons responding equally well to both tsBOS and tutor song, histograms show paired comparisons of the mean RS to tsBOS (black bars) or tutor song (white bars) to adult conspecific song. *E*, For those neurons responding equally well to tsBOS and tutor song, histograms show paired comparisons of the mean RS to tsBOS (black bars) or tutor song (white bars) and their corresponding reverse songs. In *D* and *E*, error bars indicate SEM, and asterisks denote significant differences between each pair of stimuli.



**Figure 11.** Song selectivity of the entire population of X neurons recorded in ts cut birds. Paired comparisons of mean RS show that neurons responded more to tsBOS and tutor song than to adult conspecific (**A**) and heterospecific (**B**) song. **C**, Paired comparisons also show greater responses to tsBOS than to ts cut and normal 60 d songs. In **A–C**, error bars indicate SEM, and asterisks indicate significant differences between song pairs. **D**, The mean RS to tsBOS of each neuron is plotted against its mean RS to adult conspecific song (*adult con*). The diagonal line marks where cells lie if the RS to each stimulus were equal. Black circles indicate those neurons with significantly greater responses to stimulus on the abscissa ( $p < 0.05$ , unpaired  $t$  test between abscissa stimulus trials and all adult conspecific trials). **E**, The mean RS to tutor song of each neuron is plotted against the mean RS to adult conspecific song. Conventions are as in **D**.

1997). The mean  $d'$  values of LMAN neurons from ts cut birds were significantly lower than those from normal 60 d birds when tsBOS selectivity was analyzed (unpaired  $t$  tests,  $p < 0.0001$  for tsBOS–reverse;  $p < 0.001$  for tsBOS–adult conspecific). The mean  $d'$  values of X neurons from ts cut birds were significantly lower than those from normal 60 d birds when order selectivity was examined (unpaired  $t$  test,  $p < 0.009$  for tsBOS–reverse;  $p < 0.002$  for tutor–reverse). This lower selectivity relative to normal 60 d birds was also maintained when only data from ts cut and normal birds with the same tutor (different clutches) were compared (data not shown). Thus, the lower selectivity is probably not attributable to differences in tutor bird efficacy between the normal and ts cut studies. Furthermore, this lower selectivity was apparent when the percentage of selective neurons in ts cut birds was compared with that found in normal 60 d birds. In LMAN, the percentages of selective neurons in the tsBOS > tutor and the tsBOS > reverse tsBOS categories were significantly lower in ts cut birds. In X, the percentages of selective neurons in the tsBOS > reverse tsBOS, tutor > reverse tutor, and tsBOS >

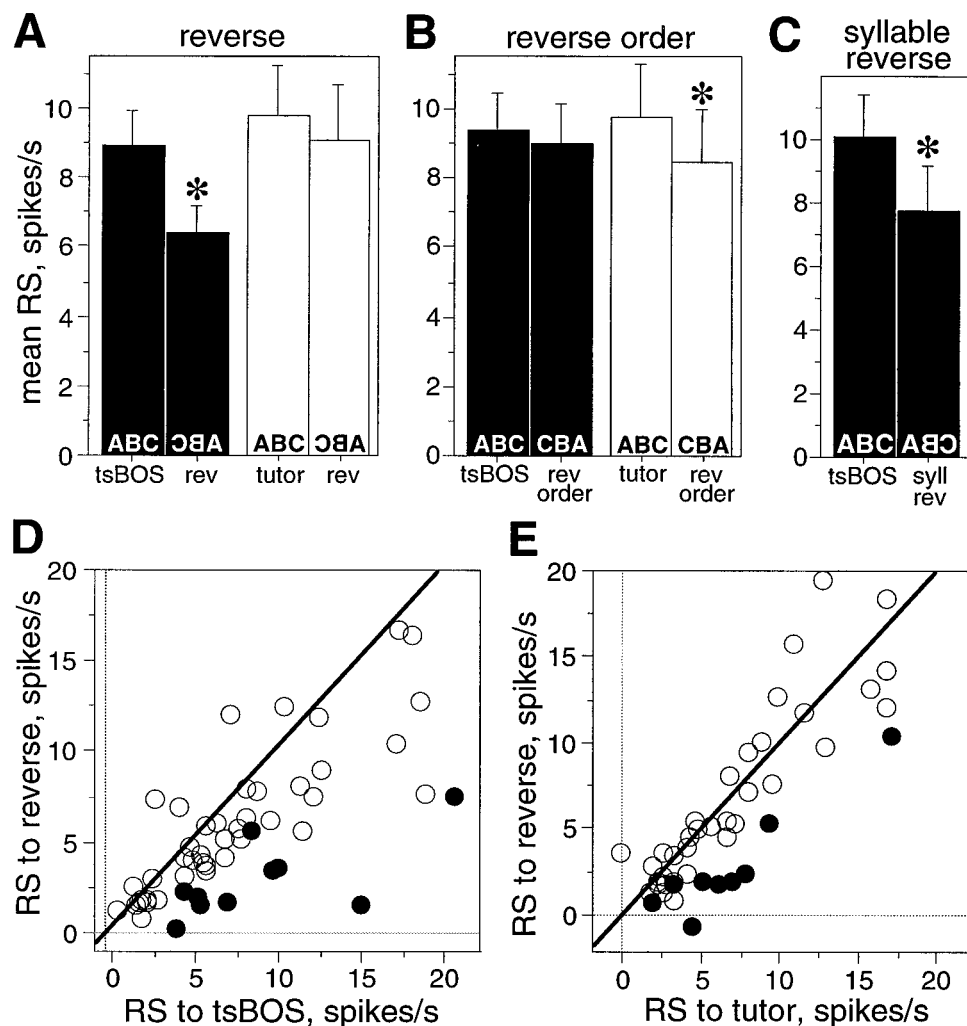
reverse order tsBOS categories were significantly lower in ts cut birds (Table 1).

Despite this lower selectivity, neurons from ts cut birds were clearly selective relative to neurons from normal 30 d birds; as seen in Figures 7, 8, 11, and 12, neurons from ts cut 60 d birds on average discriminated tutor song from adult conspecific song and reverse tutor song (except in X), which is not true for 30 d neurons. In addition, when classified according to tutor song categories of selectivity only (i.e.,  $d'_{\text{tutor-adult con}}$  and  $d'_{\text{tutor-rev}}$ ), ts cut birds had significantly more selective LMAN and X neurons than did 30 d birds (30 d data from Doupe, 1997;  $\chi^2$  tests,  $p < 0.0004$  for both comparisons).

#### RA volumes were not affected by NXIIIs transections

Transecting the NXIIIs nerve might have caused neuronal atrophy or death in upstream nuclei, which could lead to nonspecific changes in LMAN or X selectivity. To estimate potential retrograde effects of the nerve transections, RA volumes were measured. In ts cut birds, the mean  $\pm$  SD RA volume ( $0.244 \pm 0.053$





**Figure 12.** Order selectivity of the entire population of X neurons recorded in ts cut birds. Paired comparisons of mean RS show greater responses to tsBOS and tutor song than to reverse (*A*) and reverse order (*B*) versions of these songs. *C*, Paired comparisons show greater RS to tsBOS than to syllable reverse tsBOS. In *A–C*, error bars indicate SEM, and asterisks mark significant differences between song pairs. *D*, The mean RS of each X neuron to tsBOS is plotted against its mean RS to reverse tsBOS. The diagonal line indicates where cells would lie if their responses to the two stimuli were equal. Black circles indicate those cells whose responses to the stimulus on the abscissa were significantly greater than its responses to the ordinate stimulus ( $p < 0.05$ , unpaired *t* test). *E*, The mean RS of each X neuron to tutor song is plotted against its mean RS to reverse tutor song. Conventions are as in *D*.

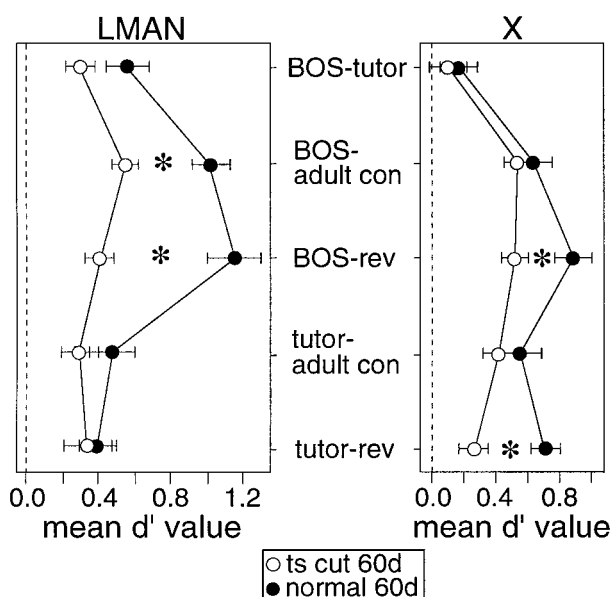
$\text{mm}^3$ ) was within the range of RA volumes previously reported for normal adults (0.220–0.372  $\text{mm}^3$ ; Gurney, 1981). Also, although syrinx weight of ts cut birds was significantly less than that of normal 60 d birds (unpaired *t* test,  $p < 0.0001$ ), the normalized RA volume was not significantly different between the same two groups of birds (unpaired *t* test,  $p = 0.300$ ); this indicates that nerve transection did not affect RA volume. Finally, the mean RA volume of a ts cut bird did not correlate with its mean  $d'_{\text{tsBOS-tutor}}$  value from LMAN or X (Table 2). Thus, possible retrograde effects of nerve transection, as measured here with RA volume, do not account for the difference in selectivity between ts cut and normal 60 d birds or for the range of tsBOS versus tutor song preferences.

#### Differences between individual ts cut birds

Neurons recorded from the same bird had similar tsBOS versus tutor song preferences (Fig. 14). Experimental  $d'_{\text{tsBOS-tutor}}$  values of LMAN neurons clustered in nine of nine birds in which multiple cells were recorded (see Materials and Methods). This frequency of clustering was greater than expected by chance (sign test,  $p < 0.002$ ). Similarly, the  $d'_{\text{tsBOS-tutor}}$  values of X neurons within a bird were also clustered more frequently than expected by chance (12 of 14 birds; sign test,  $p < 0.006$ ). The  $d'_{\text{tsBOS-tutor}}$  values obtained from both LMAN and X cells in a single bird were also more similar than expected (15 of 17 birds were clus-

tered;  $p < 0.001$ ). The similarity of  $d'_{\text{tsBOS-tutor}}$  values for neurons from the same bird suggests that factors specific to the experiment or to the bird could account for the responses to these two songs.

A bird's neural preference for tsBOS versus tutor song was not readily explained by conditions that varied between experiments (Table 2). Potential measures of anesthesia depth, such as the spontaneous rate, maximum RS of a neuron, or time at which each neuron was recorded relative to anesthesia administration, were not well correlated to the neuron's  $d'_{\text{tsBOS-tutor}}$  value. Anatomical location within the nucleus did not predict the  $d'_{\text{tsBOS-tutor}}$  value of a neuron in the anteroposterior, mediolateral, or dorsoventral dimension. Neither slight intensity differences between tsBOS and tutor song stimuli nor the ages of individual birds correlated well with the mean  $d'_{\text{tsBOS-tutor}}$  value from each bird. To weigh the contribution that each bird made to these correlations by the number of cells recorded from it, these last two comparisons were also made using the  $d'$  values of individual neurons; this did not improve their correlations (Table 2). Furthermore, no significant correlations were found among these variables when only data from experiments with equivalent tsBOS and tutor song durations were considered or when peak  $d'_{\text{tsBOS-tutor}}$  values were used (Table 2). The inclusion of immature neurons in these correlations could obscure a relation be-



**Figure 13.** The degree of selectivity of AF neurons from ts cut 60 d birds (white circles) is less than that observed in normal 60 d neurons (black circles). The mean  $d'$  values for each selectivity category (listed on the ordinate) are plotted along the abscissas; error bars indicate SEM. Asterisks denote significant differences between ts cut and normal 60 d birds.

tween the variables tested and  $d'_{\text{tsBOS-tutor}}$  values, because these cells would have equivalent responses regardless of any experimental condition tested here. Thus, these correlations were recalculated with the 10 unselective neurons excluded from the data set; this still did not reveal any significant correlations (data not shown). Thus, a bird's neural preference for tsBOS versus tutor song did not seem to depend on conditions that varied between experiments.

### Acoustic similarity between the tsBOS and tutor song does not correlate with neural responses to these two song stimuli

Because the equivalent responses to tsBOS and tutor song were not the result of inappropriate tsBOS stimulus choice nor of neuronal immaturity, we considered another possible explanation for such responses. In principle, residual acoustic similarities between tsBOS and tutor song might have produced similar responses to tsBOS and tutor song. Although the average similarity between tsBOS and tutor songs was lower than normal (Fig. 3D), it remained possible that this similarity for an individual ts cut bird predicted the song preference of that bird's neurons. If the  $d'_{\text{tsBOS-tutor}}$  values obtained from a bird reflect residual acoustic similarity between tsBOS and tutor songs, then neurons with equivalent responses to tsBOS and tutor song should come from birds with tsBOS similar to tutor song. As similarity decreases, neurons would have strong preferences for either tsBOS or tutor song, depending on the experience shaping selectivity (Fig. 15A, top panel). This "similarity hypothesis" predicts a negative correlation between the absolute value of  $d'_{\text{tsBOS-tutor}}$  values ( $|d'_{\text{tsBOS-tutor}}|$ ) and the measured similarity between tsBOS and tutor song (Fig. 15A, bottom panel).

### Matching task

To assess whether such a correlation existed in the data, we analyzed pairs of tsBOS and tutor songs to compare their acoustic

similarity to the corresponding neural data. Similarity between tsBOS and tutor song was analyzed in several ways. In the matching task described earlier, the percentage of observers correctly matching tsBOS to tutor song was used as a measure of overall similarity between the two songs. When this percentage was compared with the mean  $|d'_{\text{tsBOS-tutor}}|$  value of LMAN neurons obtained from each bird, no correlation was evident (Fig. 15B) ( $r^2 = 0.023$ ). Birds whose songs were infrequently matched to the correct tutor song, a sign of dissimilarity, still had neurons that responded equally well to tsBOS and tutor song.

Because the measure of percent correctly matched combines both spectral and temporal features of song, these features were also scored separately to control for the possibility that ts cut birds could imitate the timing of the tutor song without mimicking its spectral content. When the mean spectral similarity score for each bird was compared with the mean  $|d'_{\text{tsBOS-tutor}}|$  value from its LMAN neurons, no correlation resulted (Fig. 15C) ( $r^2 = 0.034$ ). Comparing temporal similarity scores with the mean  $|d'_{\text{tsBOS-tutor}}|$  values from LMAN neurons in each bird also failed to yield a correlation (Fig. 15D) ( $r^2 = 0.046$ ). Furthermore, when mean peak  $|d'_{\text{tsBOS-tutor}}|$  values were used, regression lines with slopes of the same sign were obtained, but the correlations were also weak (for percent of correct observers,  $r^2 = 0.0003$ ; for mean similarity score,  $r^2 = 0.008$ ; for mean timing score,  $r^2 = 0.132$ ).

### Cross-correlations

Because some similarities may have been too subtle for detection by human observers, we also used cross-correlation methods to analyze song pairs. Three types of cross-correlation measures were calculated between each tsBOS and its tutor song. First, each tsBOS spectrogram was cross-correlated with the corresponding tutor song spectrogram. Second, to analyze spectral similarity alone, all syllables from each tsBOS were isolated from the song and individually cross-correlated with the isolated syllables of the tutor song. Third, to compare temporal similarity, the amplitude envelopes of each song were cross-correlated. Figure 16A shows the mean cross-correlation measures obtained from tsBOS and tutor song comparisons. For reference, mean cross-correlation measures were also calculated between tutor song and other song types, including normal adult, normal 60 d, and randomly matched songs. Of these measures, only the syllables-only cross-correlation measures distinguished differences between these song types. Songs from ts cut 60 d birds had less similarity with their tutor songs than did normal 60 d songs (unpaired  $t$  test,  $p < 0.004$ ); however, the similarity between normal 60 d songs and tutor song was not significantly different from that between normal adult-tutor song pairs. In contrast, cross-correlation measures for the entire spectrogram and temporal envelope comparisons were uniformly low for all song types; no song type's mean cross-correlation measure was significantly different from that obtained from randomly matched song pairs. Thus, the entire spectrogram and temporal envelope cross-correlations were not sensitive enough to detect similarities that were apparent to humans (Fig. 3D). When the most sensitive syllables-only cross-correlation measure of similarity was compared with the mean  $|d'_{\text{tsBOS-tutor}}|$  value from LMAN neurons, no strong correlation resulted (Fig. 16B;  $r^2 = 0.128$ ). This was also true when mean peak  $|d'_{\text{tsBOS-tutor}}|$  values were used ( $r^2 = 0.021$ ).

Table 2. Correlations of  $d'$  values

Variable vs $d'_{\text{tsBOS-tutor}}$ value	LMAN $r^2$ values		X $r^2$ values	
	Original $d'$	Peak $d'$	Original $d'$	Peak $d'$
Background firing rate	0.210	0.018	0.036	0.075
Maximum RS	0	0.022	0.061	0.058
Experiment duration	0.036	0.040	0	0.001
Anteroposterior location within nucleus	0.056	0.045	0.117	0.007
Mediolateral location within nucleus	0.120	0.202	0.044	0.001
Dorsoventral location within nucleus	0	0.001	0.007	0.006
Relative intensity of stimuli				
Mean	0.031	0	0.002	0.027
Individual	0	0.002	0.001	0.028
RA volume				
Mean	0.106	0.010	0.192	0.025
Individual	0.043	0	0.105	0.009
Age				
Mean	0.209	0.010	0.002	0.107
Individual	0.271	0.006	0.053	0.042

The coefficient of determination ( $r^2$ ) is shown for the linear least squares fits of the data, which compare each independent variable with the  $d'_{\text{tsBOS-tutor}}$  value of a cell or group of cells (mean values). Entries of 0 refer to  $r^2$  values  $<0.001$ .

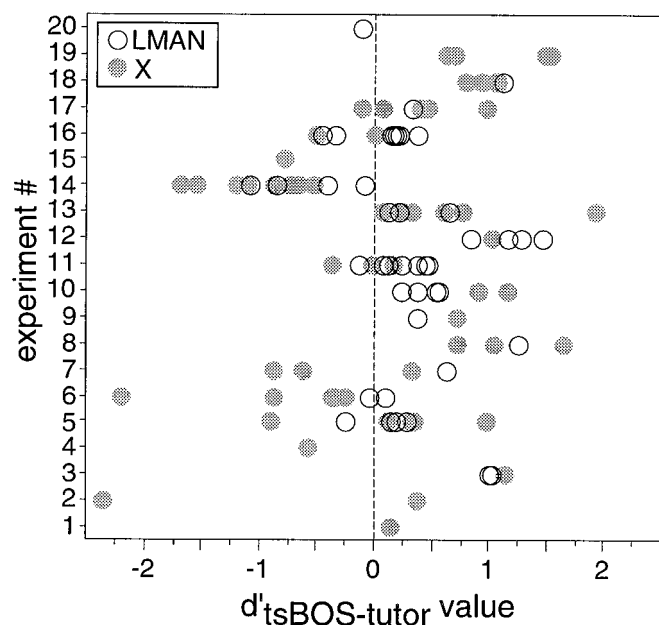


Figure 14. Clustering of  $d'_{\text{tsBOS-tutor}}$  values within each bird. The individual  $d'_{\text{tsBOS-tutor}}$  values obtained from each LMAN (open circles) and X (gray circles) cell recorded from the same bird are shown; each number on the ordinate refers to an individual bird.

### Overlap analysis

To further investigate temporal similarity between tsBOS and tutor song pairs, we used an overlap analysis (see Materials and Methods). The proportion of overlap between syllables and of intervals of an entire tsBOS string and entire tutor song string was calculated to give a song–song overlap value. To maximize the possibility of detecting temporal similarity, a motif–song overlap value was calculated between a single tsBOS motif and the entire tutor song (see Materials and Methods). Whereas the song–song overlap values did not distinguish between different

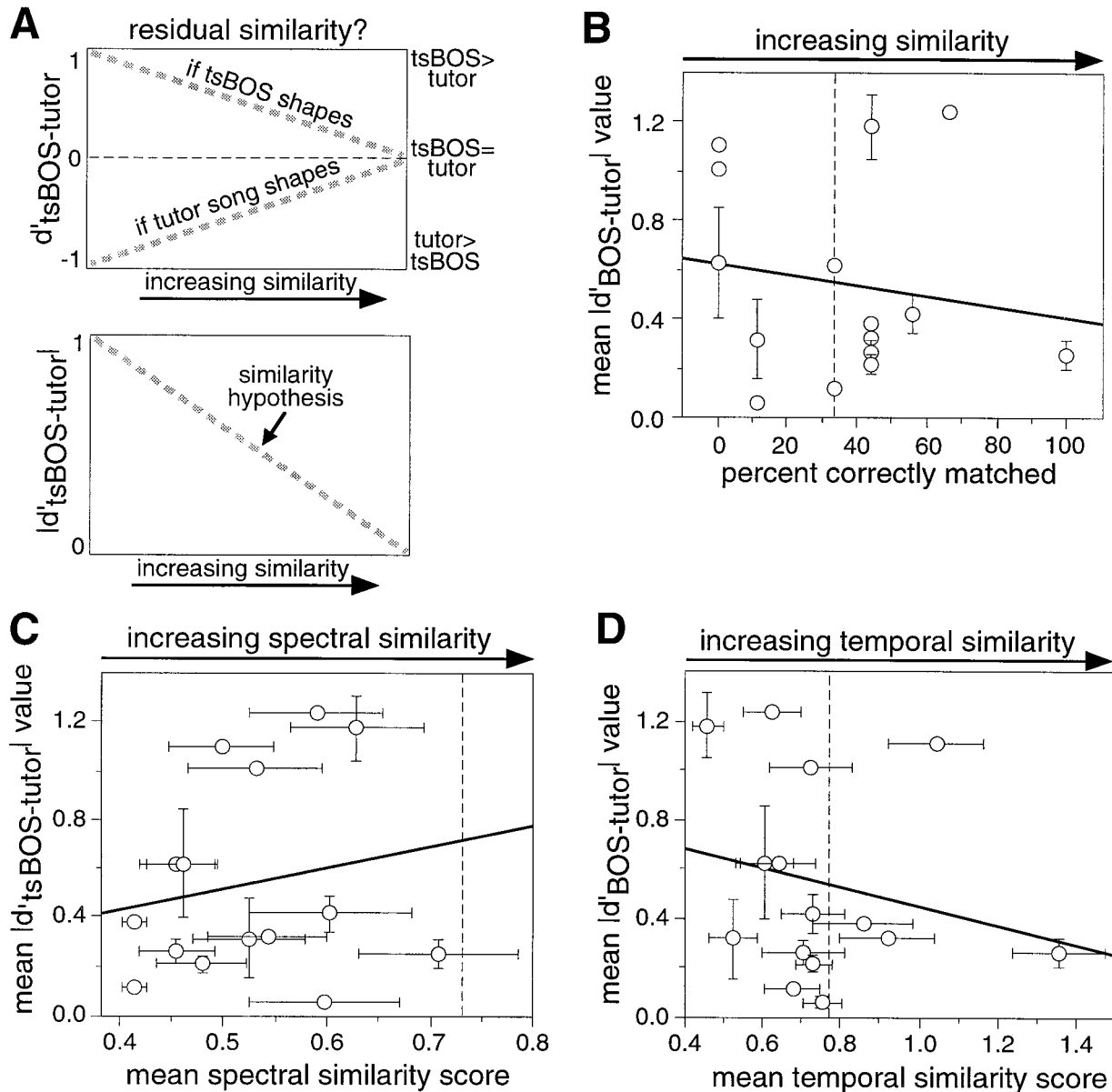
song types, the motif–song overlap values were slightly sensitive to differences in temporal similarity between them (Fig. 16A). Motif–song overlap values for ts cut–tutor song pairs were significantly less than those for normal adult–tutor song pairs (unpaired  $t$  test,  $p < 0.014$ ); this was the only significant difference between song types. When the motif–song overlap values were plotted against each bird's mean  $|d'_{\text{tsBOS-tutor}}|$  value, a weak correlation resulted (Fig. 16C) ( $r^2 = 0.341$ ;  $p < 0.022$ ); however, this positive correlation was in the opposite direction of that predicted by the similarity hypothesis. With increasing temporal similarity between tsBOS and tutor song, neurons tended to prefer tsBOS over tutor song. This correlation decreased when mean peak  $|d'_{\text{tsBOS-tutor}}|$  values were used instead ( $r^2 = 0.042$ ).

For the X data, comparisons of mean  $|d'_{\text{tsBOS-tutor}}|$  values and song similarity resulted in one weak negative correlation for the temporal similarity scores obtained from the matching task (data not shown;  $r^2 = 0.286$ ;  $p < 0.018$ ). Although this was predicted by the similarity hypothesis, it was not corroborated by the motif–song overlap measure of temporal similarity ( $r^2 = 0.015$ ). No other substantial correlations resulted for the other similarity measures (for percent of correct observers,  $r^2 = 0.163$ ;  $p = 0.087$ ; for spectral similarity,  $r^2 = 0.002$ ; for syllables-only cross-correlation,  $r^2 = 0.011$ ). When mean peak  $|d'_{\text{tsBOS-tutor}}|$  values were used, no strong correlations resulted, not even for temporal similarity scores (for percent of correct observers,  $r^2 = 0.006$ ; for mean similarity score,  $r^2 = 0.002$ ; for mean timing score,  $r^2 = 0.130$ ;  $p = 0.129$ ; for syllables-only cross-correlation,  $r^2 = 0.182$ ;  $p = 0.069$ ; for motif–song overlap,  $r^2 = 0.025$ ).

To weigh each bird's contribution to the correlation by the number of neurons recorded from the bird, the  $|d'_{\text{tsBOS-tutor}}|$  value or peak  $|d'_{\text{tsBOS-tutor}}|$  value of each cell was also compared with each measure of similarity. This did not reveal any substantial correlations (Table 3). Thus, the relative responses to tsBOS and tutor song were not strongly dependent on the similarity between tsBOS and tutor song as measured in this study.

Because correlations between  $d'_{\text{tsBOS-tutor}}$  values and measures of song similarity could be weakened by the presence of unselective neurons, the above correlations were recalculated excluding



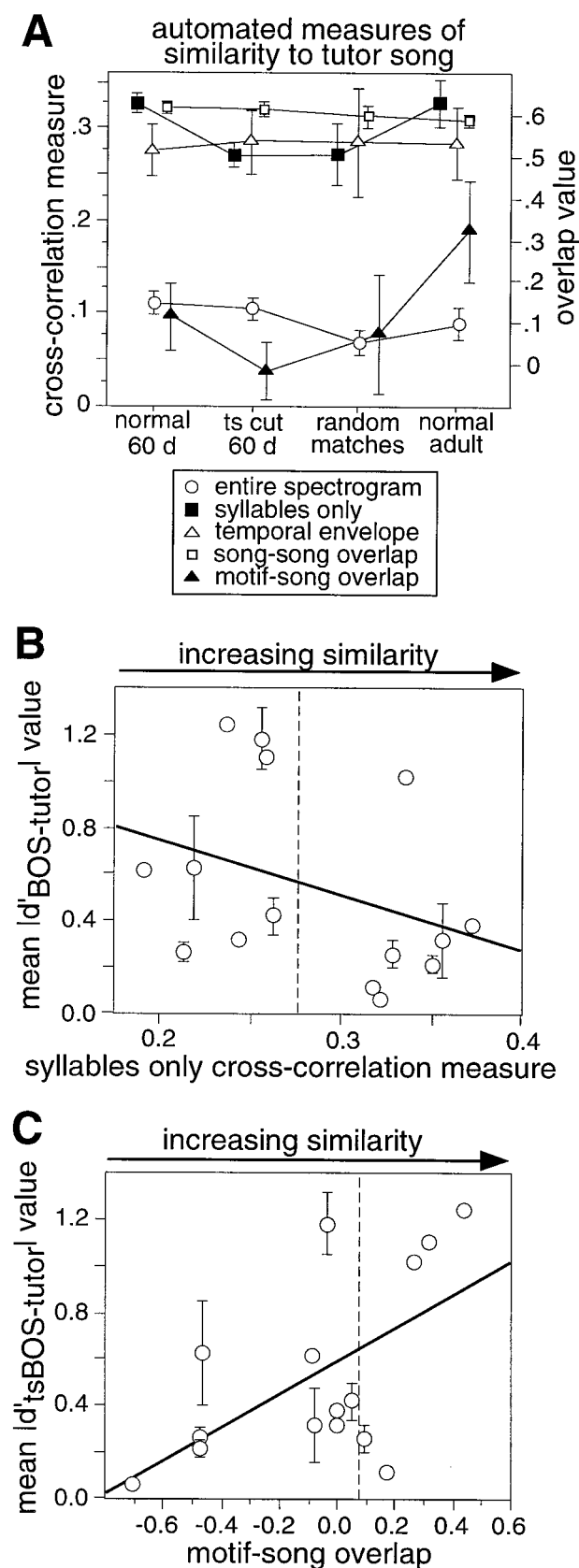


**Figure 15.** *A*, If residual acoustic similarity between tsBOS and tutor song accounts for the neural tsBOS versus tutor song preferences, then the trends shown with the gray dashed lines in the top panel are expected. For both scenarios, high similarity between tsBOS and tutor song is associated with low  $d'_{\text{tsBOS-tutor}}$  values. If tsBOS shapes neural selectivity, then low similarity should be associated with strong tsBOS preference over tutor song; conversely, if tutor song shapes selectivity, then low similarity should be associated with strong tutor song preference. The bottom panel combines these two trends by plotting the absolute values of  $d'_{\text{tsBOS-tutor}}$  against similarity. According to this similarity hypothesis, large mean  $|d'_{\text{tsBOS-tutor}}|$  values should occur in birds with low similarity to tutor song, and smaller mean  $|d'_{\text{tsBOS-tutor}}|$  values should occur in those with higher similarity to tutor song. *B*, The percent of observers who correctly matched the tsBOS song to the tutor song is plotted against the mean  $|d'_{\text{tsBOS-tutor}}|$  values obtained for each bird. The regression coefficient (slope) is  $-0.219 \pm 0.853$  ( $\pm 95\%$  confidence intervals). Points to the right of the dashed line refer to birds whose frequency of correct matches by nine observers was greater than expected by chance ( $p < 0.05$ , sign test). In *B–D*, error bars indicate SEM, and the thick black line is the linear least squares fit of the data. *C*, The mean spectral similarity score for each ts cut bird is plotted against the mean  $|d'_{\text{tsBOS-tutor}}|$  value for neurons recorded from that bird; the regression coefficient is  $0.869 \pm 2.781$ . The dashed line marks the mean spectral similarity score given to randomly matched songs. For the abscissa, the error bars refer to the variance in observer scoring; because of normalization, a value of 0.4 refers to low similarity, and a value of 2 refers to high similarity. *D*, The mean temporal similarity score for each ts cut bird is compared with the mean  $|d'_{\text{tsBOS-tutor}}|$  value for neurons recorded from that bird. The regression coefficient is  $-0.388 \pm 1.062$ . The dashed line marks the mean temporal similarity score given to randomly matched songs; the normalized scale of similarity is as described in *C*.

data from the bird that had contributed most of the unselective neurons (seven of eight unselective LMAN neurons came from this animal). The lack of correlation persisted for these comparisons; the strongest trend was a positive correlation between mean  $|d'_{\text{tsBOS-tutor}}|$  values from LMAN and motif-song overlap ( $r^2 = 0.316$ ;  $p < 0.037$ ; data not shown).

#### Song stereotypy shows little correlation with neuronal song preference

Rather than reflecting the acoustic properties of songs, the tsBOS versus tutor song preferences found in each bird might instead reflect its stage of song maturity. Although a weak correlation existed between mean  $d'_{\text{tsBOS-tutor}}$  values and age, the variability



**Figure 16.** Automated measures of similarity between experimental and tutor song pairs. **A**, Mean cross-correlation measures (left ordinate) for each song group are shown, obtained from the entire spectrogram (white circles), syllables only (black squares), and temporal envelope (white triangles).

in rate of song development between birds makes age a less reliable indicator of song maturity. Because song stereotypy also increases with song development, we used it as a more direct measure of song maturity. Song stereotypy was estimated by analyzing the similarity between multiple (usually 10) song samples from each bird. Stereotypy was measured using human subjective scoring, syllables-only cross-correlations, and motif-song overlap analyses (see Materials and Methods). Only the human scores and the overlap measures distinguished among all three song types (Fig. 17A). Songs from ts cut birds were significantly less stereotyped than normal 60 d songs (unpaired *t* test,  $p < 0.005$  for human scores;  $p < 0.003$  for overlap values). As expected, normal 60 d songs were less stereotyped than normal adult songs (unpaired *t* test,  $p < 0.001$  for human scores;  $p < 0.024$  for overlap values). The syllables-only cross-correlation measures did not find significant differences between adult and juvenile song stereotypy, regardless of whether adult song was compared with ts cut or normal 60 d songs (unpaired *t* tests,  $p = 0.148$  and  $p = 0.062$ , respectively).

When we compared the mean  $d'_{\text{tsBOS-tutor}}$  values obtained from LMAN of each bird to the two sensitive measures of song stereotypy, no strong correlations were found (Fig. 17B,C) ( $r^2 = 0.147$  for human scores;  $r^2 = 0.019$  for motif-song overlap). These stereotypy measures also failed to predict the mean  $d'_{\text{tsBOS-tutor}}$  values obtained from X (data not shown) ( $r^2 = 0.027$  for human scores;  $r^2 = 0.012$  for motif-song overlap). When mean peak  $d'_{\text{tsBOS-tutor}}$  values of LMAN cells were compared with the human stereotypy scores, a weak positive correlation resulted ( $r^2 = 0.206$ ;  $p = 0.089$ ), suggesting that an increase in stereotypy was related to an increased neural preference for tsBOS. Otherwise, no strong correlations occurred using mean peak  $d'_{\text{tsBOS-tutor}}$  values in either LMAN or X (in LMAN,  $r^2 = 0.012$  for motif-song overlap; in X,  $r^2 = 0.014$  for human scores;  $r^2 = 0.017$  for motif-song overlap). In addition, comparing the individual  $d'_{\text{tsBOS-tutor}}$  values of each cell with these stereotypy measures did not improve these correlations (Table 3), nor did excluding data from the bird contributing the majority of unselective neurons (data not shown).

## DISCUSSION

This study addressed the relative contributions of BOS and tutor song to AF selectivity. NXII's transections caused juvenile birds to produce songs that were acoustically different from tutor song. Some neurons preferred the tsBOS over the tutor song, demonstrating that BOS experience can shape AF selectivity. Other neurons responded equally well to tsBOS and tutor song, despite their acoustic differences. Many of these neurons were also selective, and several methods of song analysis did not find residual similarities between tsBOS and tutor songs that could account for

gles) comparisons. In addition, mean overlap values (right ordinate) are plotted for song-song overlap (white squares) and motif-song overlap (black triangles) values. Error bars indicate SEM. **B**, Each bird's mean  $|d'_{\text{tsBOS-tutor}}|$  value of LMAN neurons is plotted against the syllable-only cross-correlation measures obtained from comparisons of tsBOS to tutor song. The thick black line is the least squares fit of the data, the vertical dashed line indicates the mean obtained for randomly matched songs, and error bars indicate SEM. The regression coefficient (slope) is  $-2.437 \pm 3.816$  ( $\pm 95\%$  confidence intervals). **C**, Each bird's mean  $|d'_{\text{tsBOS-tutor}}|$  value of LMAN neurons is plotted against the motif-song overlap values, which measure temporal similarity between tsBOS and tutor song. Conventions are as in **B**. The regression coefficient is  $0.714 \pm 0.595$ .

**Table 3. Correlations between  $d'$  values and song characteristics**

Similarity measure vs individual $ d'_{\text{tsBOS-tutor}} $ values	LMAN $r^2$ value		X $r^2$ value	
	Original $ d' $	Peak $ d' $	Original $ d' $	Peak $ d' $
Percent correctly matched	0.059	0.019	0.093	0.001
Spectral similarity score	0.007	0.002	0.009	0.007
Temporal similarity score	0.120	0.131	0.090	0.050
Syllables-only cross-correlation	0.063	0	0.007	0.085
Motif–song overlap	0.161	0.006	0	0.027
Stereotypy measure vs individual $d'_{\text{tsBOS-tutor}}$ values	Original $d'$	Peak $d'$	Original $d'$	Peak $d'$
Human stereotypy score	0.010	0.041	0.034	0.017
Motif-song overlap	0.007	0.002	0.001	0.031

The coefficients of determination ( $r^2$ ) result from the linear least squares fit of the data; each similarity or stereotypy measure is compared to the  $d'_{\text{tsBOS-tutor}}$  value of each cell recorded. Entries of 0 refer to  $r^2$  values  $<0.001$ .

these responses. These results strengthen the idea that both BOS and tutor song can contribute to the selectivity of single AF neurons.

### **NXIIts cut song shows evidence of poor temporal and spectral learning**

The effects of NXIIts transections on song in this study indicate the importance of respiratory and vocal muscle interactions during song development. The timing of syllable occurrence within song depends on the patterns of airflow through the syrinx, which are controlled by both syringeal and respiratory musculature (Hartley and Suthers, 1989; Vicario, 1991; Goller and Suthers, 1996; Suthers, 1997). The song system coordinates these muscle groups during song through RA, which projects both to motor neurons of syringeal muscles contained in NXIIts and to respiratory premotor nuclei such as RAM and PAM (Wild, 1993, 1997; Reinke and Wild, 1998). In adults, NXIIts transections transform syllables into harmonic stacks but maintain the timing of syllable occurrence within song (Simpson and Vicario, 1990; Williams and McKibben, 1992); thus, in NXIIts-transected adults, the respiratory pathway alone can maintain the temporal structure of song. In contrast, NXIIts transections made in juveniles resulted in songs that consisted mainly of harmonic stacks but that shared little spectral or temporal similarity with the tutor song. Thus, for these juveniles, the connection from RA to respiratory centers was not sufficient to produce timing similar to the tutor song. Because RA volume was unaffected by these transections, it seems unlikely that retrograde effects of NXIIts transection on RA resulted in improper signaling to the respiratory centers. Instead, syringeal control through NXIIts appears to participate in learning song timing.

Of the song analysis methods used here, the human matching task was the most sensitive to similarities between songs from different birds. For example, humans readily detected similarities between normal 60 d songs and tutor songs that the cross-correlation and overlap measures did not. The insensitivity of some of the automated methods was surprising; some cross-correlation analyses and overlap values measured the similarity apparent to humans between normal adult song and their tutor songs to be as low as that between randomly matched songs. These automated measures may be more suitable for quantifying similarity between songs from the same bird; indeed, in the stereotypy analysis, the motif–song overlap measure was as sensitive as human scoring. Despite the subjectivity inherent in human judgment of song qualities, these methods are currently

the most sensitive and remain the standard for song analysis (Eales, 1985; Williams, 1990; Scharff and Nottebohm, 1991; Nordeen and Nordeen, 1992). An important step in the birdsong field will be the development of algorithms capable of detecting and quantifying similarities between songs from different birds at least as well as humans.

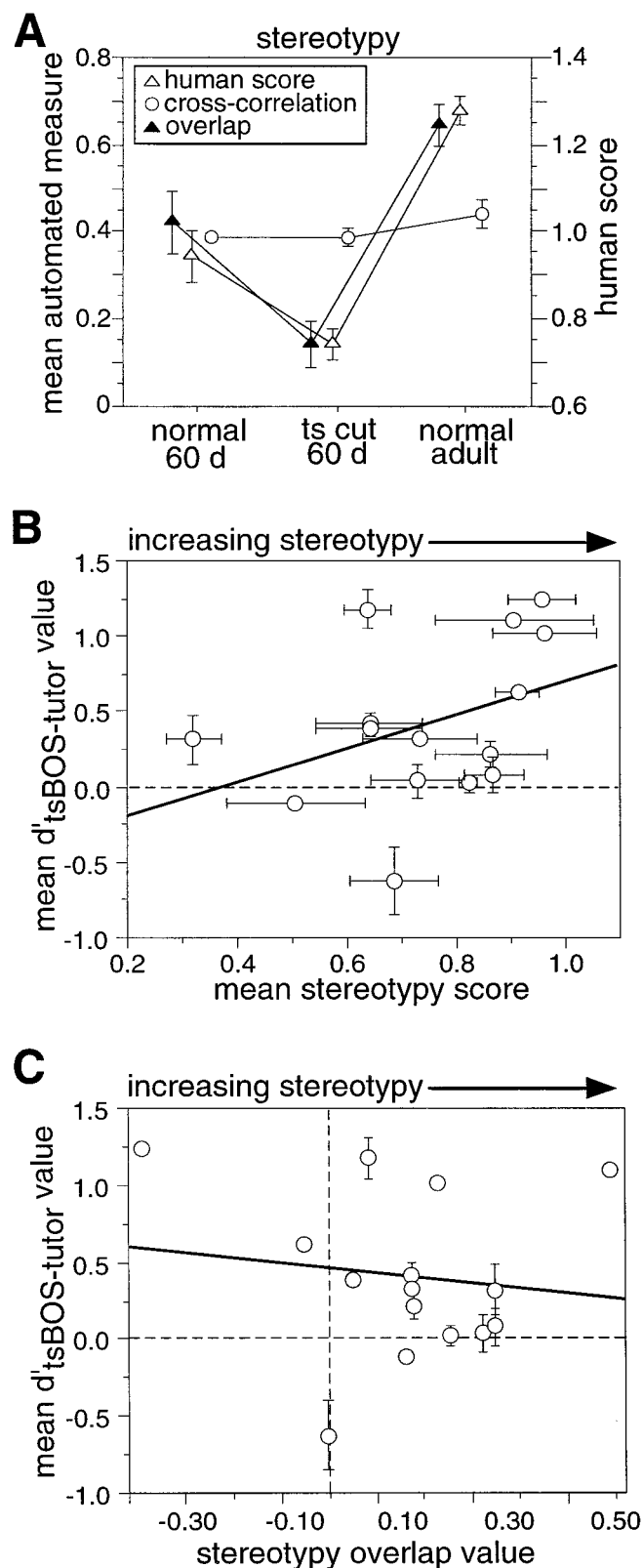
### **Contributions of two song experiences to AF selectivity**

Many AF neurons responded equally well to tsBOS and tutor song, despite the acoustic differences between these songs. These neurons were also song- and order-selective, which demonstrates that they were not indiscriminately responding to any stimulus. Thus, such neurons support the idea that both tsBOS and tutor song can shape the selectivity of single AF neurons. These neurons might be useful for comparisons of tsBOS and tutor song, which must occur behaviorally during song learning (Konishi, 1965; Price, 1979). The AF could be involved in calculating an error signal that measures the difference between tsBOS and the tutor song template; this error signal could then guide the modification of plastic song during learning.

Examples of single-neuron selectivity for two different stimuli have been found in other neural systems. Some neurons in the bat auditory cortex are tuned to two different durations, depending on whether neurons are responding to biosonar pulses or communication calls (Ohlemiller et al., 1996). Dual selectivity for more complex stimuli has been found in the inferior temporal visual cortex. There, neurons selectively respond to dissimilar fractal patterns that had been consecutively presented during training (Miyashita, 1988). This dual selectivity may reflect an association made between two stimuli consistently experienced together. Similarly, the selectivity for both tsBOS and tutor song could reflect experience of both songs during learning.

Other types of neurons that we observed also supported roles for tsBOS and tutor song experience in shaping AF selectivity. Neurons with strong preferences for the abnormal tsBOS over tutor song clearly indicate that, rather than reflecting poor tutor song copying, BOS experience itself can shape AF selectivity. The sensitivity of these neurons to tsBOS makes them well suited to provide information about the state of plastic song during sensorimotor learning, perhaps participating in the evaluation of BOS. Neurons that preferred tutor song over tsBOS, although rare, are also consistent with an influence of tutor song experience on AF selectivity. Such neurons could represent the stored memory of tutor song. Their rarity, however, combined with the





**Figure 17.** Song stereotypy analysis. *A*, Three measures of stereotypy are compared for different song groups. The mean syllables only cross-correlation measures (white circles) and motif–song overlap values (black triangles) are shown on the left ordinate, and the mean human stereotypy scores are shown on the right ordinate. Error bars indicate SEM and in one case is smaller than the symbol itself. *B*, The mean  $d'_{tsBOS-tutor}$  values of LMAN neurons from each ts cut bird are compared with the human

sizable population of neurons responding to both tsBOS and tutor, raises the possibility that tutor song information is primarily found in neurons that also respond to BOS rather than residing in neurons dedicated solely to encoding tutor song. Alternatively, the rarity of neurons with strong tutor song preferences might indicate that tutor song information is distributed sparsely or is encoded differently or elsewhere in the brain (Doupe and Solis, 1997).

Recordings from normal 60 d birds had already suggested a contribution of both BOS and tutor song experience to AF selectivity (Solis and Doupe, 1997). Thus, the tuning of these neurons to two stimuli may be a normal feature of learning in songbirds. As birds mature, tutor song responses may eventually be lost, leaving selectivity only for BOS. Consistent with this idea, the average response of LMAN neurons to tutor song was less than that to tsBOS, and in general, selectivity for tutor song over other stimuli was weaker than that for tsBOS. A transition in selectivity also occurs in the barn owl optic tectum, in which tuning to two different auditory cues precedes the establishment of selectivity for only one cue during the recalibration of auditory–visual maps (Brainard and Knudsen, 1995). Thus, raising ts cut birds to adulthood might reveal a strong preference for BOS over tutor song. Furthermore, the lower selectivity of ts cut birds relative to normal birds could reflect the delayed song development in ts cut birds, which was evident in their low song stereotypy. Again, ts cut birds raised to adulthood may show an increase in AF selectivity.

#### Measuring neural responses in LMAN and X

The differences in duration between tsBOS and tutor song stimuli in some experiments compelled us to evaluate different methods of measuring neural responses. The appropriate way to measure a neural response depends on how neurons downstream from LMAN and X respond to their inputs. If downstream neurons integrate afferent activity only for a short time, then the peak RS is more appropriate. However, if they integrate incoming signals for the duration of each song stimulus, then the original RS is appropriate. Longer integration times seem more suitable for analyzing these neurons; for a sample of AF neurons, the best discrimination between two stimuli occurred for window sizes >500 msec (our unpublished data). Other clues may also come from behavioral studies. If song duration has behavioral significance in zebra finches, then it is reasonable to measure responses over an entire song stimulus, whatever its duration. Some species respond differently to a song when its duration is varied by altering the number of repeated motifs (Kroodsma, 1976; Becker, 1982). Nevertheless, the peak RS and peak  $d'$  values gave results similar to the original measures of neural response, indicating the robust nature of the selectivity found in this study.

#### Song experience shaping HVC, the source of AF input

Neurons in HVC also become selective during song learning (Volman, 1993). Studies in both adult (Margoliash, 1986) and

scores of stereotypy. Because individual scores were normalized by the observer's mean score, a value of 0.3 refers to low stereotypy, and a value of 1.4 refers to high stereotypy. The thick black lines are least squares fit of the data, and error bars indicate SEM. The regression coefficient is  $1.133 \pm 1.64$  ( $\pm 95\%$  confidence intervals). *C*, The mean  $d'_{tsBOS-tutor}$  values of LMAN neurons from each ts cut bird are compared with the motif–song overlap measure of stereotypy. The regression coefficient is  $-0.328 \pm 1.42$ . Conventions are as in *B*.

juvenile (Volman, 1993) white-crowned sparrows have shown that HVC neurons are selective for BOS. In both studies a subset of birds had songs different from the tutor song, and their neurons showed strong preferences for BOS over tutor song. The weak tutor song responses in these birds might seem inconsistent with the significant tutor song responses found in the AF in this study, because HVC is the source of auditory inputs to the AF. It is possible, however, for AF neurons to derive their selectivity differently from their HVC inputs. HVC comprises two populations of neurons, one projecting to the AF, the other to RA (Sohrabji et al., 1989). X-projecting neurons could be shaped by both tutor song and BOS, whereas RA-projecting neurons could be shaped by BOS alone. Alternatively, tutor song responses may have been overlooked in these HVC studies. First, both used multiunit recordings, which could miss other kinds of selective neurons if they are few. Second, the strong BOS preferences of HVC neurons in this subset of birds could reflect poor copying of the tutor song model. In one study, the birds with songs unlike the tutor song had been tutored with an abnormal tutor song model, unlike natural white-crowned sparrow song (Margoliash, 1986). As juveniles, the birds may have had difficulty memorizing such an abnormal song, potentially leading to storage of something other than the tutor song model as the template. In contrast, the NXIIIs transections used in this study reduced the similarity between BOS and tutor song by manipulating the juvenile song itself rather than the tutor song model. Thus, the stronger tutor song responses found here may reflect greater agreement between what was presented as the tutor song and what was stored as the template.

### Range of tsBOS versus tutor song preferences among AF neurons

It remains intriguing that there is a variety of tsBOS versus tutor song preferences in ts cut birds, ranging mainly from neurons that strongly prefer tsBOS over tutor song to neurons that respond equally well to both songs. Although not found here, it is theoretically possible that a different method of song analysis, perhaps one that exclusively measures those elements salient to the birds themselves, could find residual similarities between tsBOS and tutor song that were responsible for the equivalent responses to these songs. Alternatively, the state-dependence of auditory responses in neurons of the song system might have influenced the preferences recorded (Hessler and Doupe, 1997; Dave et al., 1998; Schmidt and Konishi, 1998). Although measures of anesthesia depth did not correlate with  $d'_{\text{tsBOS-tutor}}$  values, it remains possible that subtle differences in arousal state may have differentially emphasized BOS or tutor song responses in neurons actually capable of responding equally well to both stimuli. A third possibility is that the  $d'_{\text{tsBOS-tutor}}$  values from a bird could reflect its maturity. One measure of song similarity weakly supported this idea: as motif-song overlap values increased (which presumably happens as a bird matures), neural preference for BOS also increased. In contrast, neither the age nor song stereotypy of a bird predicted the  $d'_{\text{tsBOS-tutor}}$  values obtained from its neurons.

Finally, differences in the accuracy of learning from a tutor bird may have contributed to the range of  $d'_{\text{tsBOS-tutor}}$  values. In another system, differences between individual animals in neural processing of a task resulted from slightly different behavioral strategies for solving that task (Seidemann et al., 1998). In our study it is difficult to know whether the tutor song presented during the experiment was similar to what the bird had memo-

rized. To favor similarity between the tutor song and the template, juveniles were housed in conditions that maximize copying from the tutor bird (see Materials and Methods). Also, the great degree of similarity between songs from normal adults and their tutor songs shows that there is a high incidence of tutor song copying in the colony used in this study. Nonetheless, because there is currently no direct way to assay what has been stored as the template, it remains possible that all neurons are actually equally shaped by tsBOS and tutor song; neurons with strong preferences for BOS may have been presented with a tutor song that did not match the template.

### Possible mechanisms underlying dual selectivity

Although the site of plasticity initially giving rise to selectivity is not addressed in this study, possible mechanisms for mediating selectivity to two different songs exist within the AF. Both LMAN and X have complex intrinsic circuitry that could form multiple, differentiated populations of synapses onto single neurons, enabling them to process tsBOS and tutor song separately (Sohrabji et al., 1993; Vates and Nottebohm, 1995; Farries and Perkel, 1998; Luo and Perkel, 1999). Indeed, synapses onto single LMAN neurons differ markedly in their pharmacological and temporal properties (Boettiger and Doupe, 1998). Moreover, in addition to their auditory responses, LMAN and X neurons are strongly active during singing (Dave et al., 1997; Hessler and Doupe, 1999). This raises the possibility that tuning to different stimuli occurs in different behavioral states, such as singing or listening. The motor-related activity in AF neurons, which could represent efference copy signals (Troyer et al., 1996), might actually contribute to the auditory tuning of these cells. Finally, the acoustical basis of the dual responses to tsBOS and tutor song could be determined in experiments that systematically decompose tsBOS and tutor song stimuli, thus revealing the song components essential for a neural response. Such experiments would contribute to our understanding of a single neuron's selectivity for two different stimuli.

### REFERENCES

- Basham ME, Nordeen EJ, Nordeen KW (1996) Blockade of NMDA receptors in the anterior forebrain impairs memory acquisition in the zebra finch (*Poephila guttata*). *Neurobiol Learn Mem* 66:295–304.
- Becker P (1982) The coding of species-specific characteristics in bird sounds. In: *Acoustic communication in birds* (Kroodsma D, Miller E, eds), pp 213–252. New York: Academic.
- Boettiger C, Doupe A (1998) Intrinsic and thalamic excitatory inputs onto songbird LMAN neurons differ in their pharmacological and temporal properties. *J Neurophysiol* 79:2615–2628.
- Bottjer SW, Miesner EA, Arnold AP (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224:901–903.
- Brainard MS, Knudsen EI (1995) Dynamics of visually guided auditory plasticity in the optic tectum of the barn owl. *J Neurophysiol* 73:595–614.
- Dave AS, Yu AC, Margoliash D (1997) Reassessing auditory responses in the birdsong system. *Soc Neurosci Abstr* 23:245.
- Dave AS, Yu AC, Margoliash D (1998) Behavioral state modulation of auditory activity in a vocal motor system. *Science* 282:2250–2254.
- Doupe A, Solis M (1997) Song- and order-selective neurons develop in the songbird anterior forebrain. *J Neurobiol* 33:694–709.
- 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.
- Eales LA (1985) Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Anim Behav* 33:1293–1300.

- Eales LA (1987) Song learning in female-raised zebra finches: another look at the sensitive phase. *Anim Behav* 35:1356–1365.
- Eales LA (1989) The influences of visual and vocal interaction on song learning in zebra finches. *Anim Behav* 37:507–520.
- Esser K, Condon C, Suga N, Kanwal J (1997) Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat *Pteronotus parnellii*. *Proc Natl Acad Sci USA* 88:11339–11343.
- Farries M, Perkel D (1998) Electrophysiological similarities between Area X neurons in the zebra finch and neurons of the mammalian striatum. *Soc Neurosci Abstr* 24:191.
- Goller F, Suthers R (1996) Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J Neurophysiol* 75:867–876.
- Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.
- Gurney ME (1981) Hormonal control of cell form and number in the zebra finch song system. *J Neurosci* 1:658–673.
- Hartley R, Suthers R (1989) Airflow and pressure during canary song: direct evidence for minibreaths. *J Comp Physiol [A]* 165:15–26.
- Hessler N, Doupe A (1997) Singing-related neural activity in anterior forebrain nuclei of adult zebra finch. *Soc Neurosci Abstr* 23:245.
- Hessler N, Doupe A (1999) Social context modulates singing-related neural activity in the songbird forebrain. *Nat Neurosci* 2:209–211.
- Immelmann K (1969) Song development in the zebra finch and other estrildid finches. In: *Bird vocalizations* (Hinde RA, ed), pp 61–74. London: Cambridge UP.
- Konishi M (1965) The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z Tierpsychol* 22:770–783.
- Kroodsmas D (1976) Reproductive development in a female songbird: differential stimulation by quality of male song. *Science* 192:574–575.
- Lewicki MS (1994) Bayesian modelling and classification of neural signals. *Neural Comput* 6:1005–1030.
- Luo M, Perkel DJ (1999) A long-range GABAergic projection in a circuit essential for vocal learning. *J Comp Neurol* 403:68–84.
- 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.
- Miyashita Y (1988) Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335:817–820.
- Mooney R (1992) Synaptic basis for developmental plasticity in a bird-song nucleus. *J Neurosci* 12:2464–2477.
- 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.
- Ohlemiller K, Kanwal J, Suga N (1996) Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. *NeuroReport* 7:1749–1755.
- Price PH (1979) Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol* 93:260–277.
- Rauschecker J, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.
- Reinke H, Wild J (1998) Identification and connections of inspiratory premotor neurons in songbirds and budgerigar. *J Comp Neurol* 391:147–163.
- Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits after lesions of the various parts of the zebra finch song system: implications for vocal learning. *J Neurosci* 11:2896–2913.
- Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. *Nat Neurosci* 1:513–518.
- Seidemann E, Zohary E, Newsome W (1998) Temporal gating of neural signals during performance of a visual discrimination task. *Nature* 394:72–75.
- Simpson HB, Vicario DS (1990) Brain pathways for learned and unlearned vocalizations differ in zebra finches. *J Neurosci* 10:1541–1556.
- Sohrabji F, Nordeen KW, Nordeen EJ (1989) Projections of androgen-accumulating neurons in a nucleus controlling avian song. *Brain Res* 488:253–259.
- Sohrabji F, Nordeen EJ, Nordeen KW (1990) Selective impairment of song learning after lesions of a forebrain nucleus in the juvenile zebra finch. *Behav Neural Biol* 53:51–63.
- Sohrabji F, Nordeen EJ, Nordeen KW (1993) Characterization of neurons born and incorporated into a vocal control nucleus during avian song learning. *Brain Res* 620:335–338.
- Solis M, Doupe A (1997) Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. *J Neurosci* 17:6447–6462.
- Sossinka R, Böhner J (1980) Song types in the zebra finch *Poephila guttata castanotis*. *Z Tierpsychol* 53:123–132.
- Suga N, O'Neill W, Manabe T (1978) Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat. *Science* 200:778–781.
- Suthers R (1997) Peripheral control and lateralization of birdsong. *J Neurobiol* 33:632–652.
- Theunissen F, Doupe A (1998) Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVC of male zebra finches. *J Neurosci* 18:3786–3802.
- Troyer TW, Doupe AJ, Miller KD (1996) An associational hypothesis for sensorimotor learning of birdsong. In: *Computational neuroscience: trends in research 1995* (Bower JM, ed), pp 409–414. New York: Academic.
- Vates G, Nottebohm F (1995) Feedback circuitry within a song-learning pathway. *Proc Natl Acad Sci USA* 92:5139–5143.
- Vicario DS (1991) Contributions of syringeal muscles to respiration and vocalization in the zebra finch. *J Neurobiol* 22:63–73.
- Volman SF (1993) Development of neural selectivity for birdsong during vocal learning. *J Neurosci* 13:4737–4747.
- Volman S (1996) Quantitative assessment of song-selectivity in the zebra finch “high vocal center.” *J Comp Physiol [A]* 178:849–862.
- Wang X, Merzenich M, Beitel R, Schreiner C (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74:2685–2706.
- Wild J (1997) Neural pathways for the control of birdsong production. *J Neurobiol* 33:653–670.
- Wild JM (1993) Descending projections of the songbird nucleus robustus archistriatalis. *J Comp Neurol* 338:225–241.
- Williams H, McKibben J (1992) Changes in stereotyped central motor patterns controlling vocalization are induced by peripheral nerve injury. *Behav Neural Biol* 57:67–78.
- Williams HW (1990) Models for song learning in the zebra finch: fathers or others? *Anim Behav* 39:745–757.