Song Tutoring in Pre-singing Zebra Finch Juveniles Biases a Small Population of Higher-Order Song Selective Neurons Towards the Tutor Song

Patrice Adret
C. Daniel Meliza
Daniel Margoliash

Author contributions
P. Adret and D. Margoliash conceived and designed the two studies; P. Adret carried out the experiments; P. Adret and C. D. Meliza analyzed the data; P. Adret, C. D. Meliza and D. Margoliash wrote the paper.

Running Head
Auditory selectivity in pre-singing zebra finches

Affiliation
Department of Organismal Biology and Anatomy
University of Chicago

Corresponding author
Daniel Margoliash
Department of Organismal Biology and Anatomy
1027 E. 57th St., Chicago IL 60637
Tel: (773) 702-3224
Fax: (773) 702-0037 dan@bigbird.uchicago.edu
Adret, Patrice, C. Daniel Meliza and Daniel Margoliash. We explored physiological changes correlated with song tutoring by recording the responses of caudal nidopallium neurons of zebra finches aged P21–P24 (days post hatching) to a broad spectrum of natural and synthetic stimuli. Those birds raised with their fathers tended to show behavioral evidence of song memorization but not of singing; thus auditory responses were not confounded by the birds’ own vocalizations. In Study 1, 37/158 neurons (23%) in 17/22 tutored and untutored birds were selective for one only out of ten stimuli comprising broadband signals, early juvenile songs and calls, female calls, and adult songs. Approximately 30% of the selective neurons (12/37 neurons in 9 birds) were selective for adult conspecific songs. All these were found in the song system nuclei HVC and paraHVC. Of 122 neurons (17 birds) in tutored birds, all of the conspecific song selective neurons (8 neurons in 6 birds) were selective for the adult tutor song; none were selective for unfamiliar song. In Study 2 with a different sampling strategy, we found that 11/12 song selective neurons in 6/7 birds preferred the tutor song; none preferred unfamiliar or familiar conspecific songs. Most of these neurons were found in NCL below HVC. Thus by the time a bird begins to sing, there are small numbers of tutor song selective neurons distributed in several forebrain regions. We hypothesize that a small population of higher-order auditory neurons is innately selective for complex features of behaviorally-relevant stimuli, and these responses are modified by specific perceptual/social experience during development.

Keywords: birdsong learning, song system, template theory, vocal development
INTRODUCTION

Developmental learning is characterized by restricted periods of heightened plasticity during which the structure of sensory experience has long-term effects on neuronal connectivity and the representation of sensory stimuli. For example, in the midbrain of young barn owls, the auditory map of space is adjusted to the map of visual space when auditory and visual cues are associated by sound localization experience (Knudsen and Knudsen 1989), and in primary visual cortex of juvenile macaque monkeys, the ocular dominance map is shaped by the relative activity from each eye (Hubel et al. 1977). It is commonly anticipated that developmental, experience-dependent plasticity also occurs in higher-order sensory and sensorimotor structures. This has been more difficult to characterize, however, in part because the topography of these areas is coarser (Romanski and Averbeck 2009) or developmentally delayed (Iyengar et al. 1999), or the topography is not known (Johnson and Bottjer 1995; Vates and Nottebohm 1995) and a more complex developmental pattern of neural selectivity dominates (Nick and Konishi 2005a,b).

Beyond topography, in higher-order areas neurons can respond selectively to discrete objects from an animal’s perceptual experience (Horn et al. 2001; Gentner and Margoliash 2003), often showing nonlinear responses to combinations of the objects’ features (Suga et al. 1978; Margoliash 1983; Rauschecker et al. 1995; Romanski and Goldman-Rakic 2002). Here we explore the emergence of high-level selectivity during development and the effects of auditory and social experience.

Early in development, juvenile songbirds form sensory memories of adult song that will guide the emergence of their own singing behavior (for recent reviews see Adret 2008; Mooney 2009;
Margoliash and Schmidt 2010). These memories are acquired from a conspecific tutor, often well before the bird makes any attempt at singing, though this claim rests primarily on behavioral evidence of early perceptual biases (Marler 1997) and the instructive effects of early tutoring on the bird’s adult song (Baptista and Petrinovich 1984; Nelson 1997; Beecher and Burt 2004). A few studies have attempted to characterize the response properties of single neurons in higher-order auditory pathways and their feed-forward targets in the vocomotor song system of young birds (Volman 1993; Whaling et al. 1997; Stripling et al. 2001), but these did not directly test if the birds had prior singing experience. If vocal learning affects auditory perceptual representations (Nick and Konishi 2005a; Keller and Hahnloser 2009), this caveat could prove important.

In the present study, a small fraction of neurons recorded in pre-singing zebra finches showed evidence of selectivity for conspecific song over other natural and synthetic stimuli. Every one of these neurons preferred the song of the bird’s tutor over songs from unfamiliar and familiar birds. We found such memory traces in three of the four regions we examined but in each region only in a small percentage of neurons. We suggest that the influence of early development on the responses of higher-order forebrain neurons is pervasive – not restricted to a single forebrain region within a sensory modality, but limited to a restricted population of neurons within each region. For early birdsong development, we hypothesize that such processes contribute to a representation of the tutor song memory that is distributed across multiple nuclei.
MATERIALS AND METHODS

All procedures were approved by the University of Chicago Institutional Animal Care and Use Committee.

Subjects

All of the zebra finches (*Taeniopygia guttata*) used in these experiments were males hatched in our breeding colony \((n = 36 \text{ birds})\), 25 clutches; mean ± SD = 4.0 ± 1.5 birds/clutch). We used only those clutches in which young males could be readily identified by plumage appearance: the presence at fledging (≈P19) of black feathers on the chest (badge). Breeding pairs were kept visually isolated from each other in cages arranged as a wall unit. Birds were either reared normally by both parents (hereafter tutored birds, \(n = 31\)), or they were isolated from male song at P1–P4 by removing the father from the cage and transferring the nestlings with their mother to a large, communal sound booth that held up to four cages in visual isolation (hereafter untutored birds, \(n = 5 \text{ birds, 1 per clutch}\)). The tutored birds were exposed to a complex auditory scene dominated by calls from both sexes, crystallized songs of adult males and varying amounts of begging calls, subsong and plastic song from other juveniles. Subjects could only interact physically with siblings, their mother, and one adult male, their father, who served as a song tutor. Young males raised in these social conditions copy their tutor’s rather than their neighbors’ songs (Slater et al. 1988). In contrast, the untutored birds could hear the calls of adult females (who don’t sing), other nestlings’ begging calls, and perhaps occasional subsong from older fledglings (>P25), but were not exposed to adult song at any point after P4.
Four of the tutored birds (2 birds each from 2 clutches) were used to assess the degree of song learning that took place prior to the age at which we conducted electrophysiological studies. These birds were acoustically isolated at P23–P25 with their mothers until P30 and thereafter lived independently in isolation until adulthood. Sonograms from 20 exemplars of undirected song from each bird, collected on the last day of recording (P95–P110), were visually inspected to calculate three measures of vocal learning. We assessed the proportion of song material imitated from the father; the proportion of copied syllables within each tutee’s song; and how well the sequence of copied syllables matched the tutor song syntax. Complex syllables that included two or more notes in the father’s song were often partially copied on a note basis; these syllables were scored by the proportion of notes imitated. For instance, imitation of a three-note syllable could potentially obtain a ratio ranging from 1/3 to 3/3. Song imitation was calculated by averaging the scores over the total number of syllables in the song phrase. Although young zebra finches preferentially learn from tutors with whom they can interact, we also compared each tutee’s song with those of two adult males located next to or above the subject’s cage.

The remaining birds (tutored: \( n = 27 \); untutored: \( n = 5 \)) were used to measure responses of single units to auditory stimulation. These experiments were conducted at P21–P24, at which age the dorsal skull of zebra finches is minimally calcified, requiring modified surgical procedures to achieve secure attachment of a metal pin to the skull. On days of experiments, each fledgling was weighed (mean ± SD = 12.0 ± 0.8 g, \( n = 26 \)) and anesthetized with an intramuscular injection of
90–110 µl of 20% urethane, administered in 3 aliquots at ~20 min intervals. After application of
topical lidocaine, the scalp was incised and retracted and the fascia over the dorsocaudal skull
was retracted with a blunt surgical instrument. A small metal pin was rigidly affixed to the soft,
non-sutured and translucid skull. This delicate operation consisted of three steps. First, we made
a small craniotomy on each side of the parietal plates, approximately 3.5 mm lateral and 3 mm
caudal to the midsagittal sinus. To access the skull at these lateral coordinates, the neck muscles
were partially retracted from the supra-occipital plates and a small fenestra was opened. The
exposed cavity, rich in trabecula cranii, was filled with dental acrylic (Grip cement, Division
Dentsply, Int. Inc. Milford), thus providing anchors for the metal pin. Next, the forebrain song
system nuclei HVC (the acronym is used as a proper name) and paraHVC were accessed on both
sides by opening a small fenestra in the skull (~1 mm square) centered 300 µm caudal and 2 mm
lateral to the mid-sagittal sinus, immediately caudal to the non-fused suture separating the frontal
and parietal plates. The exposed brain was protected from dehydration by silicone fluid (Dow
Corning). In the last step, the metal pin was slowly lowered onto the skull over the cerebellum
and embedded in dental acrylic cement. Hardening of the cement was then accelerated with
cyanoacrylate glue.

Single unit recordings from both hemispheres were then conducted using single solder-glass
coated Pt/Ir in-house electrodes (impedance = 0.7–1.5 MΩ). At the end of an experiment, birds
were perfused transcardially, the sex was confirmed (testes were observed in all birds with a
badge of black feathers) and each brain was stored in 4% paraformaldehyde. Sagittal brain
sections (40 µm thick) were collected in a cryostat and stained with cresyl violet for later
anatomical reconstruction of electrode penetrations.
Two different sets of auditory stimuli were used to probe the neural selectivity of the units we recorded. In Study 1, each unit was presented with 20 repetitions of ten stimuli drawn from nine different classes: early subsong from zebra finches (Juv), subsong from white-backed munia, *Lonchura striata* (Mun), subsong from the domesticated form of the white-backed munia, the Bengalese finch (BF), song from an adult zebra finch male (Con), zebra finch begging calls (Beg), zebra finch female long calls (FLC), white noise with the amplitude envelope of subsong (JuvAM), white noise with the envelope of unfamiliar conspecific song (ConAM), and unmodulated white noise (WN). To avoid pseudoreplication (in this context, the use of the same exemplar in multiple tests of categorical differences; see Kroodsma 1989), exemplars were drawn randomly for each neuron from the following classes: Juv (n = 9 exemplars), Mun (n = 2), BF (n = 2), and JuvAM (n = 9). Only one exemplar of WN was used. Each neuron was presented with two different conspecific songs. For tutored birds (n = 17), one song was randomly chosen from a male the bird had not heard before the experiment (Unf; n = 11 exemplars), and the other was the song of the bird’s tutor (Tut; n = 12). Exemplars of unfamiliar conspecific songs were drawn at random from a sample of 22 songs (mean ± SD = 5.3 ± 2.3 songs/bird; range: 1–10). For untutored birds (n = 5), both conspecific songs were unfamiliar. The ConAM stimulus was either derived from the bird’s tutor (n = 63 units, tutored birds only) or from an unfamiliar male (n = 97 units, both rearing protocols). For 19 birds (tutored: n = 15; untutored: n = 4), the exemplars of Beg and FLC were recorded from the test bird and its mother on days prior to the experiment and were therefore not randomly assigned; for the remaining 3 birds, exemplars from another subject were used.
In Study 2 (n = 10 birds, all tutored), each neuron was presented with one exemplar from JuvAM, Juv, and WN (same number of exemplars from each class as above), as well as three exemplars of adult conspecific song: the tutor song (n = 6), an unfamiliar song (n = 12), and a song from a familiar male (Fam; n = 5). The familiar songs were from adult males housed in cages adjacent to the test birds in the aviary. Thus, the familiar songs were a prominent part of the acoustic experience of the young birds, although we do not have independent confirmation that these songs were categorically distinct to the test birds from the songs of unfamiliar adult conspecifics. All the neurons in each bird were tested with the same Fam and Tut exemplars, which were determined by the bird’s rearing; exemplars of unfamiliar conspecific songs were drawn at random for each neuron (mean ± SD = 2.6 ± 2.1 songs/bird; range = 1–7). At each recording site, 5 repetitions of each stimulus were presented to make an online assessment of the selectivity. If the site appeared to be selective for adult song (19/128, 15%; n = 7 birds), an additional 15 repeats of the stimuli were presented.

The stimuli had a duration of 2.5 ± 0.3 s (mean ± SD, n = 95); amplitude was normalized to approximately 70 dB RMS. Including both studies, there were 18 different unfamiliar songs, 16 different tutor songs, and 5 different familiar songs, two of which were tutor songs for birds in Study 1. An analysis (Sound Analysis Pro; http://www.soundanalysispro.com) of three acoustic parameters (Wiener entropy, frequency modulation and goodness of pitch) for all the stimuli revealed four non-overlapping clusters: noise stimuli, female calls, subsongs and juvenile begging calls, and adult songs. Within adult songs, tutor, unfamiliar, and familiar songs were intermingled, indicating similar global, low-level acoustical properties (data not shown).

Spike analysis
Spike times were extracted offline with a spike-sorting algorithm based on principal components. An initial phase of automated clustering (Klustakwik) was followed by manual refinement (Klusters) (Hazan et al. 2006). A unit was considered to be well isolated only if no more than one or 0.5% (whichever was greater) of the interspike intervals were <1 ms. Multiple well-isolated units were often recovered from a single site. In Study 1, a total of 158 units with excitatory responses to at least one auditory stimulus were recorded from 109 sites (1–4 units per site). In Study 2, 24 auditory units were recovered from 19 sites (1–4 units per site).

Subsequent analysis followed conventional methods. Neuronal response strength \( (r) \) was defined as \( r_{i,k} = n_{i,k}/T_i - s_{i,k}/T_0 \) where \( n_{i,k} \) is the spike count in response to the \( k \)-th presentation of stimulus \( i \), \( T_i \) is the duration of stimulus \( i \), and \( s_{i,k} \) is the spike count measured during a prestimulus baseline period (\( T_0 = 1 \) s for all trials). A unit was considered “auditory” if \( r_i \) (i.e., the mean response strength across presentations) for any stimulus was significantly greater than zero (one-sample \( t \)-test; \( P < 0.05 \)).

Single unit selectivity

Each unit was tested to determine if it responded more to one of the stimulus classes than to any other. We used d-prime (\( d' \)), a measure of how discriminable the responses to two stimuli are. Given stimuli \( A \) and \( B \), \( d' = 2(r_A - r_B)/\sqrt{\sigma_A^2 + \sigma_B^2} \), where \( r_A \) and \( r_B \) are the mean responses to \( A \) and \( B \), and \( \sigma_A^2 \) and \( \sigma_B^2 \) the corresponding variance (both statistics calculated across trials). Because \( d' \) takes into account both the strength and the variance of the neuronal response to a given stimulus, it has been used extensively to assess the selectivity of songbird auditory
forebrain neurons (Theunissen et al. 2004). Note the similarity between $d'$ and the $t$ statistic:

$$t = \sqrt{n}(r_A - r_B)/\sqrt{\sigma_A^2 + \sigma_B^2},$$

where $n$ is the number of samples for each stimulus type; thus, $d' = 2t/\sqrt{n}$.

For each auditory unit, we calculated $d'$ between the stimulus with the strongest average response and each of the other stimuli. A unit was considered selective if $d' > 0.77$ for all these comparisons. For 20 repetitions, this threshold corresponds to a $t$ statistic of 1.72 ($P < 0.05$ for a one-tailed $t$-test). This is a more stringent threshold than is usually applied (0.5 is common). No correction for multiple comparisons was necessary because the tests were explicitly linked and the null hypothesis was that any one of the means is equal to the largest mean. For example, consider the case of two tests, each with a type-I error threshold of $\alpha$. The probability of a false positive in either test is $2\alpha$, but the probability of a false positive in both tests is $\alpha^2$. In our analysis, the sequential $t$-tests were not independent, and although the probability of a false positive was greater than $\alpha^2$, it was still bound to be less than or equal to $\alpha$. Note also that in contrast to the normal logic of a multiple comparisons correction, in our case the null hypothesis could be falsified by a single pair-wise comparison.

Population selectivity

We also measured selectivity at the population level using a principal components-based analysis. Response strength was first standardized across cells and across animals by applying a $z$ transformation to each unit (i.e., subtracting the mean and dividing by the standard deviation). This normalization emphasizes the relative differences in each unit’s responses to the test.
stimuli. Following normalization, the responses of the population were arranged into a matrix $R$ with columns corresponding to stimulus classes and rows corresponding to units. Singular value decomposition was used to identify the principal components (PCs) and the projections of the neurons along each PC. We used the first two PCs, which together accounted for 44% of the variance in $R$. Because PCA maximizes the variance of the projections into this limited subspace, the values (or loadings) associated with each stimulus in the PCs are informative about the variance-covariance structure of $R$. If a stimulus has a high loading in the first PC, it indicates that the variance among neurons for that stimulus is relatively large. Furthermore, if two stimuli have similar loadings, it indicates that responses to those two stimuli tended to covary across neurons. This feature of PCA makes it useful to plot the loadings and the projections in the same “biplot” (Gabriel 1971) as we have done in Figure 4; in such plots, the length of the loading vectors approximates the variance associated with each stimulus, the dot product of two vectors approximates the covariance, and the cosine of the angle between vectors approximates the correlation. We confirmed the statistical significance of any apparent correlations between stimuli using Pearson product-moment tests.

A potential issue with the $z$-score normalization is that the values are scaled by the relative differences between responses to the different stimuli, and neurons with low selectivity (small $\sigma_{Ri}$) may be overrepresented. Therefore, the PCA analysis was repeated using a different normalization scheme that retains differences in absolute firing rate, $z_i = r_i \sqrt{K} / \sigma_{i,k} - \bar{z}$, where $K$ is the number of trials (20), $\sigma_{i,k}$ is the standard deviation of $r_i$ across trials, and $\bar{z}$ is the mean of $z_i$ across stimuli. Note that this is a mean-corrected version of the commonly-used “$z$-score” (Theunissen et al. 2004). Results using this normalization were essentially the same.
RESULTS

Juvenile zebra finches form sensory memories of song at an early age

Even with limited, early exposure to their fathers’ songs, adult zebra finches showed evidence of incorporating syllables from these songs into their own adult songs. Four birds that received early tutoring (through P23–P25) and then were raised to adulthood in isolation copied 45.5 ± 28.7% (mean ± SD) of the 7 or 8 syllable types in the tutor songs. Siblings in each clutch were exposed to the same tutor song; birds from one clutch copied 34% and 86% of their father’s song (Fig. 1A), and birds in the other clutch copied 19% and 43% of their father’s song. The copied syllables represented between 37% and 98% of the total number of syllables in the most common versions of the final, adult songs. Importantly, in one pair of siblings, the copied syllables were sung with the same sequence as the tutor song (Fig. 1A). By comparison, the degree of song similarity with the songs of two neighbors (5 and 7 syllables) was considerably less (mean ± SD = 3.6 ± 4.4%), with at most only one syllable in each tutee’s song that resembled a syllable from a neighbor. Two additional observers participating in a blind test ranked the tutor song as more similar to the song of each tutored bird’s song than to the songs of two neighbors.

Collectively, this is a strong indication that the birds learned even from limited exposure to song early in life, and acquired their song from their father, not from their neighbors. Overall, the copy quality of individual syllables was lower than what juvenile zebra finches with exposure to song throughout development often achieve, reminiscent of “improvised” syllable copying known for other species. Since the formation of early sensory memories was assessed – both in youth and adulthood – by singing (motor performance), these data reflect a lower bound on what the birds
had memorized.

Study 1: auditory responses in tutored and untutored juveniles

To assess whether this learning is reflected in early auditory responses, we recorded from single units in 22 juvenile birds (P12–P24) divided into two treatment groups: 17 were raised normally (tutored) and 5 were denied access to adult song soon after hatching (untutored). We recorded from 109 auditory sites in the song system nucleus HVC and surrounding nidopallial regions paraHVC, HVCshelf, and the caudal lateral nidopallium (NCL), achieving single-unit isolation from 158 neurons, 1–4 at each site (mean ± SD = 1.46 ± 0.63). These are among the earliest ages at which single neurons have been recorded from the songbird brain (Stripling et al. 2001). Each neuron was presented with a broad battery of 10 stimulus classes (see Materials and Methods).

The temporal patterns of auditory responses varied across neurons; some neurons exhibited a sustained increase in firing rate over the entire duration of the stimulus whereas other cells displayed sparse spiking activity apparently locked to specific sound features. Other cells gave transient excitatory responses to stimulation or showed relative inhibition interrupted by intermittent bursts of excitatory discharges. There were no qualitatively apparent differences in the response patterns of neurons in tutored and untutored birds.

For many neurons, some stimuli were more effective than others. A large proportion of neurons (79%, 125/158) showed a statistically significant difference in their average rate of response to the stimulus classes (one-way ANOVA at $P < 0.05$), which is unsurprising given how varied the stimuli were. There was no effect from tutoring (i.e., comparing tutored and untutored birds) in the proportion of neurons that passed this test ($\chi^2 = 0.44, df = 1, P = 0.49$). A much smaller
proportion of neurons passed a more stringent test of selectivity that required that the neuron respond more strongly to one of the stimuli than to any of the other nine, with $d' > 0.77$ (equivalently, $P < 0.05$ for all one-tailed $t$ tests comparing the strongest response to the other stimuli). Out of all the neurons, 23% (37/158) distributed across three brain areas (HVC, paraHVC and HVCshelf) were deemed “selective” because they responded more to one stimulus than all others. There was no significant effect of tutoring on the proportion of neurons that were selective (tutored = 28/122; untutored = 9/36; $\chi^2 = 0.07$, $df = 1$, $P = 0.80$).

The distribution of selectivity among the different stimulus types was not uniform (Fig. 2). Most of the selective neurons preferred either WN (32%, 12/37) or one of the conspecific songs (including Tut) (12/37). There was no significant difference in the distribution of selectivity between tutored and untutored birds ($\chi^2 = 3.05$, $df = 6$, $P = 0.80$; Fig. 2A,B). For untutored birds, 4/9 selective neurons preferred one of the conspecific songs (3 birds; 1–2 cells/bird). However, for tutored birds, all 8 conspecific song selective (CSS) neurons preferred the tutor song (6 birds; 1–2 cells/bird; Table 1). This observation was unlikely to have occurred through sampling error. If CSS neurons were selective to general features of conspecific songs, there would be an equal probability that a CSS neuron would respond most strongly to either Unf or Tut, but this hypothesis is inconsistent with the data ($\chi^2 = 8$, $df = 1$, $P = 0.005$). Each neuron was tested with a different unfamiliar song, but all the CSS neurons preferred the tutor song. Given that we found no differences in the overall distribution of selectivity between tutored and untutored birds, it is particularly striking that all the CSS neurons in tutored birds preferred Tut over Unf. We term these neurons tutor song selective (TSS).

Our sample sizes were small because these were technically difficult experiments in young
anesthetized birds, and we adopted a strict criterion for selectivity. The distribution of selectivity among stimuli, however, was not dependent on the criterion used to decide whether a neuron was selective or not. We tested a range of more and less stringent criteria ($d'$ thresholds from 0 to 1.34), and found that the proportion of TSS neurons was significantly above the upper bound expected by chance, and the proportion of Unf-selective neurons was well below the lower bound for all values of $d'$ (except $d' = 1.12$, where the proportion of TSS neurons was equal to the upper bound; chance distributions generated by bootstrap sampling). In this regard, note that the response strength to the tutor song was on average 3.75 times stronger than the response to the second strongest stimulus (Table 1; $n = 17$ TSS cells where the second strongest response was above baseline, i.e. non-negative).

Apart from the fact that all the CSS neurons in tutored birds preferred the tutor song, the responses of CSS neurons in tutored and untutored birds were qualitatively similar (Fig. 3). Some neurons showed transient responses and others showed sustained responses, and typically responses were locked to specific acoustic features of the tutor song. We note that neurons in untutored birds were tested with two different conspecific songs, both unfamiliar, and thus the CSS neurons in untutored birds preferred one conspecific song over all the other stimuli, including the other conspecific song. The proportion of selective neurons in untutored birds was significantly greater than expected from the type-I error rate of the selectivity test ($\chi^2 = 30, df = 1, P < 10^{-7}$), as was the proportion of selective neurons that preferred a conspecific song ($\chi^2 = 13.6, df = 6, P = 0.035$). Although the number of CSS neurons found in untutored birds is small, this result suggests that even without tutoring some auditory neurons are narrowly tuned to acoustic features present only in a limited number of conspecific songs.
Due to the young ages of the birds, the selectivity we observed here is highly unlikely to reflect singing-related auditory feedback. Furthermore, prior to electrophysiological experiments 19 of the subjects were transferred to isolation booths with one or both parents to record the juvenile’s begging calls, the mother’s long calls, and to assess the juvenile’s singing behavior. In a total of 212 hr of recordings (mean ± SD = 19.3 ± 16.4 hr/bird) from 11 individuals, only one bird produced rudimentary subsong, indicating that recordings took place before the onset of singing. Although moving the birds into a sound booth may have suppressed singing, the lack of singing at this early age that we observed is also consistent with prior observations (Zann, 1996).

Population selectivity of auditory neurons

We also examined selectivity at the population level to see whether the distribution of responses to different classes of stimuli tended to be correlated across the population, and whether there were distinct clusters of neurons that showed similar responses. We set up a matrix of units vs. response strengths to different stimulus classes (see Materials & Methods) and used principal components analysis to examine the distribution of neurons in a reduced space. Neurons from untutored birds were not included in this analysis, because one of the stimulus classes presented to the tutored birds, Tut, by definition could not be presented to untutored birds.

Figure 4 shows the projection of each neuron’s response profile (a 10-dimensional vector) into this reduced space. Overlaid on the plot are arrows indicating the loadings of each stimulus for the first two PCs. A distinct cluster of neurons in the lower left quadrant is defined by strong responses to Tut and weak responses to other stimuli (Fig. 4, triangles). Many of the neurons in this cluster were identified as selective for tutor song ($d' > 0.77$) as described in the previous
section (TSS neurons; numbered symbols); other neurons in the cluster failed to meet this
criterion but also showed a clear preference for the tutor song by this analysis. Throughout the
rest of the population, other neurons gave strong responses to Mun and BF, or to JuvAM and
WN (squares and open circles, respectively), as indicated by their projections along the vectors
corresponding to those stimuli. In contrast to the Tut-preferring neurons, these neurons did not
form distinct clusters (i.e., with little or no contamination from other clusters). For example, a
line surrounding all the triangles (Tut-preferring neurons) defined an area that contained only a
single neuron that was not Tut-preferring. Similar areas defined for each of the other clusters
contained more contaminating neurons (Fig. 4, white noise (square) and JuvAM (open circle);
other clusters not shown).

The loadings of the ten stimulus classes on the first two PCs reflect correlations in the
population’s response to different stimuli. Responses to stimuli with similar acoustic properties
tended to show significant positive correlations. The Mun and BF stimuli are similar
vocalizations from wild and domesticated strains of the same species, and the correlation
coefficient of the neuronal responses associated with those stimuli was 0.37 (P = 0.0001).
JuvAM and WN have similar spectral properties but different amplitude envelopes, and their
associated neuronal responses correlation coefficient was lower, only approaching significance (r
= 0.17, P = 0.056). In contrast, the correlation between responses to Tut and all of the other
stimuli was either negative (Juv, JuvAM, WN) or not significantly different from zero (BF, Beg,
FLC, Mun, Unf, ConAM). No other stimulus class had neuronal responses as isolated from
responses to other stimuli as did Tut. Responses to Unf stimuli showed the highest correlation
with tutor song responses, but this was not significant (r = 0.11, P = 0.22). Given that Unf and
Tut stimuli share the most acoustic features with each other, some degree of correlation is
expected (and might be statistically significant with a large enough sample size), but the weakness of this correlation relative to the correlations observed between other acoustically similar stimuli is unexpected if the correlations between songs were also based only on acoustic similarity. Instead, this suggests that the population discriminates between the tutor song and all other stimuli on the basis of experience. This result, which obtains from an analysis that makes no a priori assumptions about measures of neuronal stimulus selectivity, emphasizes that tutor-song selective neurons are part of a larger, distinct group of neurons within the population we recorded. Note that the analysis identifying TSS neurons and the PC analysis of the population were distinct from each other. The analyses could have identified TSS neurons without finding a larger population of cells with similar response distributions, and we could have found the larger population without any of the cells being TSS.

Study 2: controlled search strategy for tutor song selectivity

We conducted a second study to confirm the results of Study 1 and to determine whether selectivity of TSS neurons is specific to the tutor’s song, or if it simply reflects familiarity. In this study, responses to tutor songs (Tut) were compared with responses to two other conspecific songs, one unfamiliar (Unf) and one the familiar song of a neighboring bird (Fam). In addition, in order to sample more from CSS neurons, we used a more limited stimulus set comprising the three adult songs as well as JuvAM, WN, and Juv. As in Study 1, different neurons in each bird were presented with different unfamiliar conspecific songs. The response strength of the units was assessed online with rasters and PSTHs to identify neurons that were selective for any adult zebra finch song. At 128 recording sites in ten birds (P21–P24), we first recorded responses to five repetitions of each of the six stimulus exemplars. Of these sites, 109 were determined online
(without the benefit of offline spike sorting) to respond either more strongly to WN, JuvAM or Juv, or indiscriminately to all stimuli, and were not tested further. At the other 19 sites (7 birds) where the response was strongest to one of the three adult songs, 15 additional stimulus repetitions were presented. Thus, this experiment was designed to focus on selectivity within the class of adult songs, for neurons known to prefer this class of stimulus.

Of the 24 units extracted offline after the experiments from the 19 sites, 12 of them (from 11 sites) were selective ($d' > 0.77, P < 0.05$). One unit preferred JuvAM, and 11 others (5 birds; 1–5 cells/bird; Table 1) preferred the tutor song. The TSS cells exhibited robust firing to the tutor song but only weak responses to the other stimuli. In eight of those cells, adult conspecific songs – familiar and unfamiliar – evoked the second largest response (Table 1), but the response to tutor song was significantly stronger. Assuming the same yield of single units per site from the unanalyzed sites in Study 2 as observed in Study 1 (1.63 units/site), the proportion of CSS neurons was similar in both studies (including both conditions of Study 1), and the proportion of TSS units in tutored birds from both studies was nearly identical (Fig. 5).

Locations of tutor song selective cells

Anatomical reconstruction of the recording location on each side of the brain was successful for 18 out of 19 tutor song selective sites from both studies. TSS sites were distributed among three nidopallial regions: the HVC ($n = 7$), the paraHVC ($n = 5$), or deeper in NCL ($n = 6$). A similar distribution was observed if limiting consideration to sites whose location was most securely determined, based on fiduciary lesions made at the recording sites ($n = 8$) or within 300 µm along the same electrode penetration ($n = 7$). Those 15 sites were distributed in the HVC (6 sites,
4 birds), the paraHVC (3 sites, 3 birds) and deeper in NCL (6 sites, 3 birds) (Fig. 6). The location of the recording sites differed somewhat between the two studies. In Study 1, most sites were primarily in HVC, paraHVC and HVCshelf, whereas in Study 2 the electrodes went deeper into NCL, in addition to HVC, paraHVC and HVCshelf. A discriminant analysis applied to the stereotaxic coordinates of all recording sites (rostro-caudal position, medio-lateral position and depth) revealed that depth was the variable that best separated the data from the two studies. Subsequent analysis confirmed that in Study 1 our recording sites were more superficial (mean ± SD = 720.1 ± 273.5 µm; n = 108) than they were in Study 2 (918.3 ± 178.4 µm; n = 18) (t-test corrected for unequal variances: t = 4.0; P < 0.0004). Correspondingly, Study 1 accounted for most of the TSS neurons from HVC and paraHVC, and Study 2 accounted for most of the TSS neurons in NCL. In Study 1, selective neurons were found in all four areas, and the distribution of selectivity was similar for neurons in HVC and in other regions ($\chi^2 = 10.5$, $df = 12$, $P = 0.57$; Table 2). Note that we only placed fiduciary lesions at or near sites of putative tutor song selective cells, so the location of other selective neurons was overall less reliable. Both studies included recordings from sites in the HVC shelf, but none of the units in that region were selective for the tutor song (Margoliash 1983, 1986).

**DISCUSSION**

These results demonstrate the existence in pre-singing zebra finch juveniles of a small population of higher-order neurons selective for specific adult songs. We recorded the responses of neurons in HVC, HVCshelf, paraHVC, and NCL and found that 23% of the units were selective for one of ten natural and synthetic stimuli. Of these neurons about 30% preferred the song of a conspecific, whereas most of the remainder were selective for broadband stimuli like white noise.
and amplitude-modulated noise (Fig. 2). In normally-raised, tutored birds all of the conspecific
song selective neurons preferred the song of the bird’s tutor, and we found no neurons that were
selective for any unfamiliar conspecific songs. A principal component analysis of the responses
of the neurons in normally raised birds identified a larger population of neurons that had
qualitatively similar properties. This cluster of cells responded more strongly to the tutor song
and more weakly to other stimuli, including the songs of unfamiliar conspecifics (Fig. 4).

We conducted a second set of recordings in normally raised birds in which we presented a song
from a familiar bird in addition to the songs of the tutor and an unfamiliar male. In these
recordings we specifically targeted neurons that were identified during the recording session as
selective for conspecific song, in order to maximize our sample from these relatively rare
neurons. Out of 12 selective neurons, 11 preferred the tutor song and none preferred the familiar
or unfamiliar song. This indicates that the observed selectivity for tutor song was not
confounded with a familiarity effect – an issue that has not been addressed in previous studies
(Solis and Doupe 1997, 1999; Bolhuis et al. 2000, 2001; Nick and Konishi 2005a, 2005b; Phan
et al. 2006).

The proportion of TSS neurons we found is relatively small (about 6% in both studies), though it
is highly significant statistically, and the neurons that met our criterion for selectivity appear to
be part of a larger cluster with weaker selectivity but similar response profiles (Fig. 4). Higher-
order neurons with mnemonic responses for behaviorally relevant stimuli are broadly known in
adult sensory systems (O’Neill and Suga 1979; Margoliash 1983; Tsao et al. 2006), often in a
limited proportion of the population (Desimone et al. 1984). In many cases the selectivity can be
shown to reflect the animal’s prior sensory experience, confirming that learning can radically
alter the processing of sensory signals (Gentner and Margoliash 2003; Nick and Konishi, 2005a,b; Quiroga et al. 2005). The present results indicate that high-order selectivity for specific stimuli can emerge very early in development, on the basis of limited sensory experience. A form of tutoring (instrumental conditioning with song against a background of isolation rearing; Adret 1993) in young zebra finches has been shown to shape spontaneous activity in RA neurons, downstream from HVC (Shank and Margoliash 2009).

A further question is whether tutor song selectivity arises from a population of neurons that are broadly tuned to generic acoustic features, or whether some degree of selectivity for conspecific acoustic features already exists. To begin to address this question, we also recorded from juvenile birds denied exposure to adult song tutoring, obtaining however only a small sample of neurons from the juveniles. We found a similar proportion of neurons in these birds that were selective at our criterion ($d’ > 0.77$), and the distribution of selectivity among the stimuli was statistically indistinguishable from normally raised birds (Fig. 2B). A significant proportion of the selective neurons preferred one of the two conspecific songs, though both were unfamiliar and the birds were not exposed to adult song after P4. This result suggests that even without exposure to song, some neurons may be narrowly tuned to acoustic features of conspecific song. Given the small number of neurons and selective neurons we recorded from untutored birds, additional work is needed to confirm this result and to examine whether tutor song selectivity arises from neurons already selective for conspecific song or from the broader population.

Most of the TSS cells appeared to respond to a specific part of the song (Fig. 3), suggesting that the cells were responding to a specific set of learned features of the tutor song. Whereas the parametric acoustic basis of the selectivity of TSS cells remains to be established, in particular
the integration time and whether the cells require temporal sequences of tutor song syllables for excitation (Margoliash 1983; Margoliash and Fortune 1992) the inclusion of several synthetic, broadband stimuli in the design of these experiments renders unlikely that the apparent selectivity for learned stimuli actually arises from selectivity for relatively simple acoustic features, because such neurons would be likely to also respond strongly to one of the other stimuli. Neurons with strong preferences for WN and other stimuli were indeed present in these areas, but the population analyses (Fig. 4) indicate that they belong to a distinct class from TSS neurons.

We sampled somewhat different regions of the caudal nidopallium in the two studies but found TSS cells in both studies. A fraction of TSS cells was located in the forebrain song system nucleus HVC (and paraHVC), which exhibits pre-motor activity patterns during singing (McCasland and Konishi 1981). Perhaps some of these cells – particularly those with reliable firing – will later express motor patterns of activity mirroring juvenile auditory patterns of response to tutor song (Prather et al. 2008). Indeed, HVC has little or no influence on motor output at the onset of singing (Aronov et al. 2008), which gives additional focus to its potential role in sensory acquisition (Roberts et al. 2010). Physiological experiments have reported TSS neurons in the basal ganglia pathway of the song system at intermediate stages of song development (Solis and Doupe 1997), and pharmacological experiments also implicate the basal ganglia pathway in sensory acquisition (Basham et al. 1996). The projection of HVC to Area X, which contains the basal ganglia components of the song system, is fully established in zebra finches at P25, as are the connections of paraHVC, which receives direct input from HVC as well as from the song control nuclei MMAN and the medial region of LMANshell. (Foster and Bottjer 1998). ParaHVC in turn projects to both area X and to adjacent regions in NCM.
(Nordeen et al. 1987; Johnson and Bottjer 1995; Foster and Bottjer 1998). The sensory memory represented by TSS neurons within HVC and paraHVC in pre-singing birds could help shape downstream selectivity in the basal ganglia.

Another fraction of TSS cells was located in the caudal lateral nidopallium (NCL), which lies ventral and caudal to HVCshelf and paraHVC in the songbird auditory forebrain (Müller and Leppelsack, 1985). The dorsal region of NCL receives input from LMANshelf and projects back to that structure as well as to the dorsal archistriatum, and lesions of this parallel vocal pathway in young birds suggest a role in auditory-motor integration (Bottjer and Altenau 2010). Perhaps TSS cells in NCL will prove to establish synaptic connections with this parallel vocal pathway.

Auditory gating and state-dependent neural sensitivity

We recorded from very young birds to avoid ambiguity in interpretation that would otherwise arise from singing-related auditory feedback. At P21–P24, most birds did not sing, and the one that did only produced rudimentary subsong. A more faithful copy of song only emerges later in song development, avoiding the caveat that the feedback could induce corresponding changes in sensory neurons to match the motor output. Furthermore, the very act of singing could induce changes in the representation of the tutor song – the representation of the tutor song need not be static but may change over time. For example, TSS multiunit responses have been described in HVC at intermediate stages of song development in awake birds, whereas during sleep the neurons preferentially responded to the bird’s plastic songs (Nick and Konishi 2005a). Our results, obtained in anesthetized birds, suggest that these state-dependent mechanisms (Dave et al. 1998; Schmidt and Konishi 1998) are not yet in place in pre-singing birds.
Functional implications

Supporting the two-step model of song learning i.e., memorization followed by production (Konishi 1965), we identified a narrow window during zebra finch development just prior to the onset of singing (P25) when birds showed evidence of tutor song memorization but had yet to begin subsong (early singing). This confirms Immelmann’s (1969) anecdotal observation of one zebra finch memorizing tutor song prior to P25. However, with such limited experience, the bird failed to copy the model faithfully (see also Roper and Zann 2006). Likewise, our four juveniles that were tutored until P25 developed structured songs with introductory syllables followed by repeated song phrases but there was a large between-individual variability in the amount of material learnt and song imitation was imperfect. Perhaps it is not surprising that we identified only a small percentage of TSS cells (1 to 5 cells/bird) and in only about half of our subjects. Thus, the neuronal representation of the tutor song may be immature at the ages when the experimental design dictated we perform the recordings. By contrast, zebra finches trained with a live tutor over a 10-day period starting at P25 have been shown to develop accurate song copies (Roper and Zann 2006). Many birds at P25–P35 will also have begun singing. Yet sensory and motor representations of song interact, even in the same neurons (Dave and Margoliash 2000; Nick and Konishi 2005a; Prather et al. 2008), and at the earliest stages of song learning (Shank and Margoliash 2009). Such confounds were largely avoided by restricting the present study to younger, pre-singing birds.

The behavioral significance of the small populations of selective neurons we have described remains to be established. We hypothesize that these cells represent the early emergence of a sensory memory, contributing to the long-hypothesized “acquired template” of the tutor song.
(Konishi 1965, 1978) formed early in development that guides subsequent sensorimotor song learning. The selective responses to conspecific adult songs in untutored birds could represent part of an “innate template” birds use to guide their songs in the absence of environment cues (Marler and Sherman 1983; Konishi 2004). This hypothesis conceives of the acquired sensory template as being represented in small populations of neurons distributed across multiple nuclei. These data cannot rule out the possibility that the acquired template is also represented more robustly elsewhere (Bolhuis and Gahr 2006). In fact, a wealth of observations derived from immediate early gene studies in caudal nidopallial areas outside of the song system have implicated NCM in the formation of auditory memories (Chew et al. 1995; Jarvis et al. 1995; Mello et al. 1995; Chew et al. 1996). Consistent with a tutor song memory effect, manipulations of the tutoring experience have established a linear relationship between gene expression in NCM and the degree of imitative learning (Bolhuis et al. 2000, 2001; Terpstra et al. 2004). Further evidence for a mnemonic role of NCM in model acquisition is supported by recent electrophysiological, pharmacological and lesion studies (Phan et al. 2006; Gobes and Bolhuis 2007; London and Clayton 2008). Still, our results suggest that by the time a bird makes his first efforts at singing, a sensory memory of the tutor song is already distributed throughout multiple forebrain regions in small populations of highly selective neurons.
ACKNOWLEDGMENTS

We thank Kazuo Okanoya for kindly providing recordings of Bengalese finch and white-backed munia subsong, and David J. Freedman and Howard C. Nusbaum for commenting on an earlier draft.

GRANTS

This work was supported by NIDCD grants DC03996 to P. Adret, DC007206 to D. Margoliash, F32DC008752 to C. D. Meliza, and NIMH grant MH59831 to D. Margoliash.
REFERENCES

Adret P. Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Anim Behav* 45: 149-159, 1993


Bolhuis JJ, Zijlstra GGO, Den Boer-Visser AM, and Van der Zee EA. Localized neuronal activation in the zebra finch brain is related to the strength of song learning. *Proc Natl Acad Sci (USA)* 97: 2282-2285, 2000.


Fig. 1. Song imitation from limited exposure to a live tutor. The songs of two sibling juveniles that were exposed to their father’s singing until P25 then isolated. Two versions of the final adult songs that the juveniles developed are shown, the song sung in isolation (“undirected”) and the song sung in the presence of a female (“directed”), as well as an exemplar of developmental plastic singing. Each syllable of the tutor song (model) is labeled; there are many corresponding syllables, which are labeled in the offspring’s songs. Tutee Or305 copied the entire tutor song; imitation of complex syllable D was imperfect. The song syntax was close to the model, with only a temporal inversion for syllables F and G and omission of syllables A and B from some motifs (song matching index: 86%). Tutee Or304 copied the model’s introductory note as well as three syllables, all produced with the correct syntax; imitation of complex syllable D was imperfect (song matching index: 34%).

Fig. 2. The distribution of stimulus preference among selective neurons in tutored and untutored birds. A, Bar plot showing the proportion of 28 selective neurons from tutored birds selective for a given stimulus class. All eight neurons selective for Con were selective for the song of the bird’s tutor. B, Bar plot for 9 selective neurons from untutored birds. All four Con-selective neurons were selective for one of two unfamiliar songs.

Fig. 3. Responses of conspecific song selective neurons in the caudal nidopallium of tutored and
untutored young birds from Studies 1 and 2. Extracellular activity recorded from exemplar neurons drawn from recordings in HVC, paraHVC, and NCL in response to the song of the bird’s tutor (for TSS neurons) or to the song of a conspecific male (CSS neurons). All the CSS neurons were recorded in birds that were untutored; no units in any tutored bird were selective for conspecific song other than the tutor’s. Responses are shown as rasters of spike times and raw electrical activity. Insets (scale 2.4 ms) show 100 random spike waveforms (gray traces) and mean (black trace). One unit (TSS9) had excellent isolation and clipped during recording; this inset shows the first 100 spikes, when clipping was minimal. Labels for the TSS units correspond to numbers in Figure 4 (TSS 2 and 8) and Figure 6 (TSS 9, 11, 13, 19). Based on stereotaxic coordinates the CSS cells in untutored birds were recorded in the same vicinity as were TSS cells in tutored birds, but lesions unambiguously identifying the locations of the CSS cells were not recovered.

**Fig. 4.** Population selectivity of neurons in tutored birds. Each neuron’s response to the test stimuli can be represented by a 10-dimensional vector, which is reduced to two dimensions by principal component analysis. Each point represents the projection of a neuron along the first two PCs (PC1, PC2). The loadings of each class in the first two PCs are plotted as arrows labeled with the name of the class. Clustering in the PC space indicates that neurons respond similarly to the same stimuli, and the position of points relative to the loading arrows indicates which stimuli evoked strong responses for those neurons. Arrow length indicates the variance in responses to a given stimulus, and angles between arrows indicate the correlation in responses to the two stimuli. For example, neurons tended to respond similarly to WN and JuvAM (low population selectivity) but differently to Tut and WN (high population selectivity). Neurons were
categorized by the stimulus that evoked the largest response. Neurons preferring the top three classes of stimuli are indicated by different symbols. The axes values for the projections and loadings are indicated on bottom/left and top/right axes, respectively. Neurons that were significantly selective \( (d' > 0.77) \) for the tutor song over all other stimuli (TSS) are numbered. The proportion of the total variance explained by the first two PCs was 44%.

**Fig. 5.** Proportion of neurons selective for conspecific song from the two conditions of Study 1 and from Study 2. Neurons are subdivided according to which class of conspecific song they preferred (unfamiliar, familiar, tutor). Proportions are relative to the total number of auditory units (for Study 2, this is based on assuming a similar yield of single units from sites that were not conspecific song selective; see Results).

**Fig. 6.** Anatomical reconstruction of recording sites selective for the tutor song in the caudal nidopallium. The composite illustration shows superimposed outlines of structures and landmarks that were drawn from 40 µm thick sagittal sections. The loci of 15 recovered sites (of 19 total — TSS neurons 3, 16, 17, and 18 were > 300 µm from the nearest fiduciary lesion and are not shown) where tutor song selective cells were recorded are shown by different symbols. Sites 1–8 are from Study 1; 9–19 from Study 2. Sites were located in HVC (6 sites; circles), paraHVC (3 sites; triangles) and deeper nidopallium (6 sites; squares). Numerical labels correspond to cell numbers in Table 1. Sites with bold numerals were localized from electrolytic lesions made at the site. Examples of lesions outside of HVC are shown with arrowheads in the micrographs at right. LaM: Lamina mesopallialis; LAD: Lamina arcopallialis dorsalis.
Table 1. Locations and responses of TSS cells. For each neuron, the responses to the two strongest stimuli are shown. RS: response strength (Hz), mean ± SD across trials; z: z-score of RS (Theunissen et al. 2004); $d'$: d-prime statistic quantifying discriminability between tutor song and next strongest stimulus. Cells 1–8 are from Study 1; 9–19 from Study 2. The recording site of cell 3 was not recovered.

<table>
<thead>
<tr>
<th>CELL</th>
<th>SITE</th>
<th>STIM</th>
<th>RS</th>
<th>z</th>
<th>$d'$</th>
<th>BIRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HVC</td>
<td>Tut</td>
<td>10.91 ± 4.83</td>
<td>10.10</td>
<td>0.93</td>
<td>B84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juv</td>
<td>6.27 ± 8.76</td>
<td>3.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>HVC</td>
<td>Tut</td>
<td>5.02 ± 1.45</td>
<td>15.52</td>
<td>2.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mun</td>
<td>1.40 ± 2.28</td>
<td>2.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>Tut</td>
<td>6.60 ± 2.11</td>
<td>14.01</td>
<td>1.55</td>
<td>B90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FLC</td>
<td>4.22 ± 2.26</td>
<td>8.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>HVC</td>
<td>Tut</td>
<td>2.00 ± 1.76</td>
<td>5.07</td>
<td>0.93</td>
<td>G366</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ConAM</td>
<td>0.85 ± 1.76</td>
<td>2.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>HVC</td>
<td>Tut</td>
<td>5.19 ± 3.55</td>
<td>6.55</td>
<td>0.87</td>
<td>Y45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unf</td>
<td>3.40 ± 2.15</td>
<td>7.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>HVC</td>
<td>Tut</td>
<td>2.32 ± 2.49</td>
<td>4.18</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BF</td>
<td>0.92 ± 1.89</td>
<td>2.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>pHVC</td>
<td>Tut</td>
<td>2.77 ± 1.60</td>
<td>7.75</td>
<td>1.86</td>
<td>Y81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mun</td>
<td>0.71 ± 1.54</td>
<td>2.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>pHVC</td>
<td>Tut</td>
<td>4.31 ± 1.96</td>
<td>9.83</td>
<td>1.74</td>
<td>Y100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Beg</td>
<td>1.04 ± 3.20</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>HVC</td>
<td>Tut</td>
<td>3.52 ± 2.21</td>
<td>7.13</td>
<td>1.14</td>
<td>B183</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unf</td>
<td>1.46 ± 2.88</td>
<td>2.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>NCL</td>
<td>Tut</td>
<td>5.55 ± 3.39</td>
<td>7.33</td>
<td>2.46</td>
<td>B190</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unf</td>
<td>0.36 ± 2.51</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>NCL</td>
<td>Tut</td>
<td>0.83 ± 1.14</td>
<td>3.23</td>
<td>1.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fam</td>
<td>−0.20 ± 0.75</td>
<td>−1.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td>Tut</td>
<td>1.70 ± 4.08</td>
<td>1.87</td>
<td>1.28</td>
</tr>
<tr>
<td>Unf</td>
<td></td>
<td></td>
<td></td>
<td>−1.83 ± 3.74</td>
<td>−2.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pHVC</td>
<td></td>
<td></td>
<td>2.20 ± 1.66</td>
<td>5.94</td>
<td>1.59</td>
</tr>
<tr>
<td>Unf</td>
<td></td>
<td></td>
<td></td>
<td>0.64 ± 1.04</td>
<td>2.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td></td>
<td>1.42 ± 0.97</td>
<td>6.54</td>
<td>1.22</td>
</tr>
<tr>
<td>JuvAM</td>
<td></td>
<td></td>
<td></td>
<td>0.44 ± 1.28</td>
<td>1.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td></td>
<td>7.90 ± 3.79</td>
<td>9.32</td>
<td>2.40</td>
</tr>
<tr>
<td>Fam</td>
<td></td>
<td></td>
<td></td>
<td>2.28 ± 2.76</td>
<td>3.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td></td>
<td>0.99 ± 1.24</td>
<td>3.60</td>
<td>1.21</td>
</tr>
<tr>
<td>Fam</td>
<td></td>
<td></td>
<td></td>
<td>0.17 ± 0.59</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td></td>
<td>1.71 ± 2.03</td>
<td>3.76</td>
<td>1.05</td>
</tr>
<tr>
<td>Fam</td>
<td></td>
<td></td>
<td></td>
<td>0.40 ± 1.44</td>
<td>1.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td></td>
<td>3.82 ± 2.99</td>
<td>5.71</td>
<td>1.52</td>
</tr>
<tr>
<td>JuvAM</td>
<td></td>
<td></td>
<td></td>
<td>1.39 ± 1.12</td>
<td>5.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCL</td>
<td></td>
<td></td>
<td>6.88 ± 2.55</td>
<td>12.06</td>
<td>1.52</td>
</tr>
<tr>
<td>WN</td>
<td></td>
<td></td>
<td></td>
<td>4.28 ± 2.27</td>
<td>8.41</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Distribution of selectivity by area for neurons from Study 1 in tutored birds. Numbers in each column indicate percentage of all neurons (top) or selective neurons (bottom) selective ($d' > 0.77$) for each class of stimulus. “Non-HVC” represents the sum of cells from paraHVC, HVC<sub>shelf</sub>, and NCL. Note the similarity of value for HVC and non-HVC. N: number of cells.

<table>
<thead>
<tr>
<th>Area</th>
<th>All neurons</th>
<th>WN</th>
<th>JuvAM</th>
<th>ConAM</th>
<th>BF</th>
<th>Mun</th>
<th>Beg</th>
<th>Juv</th>
<th>Tut</th>
<th>FLC</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVC</td>
<td>7.2</td>
<td>3.6</td>
<td>1.2</td>
<td>1.2</td>
<td>2.4</td>
<td>0</td>
<td>1.2</td>
<td>7.2</td>
<td>0</td>
<td></td>
<td>83</td>
</tr>
<tr>
<td>paraHVC</td>
<td>0</td>
<td>5.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.6</td>
<td>11.1</td>
<td>0</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>HVC&lt;sub&gt;shelf&lt;/sub&gt;</td>
<td>16.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>NCL</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>non-HVC</td>
<td>7.7</td>
<td>2.6</td>
<td>0</td>
<td>0</td>
<td>2.6</td>
<td>0</td>
<td>2.6</td>
<td>5.2</td>
<td>0</td>
<td></td>
<td>39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Selective neurons</th>
<th>HVC</th>
<th>30</th>
<th>15</th>
<th>5</th>
<th>5</th>
<th>10</th>
<th>0</th>
<th>5</th>
<th>30</th>
<th>0</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>paraHVC</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>HVC&lt;sub&gt;shelf&lt;/sub&gt;</td>
<td>75</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>NCL</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>non-HVC</td>
<td>37.5</td>
<td>12.5</td>
<td>0</td>
<td>0</td>
<td>12.5</td>
<td>0</td>
<td>12.5</td>
<td>25</td>
<td>0</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>
Song selective neurons (proportion of total)

Study 1

<table>
<thead>
<tr>
<th>Stimulus Class</th>
<th>Untutored</th>
<th>Tutored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unf</td>
<td>4/36</td>
<td></td>
</tr>
<tr>
<td>Tut</td>
<td>8/122</td>
<td>0/122</td>
</tr>
<tr>
<td>Unf</td>
<td>0/122</td>
<td></td>
</tr>
</tbody>
</table>

Study 2

<table>
<thead>
<tr>
<th>Stimulus Class</th>
<th>Tutored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tut</td>
<td>11/184</td>
</tr>
<tr>
<td>Fam</td>
<td>0/184</td>
</tr>
<tr>
<td>Unf</td>
<td>0/184</td>
</tr>
</tbody>
</table>