Title: Distributed and selective auditory representation of song repertoires in the avian song system

Authors: Paul M. Nealen$^{1,2}$
Marc F. Schmidt$^{1,3}$
$^1$Department of Biology
$^3$Neuroscience Graduate Group
University of Pennsylvania
415 S University Ave
Philadelphia, PA 19104-6018

Running head: Avian song repertoire representation in nucleus HVC

$^2$Correspondence to: Paul M. Nealen
Department of Biology
Indiana University of Pennsylvania
114 Weyandt Hall
975 Oakland Avenue
Indiana, PA 15705-1001
pnealen@iup.edu

* tel: 724-357-1325
* fax: 724-357-5524

* For use during production only; please do not include in published form
ABSTRACT
For many songbirds, the vocal repertoire comprises acoustically distinct songs that are flexibly used in various behavioral contexts. To investigate how these different vocalizations are represented in the song neural system, we presented multiple song stimuli while performing extracellular recording in nucleus HVC in adult male song sparrows *Melospiza melodia*, a species known for its complex vocal repertoire and territorial use of song. We observed robust auditory responses to natural song stimuli in both awake and anesthetized animals. Auditory responses were selective for multiple songs of the bird’s own repertoire (BOR) over acoustically-modified versions of these stimuli. Selectivity was evident in both awake and anesthetized HVC, in contrast to auditory selectivity in zebra finch HVC which is apparent only under anesthesia. Presentation of multiple song stimuli at different recording locations demonstrated that stimulus acoustic features and local neuronal tuning both contribute to auditory responsiveness. HVC auditory responsiveness was broadly distributed and non-topographic. Variance in auditory responsiveness was greater among than within HVC recording locations in both anesthetized and awake birds, in contrast to the global nature of auditory representation within zebra finch HVC. To assess the spatial consistency of auditory representation within HVC, we measured the repeatability with which ensembles of BOR songs were represented across the nucleus. Auditory response ranks to different songs were more consistent across recording locations in awake than in anesthetized animals. This spatial reliability of auditory responsiveness suggests that sound stimulus acoustic features contribute relatively more to auditory responsiveness in awake than in anesthetized animals.

KEYWORDS
Auditory selectivity, birdsong, communication, song system, vocal repertoire, HVC
INTRODUCTION

Passerine birds comprise a powerful model system for the study of vocal learning. Their utility stems, in part, from the diversity of vocal behaviors passerines exhibit (Brenowitz and Beecher 2005). While most studies have utilized species that sing a single adult song, study of other species possessing multiple songs remains essential for full evaluation of models of vocal learning, perception and production (Nealen and Schmidt 2002). Much of the theoretical work in songbird vocal learning makes use of a single “bird’s own song” (BOS) as an immutable template or target, but to learn and use multiple vocalizations may be the more common strategy (e.g., Kroodsma and Baylis 1982).

One site of considerable interest within the avian song neural system, nucleus HVC (used here as the proper name), lies at the confluence of auditory and motor streams (McCasland and Konishi 1981; Nottebohm et al. 1976; Figure 1). HVC neurons exhibit complex auditory properties (Lewicki and Arthur 1996; Theunissen and Doupe 1998), and may function in the auditory-motor transformation necessary for song learning and maintenance (Troyer and Doupe 2000). Under anesthesia, HVC auditory responses in several species are consistently selective for BOS over conspecific (CON) song (Cardin and Schmidt 2003; Janata and Margoliash 1999; Margoliash 1986; Mooney et al. 2001). Mooney et al. (2001) provided the first description of song repertoire representation within nucleus HVC of the anesthetized swamp sparrow *Melospiza georgiana*, demonstrating that HVC neuronal subtypes have distinct auditory properties and may exhibit single-syllable selectivity. Given that swamp sparrow songs consist of repetitions of a single syllable, it is difficult to predict how HVC neurons would respond to multiple songs consisting of variable sequences of different syllables. It is also unclear why the robust and selective auditory responses observed under anesthesia in the zebra finch and other species have not been observed in awake animals (Cardin and Schmidt 2004, 2003; George et al. 2005).

To examine the auditory representation of song repertoires, we performed extracellular neural recordings in nucleus HVC of awake and anesthetized male song sparrows *M. melodia*. Song sparrows learn and use
repertoires of 5 – 15 different song types constructed from a large, shared pool of syllables (Podos et al. 1992; Figure 2b). Neighboring males may song type-match, singing nearly identical copies of subsets of their song repertoires (Beecher et al. 2000). In contrast to the colonially-breeding zebra finch, song sparrows are strongly territorial, are highly responsive to conspecific song, and can discriminate among individuals and among syllable types on the basis of song (Nelson 1987; Stoddard et al. 1991).

In accord with prior studies, we demonstrate robust and selective auditory responses from anesthetized HVC. We also provide the first evidence of consistently robust and selective auditory responses in awake animal HVC. We provide estimates for the independent contributions of stimulus and recording location to the dynamic range of auditory response magnitudes, and demonstrate greater variability in auditory responsiveness among than within HVC recording locations, in contrast to the “global” auditory representation observed in anesthetized zebra finch HVC (Sutter and Margoliash 1994). Using measures of the spatial reliability of auditory responses across multiple HVC recording sites, we show that auditory responses are more spatially consistent in awake than in anesthetized animals, a finding which suggests that the neural basis for auditory selectivity within nucleus HVC is not entirely conserved between the awake and anesthetized conditions.
MATERIALS AND METHODS

Animal song recording and collection from natural habitats

Adult song sparrows were located on breeding territories in natural habitats in southeastern Pennsylvania. Spontaneous song was recorded from territorial males via portable analog recording equipment (Sony PBR-330 parabolic reflector, Optimus 33-3014 microphone, Marantz PMD-430 cassette recorder). Song was recorded from both focal and neighboring males (Figure 2a). While it was not possible to exhaustively sample the entire song repertoire of individuals (bird’s own repertoire; BOR), song recording for several hours was sufficient to identify a frequently-used subset of an individual’s song repertoire, consisting of multiple, different BOS (bird’s own song) song types. When possible, song was recorded from birds which had been color banded for individual identification; song was collected from unbanded males only when individuals could reliably be followed on their territories. Focal males were then captured via mist nets, and brought to the laboratory. Captured birds were held in individual cages in a common aviary. State and federal permits for the capture and use of wild birds were in hand. All methods were approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania.

Laboratory sound recording and sound stimulus preparation

Prior to neural recordings, individual male song sparrows were housed for several days in sound attenuation chambers (Acoustic Systems, Austin, TX) for computer-based recording of spontaneous songs. A minimum of four “bird’s own song” exemplars representing unique song types (BOS1-BOS4; Figure 2b) were chosen from the field and laboratory song recordings available for each individual for use as auditory stimuli during neural experiments.

Song stimuli used for playback were digitized at 40 kHz and high-pass filtered at 300 Hz. Four acoustic variants of individual BOS sound stimuli were constructed for use as test stimuli, including BOS (in its original, forward orientation) and REV (whole-song reversed) sound stimuli. Other acoustic variants
included song stimuli whose syllables were presented in the original sequence but in individually-reverse orientation (R-SYL), and stimuli consisting of the original song syllables played in correct orientation but in reverse-order sequence (R-ORD). These latter classes of stimuli were used to assess the effects of syllable sequencing and syllable context in eliciting auditory responses. If, for example, the acoustic features of individual syllables contribute most to HVC auditory responsiveness, R-ORD sound stimuli should elicit robust auditory responses. If, however, the manner in which acoustic spectra evolve over the time course of a sound stimulus is critical to auditory responsiveness, R-SYL sound stimuli would be expected to provide auditory drive similar to that of natural BOS stimuli. Short segments of a white-noise stimulus (WN) and a heterospecific (zebra finch *Taeniopygia guttatta*) song stimulus were used during some experiments to measure the degree of non-specific auditory responsiveness. In addition to the spontaneous neural data collected before each auditory presentation, during some experiments neural data were collected during experimental trials containing no auditory stimulus (BLANK) which were interleaved among auditory test stimuli. Auditory trials involving BLANK, WN, conspecific (CON) and HET stimuli were excluded from present analyses.

**Experimental procedure – acute experiments**

Animals were removed from food/water for 1-2 hr prior to administration of either Diazepam (10 mg / kg) or a combination of Ketamine / Xylazine (35 / 7 mg / kg), injected IM. After drug onset, subjects were placed into a protective surgical jacket heated to 35 °C. Scalp feathers were plucked; topical antiseptic (Betadine, Purdue Frederick, Stamford, CT) and anesthetic (2% lidocaine, Copley Pharmaceutical, Inc., Canton, MA) were administered to the scalp. Subjects were then placed into a stereotaxic apparatus. The scalp was split and retracted, and small ink marks were made on the skull surface above nucleus HVC, as located via stereotaxic coordinates. A small post was cemented to the anterior skull with dental acrylic (Denstsply, York, PA) and then used to support the animal’s head, leaving the auditory canals unobstructed.
Recordings were made unilaterally in nucleus HVC (Figure 1a); both left and right hemispheres were sampled. A craniotomy (150 – 300 µm) was made over the desired recording location, and the dura was cut with a sharpened tungsten wire. Extracellular recordings were made using commercial tungsten electrodes (2 – 5 MΩ, A-M Systems, Carlsborg, WA) or glass micropipettes (A-M Systems) of tip diameter 8 – 15 µm filled with 4 M K-acetate. In several subjects, recordings were performed via cell-attached recording, providing high quality single-neuron isolation. Neural signals were amplified 500x and filtered between 300 – 10,000 Hz (HS4 headstage and DB4 Bio-amp controller, Tucker-Davis Technologies, Alachua, FL; Brownlee 410 amplifier, Brownlee Precision Co., San Jose, CA). Neural signals were monitored on-line via oscilloscope (Hameg Instruments, Frankfurt, FRG) and sound monitor (Grass-Telefactor, West Warwick, RI).

Correct neural recording location was achieved via stereotaxic electrode placement, confirmed by diagnostic spontaneous neural activity. Recording locations in initial study animals were confirmed via histological examination; all were confined to within the boundaries of nucleus HVC, as defined by cresyl-violet stain for Nissl substance. Once neural recording was achieved, sound stimuli were presented via passive playback in an interleaved fashion, with variable inter-trial intervals (5 – 15 sec). Sound stimuli were normalized to achieve maximum amplitude of 70 – 75 dB at the animal. Auditory presentation and concomitant neural data were digitized at 20 kHz and recorded via custom software (A. Leonardo, Cal Tech) for off-line analysis.

At each recording location, a small number (1 – 18, mean: 3.92) of the available sound stimuli for each animal was presented, with each stimulus presented between 5 – 100 (mean: 31) times. Following a sound presentation block, the recording electrode was moved a minimum of 50 µm to a new recording location before further presentation of sound stimuli.
After 2 – 3 hours of auditory stimulation, the recording electrode was withdrawn. The skull craniotomy was filled with GelFoam (Pharmacia & Upjohn, Kalamazoo, MI), and the scalp closed with sutures. The animal was removed from the surgical apparatus, and allowed to recover from anesthesia in an undisturbed, heated cage. Recovery from anesthesia was rapid; animals typically resumed perching and feeding within 1 – 2 hours. Subjects were allowed to recover a minimum of 1 week before further neural recording.

**Experimental procedure – chronic experiments**

Subjects were prepared for neural recording as described above. Several animals received bilateral, fixed electrodes in the following manner. Following stereotaxic location of left and right HVC, a single blunt-end recording electrode (0.5 – 1.5 MΩ) made from formvar-insulated nichrome wire (A-M Systems) was placed into each left and right HVC, as described previously (Schmidt 2003). Electrode leads were connected to a microconnector (Omnetics, Minneapolis, MN) cemented to the animal’s skull. Other animals received unilateral tungsten electrodes (from 1 –3 in number), moveable via custom-built manual or motorized microdrives (Fee and Leonardo 2001).

After a 24 hr recovery period, animals were connected via a wire tether to a custom recording chamber, which allowed full range of motion and continuous neural recording. Subjects exhibited all normal behaviors while in the recording chamber, including perching, feeding, calling and singing. The recording chamber itself was housed within a sound attenuation chamber containing a speaker over which sound stimuli were presented (as above) for assessment of neural responses during auditory presentation in awake animals. A camera and microphone were used to remotely monitor the animals during these experiments; at no time were animals observed to sleep during sound presentation.
Data processing and analysis – extraction of spike events and general methods

Neural recordings were processed via custom software in the MatLab environment (MathWorks, Natick, MA). All statistical analyses were performed using SAS ver. 8.00 (SAS Institute, Cary, NC) and Microsoft Excel (version 2002). Data means are presented as mean ± SE. Distribution central tendencies were compared using analysis of variance (ANOVA; F statistic reported) or via t-test (t statistic reported). T-test comparisons were paired by recording location where reported, and are unpaired otherwise. Prior to t-test computation, distributions were evaluated for equivalence of variance. Contingencies were evaluated using sign tests (M value reported).

Neural traces dominated by the activity of one or a few individual neurons were subject to template-based spike detection and sorting (Spike2 software, version 5, CED Ltd., Cambridge, UK). Other, multi-unit neural traces were subjected to thresholding spike detection. Single- and multi-unit recordings are combined in analyses except where indicated. All spike event times were binned at 10 msec intervals for analyses. For all repetitions of a given sound stimulus at one recording location, neural activity event times were summarized across trials into a peri-stimulus time histogram (PSTH; see Figure 3). PSTH bin values were normalized to the number of auditory trials in each experimental block, and thus ranged from 0 (no neural activity over a given 10 msec interval during any of the auditory presentations of a sound stimulus) to 1 (neural activity in a given 10 msec interval during every presentation of a sound stimulus). The vast majority of PSTH bin values reflect values between these two extremes. For each of the three phases of an experiment (before [duration 1.0 – 1.5 sec], during [duration 2 – 4 sec], and after [duration 2 – 4 sec] auditory stimulation), the minimum, maximum, mean and variance were calculated from PSTH bin values. Minimum, maximum, and mean of PSTH bin values reflect the overall levels of neural activity, while variance in PSTH bin values is a measure of the modulation of excitation during auditory presentation; high variance suggests sensitivity to temporal structure in acoustic stimuli.
Recording locations were classified as “auditory” if any of the sound stimuli presented at that location elicited an elevation of neural activity over pre-stimulus levels. Recording locations at which no stimuli elicited above-background levels of excitation were not included in subsequent analyses. Subsequent representations of auditory-evoked neural activity were expressed as response strength (RS) measures, in which firing rate during stimulus presentation is normalized to the pre-stimulus rate of spontaneous activity:

\[
\text{Response strength (RS)} = \frac{\text{mean(PSTH}_{\text{during}})}{\text{mean(PSTH}_{\text{before}})}
\]

The degree of selectivity for one stimulus over another in pairwise comparisons was expressed in terms of the psychophysical metric d’ (Green and Swets 1966):

\[
d'_{A-B} = \frac{2(\bar{RS}_A - \bar{RS}_B)}{\sqrt{\sigma^2_A + \sigma^2_B}}
\]

where RS and \( \sigma^2 \) refer to the strength and variance, respectively, of auditory responses to song stimuli A and B.

No selectivity (e.g., equivalent responses to two stimuli) would result in a selectivity index of 0 (Green and Swets 1966; Janata and Margoliash 1999). Given that this metric is symmetric about 0 (Janata and Margoliash 1999), the degree to which observed indices reflect consistent selectivity was tested by comparing the distributions of observed selectivity indices to 0.

Stimulus ensemble reliability measurements

Multiple BOS stimuli from any one bird’s repertoire were used as test stimuli at multiple recording locations within HVC in that individual. As described above, the size of the auditory response evoked by any one test stimulus was measured as the elevation in neural activity over pre-stimulus levels (RS value).
To test the degree to which auditory representations in nucleus HVC are global in nature (Sutter and Margoliash 1994), the magnitude of response evoked by a given stimulus was compared across the recording locations at which a stimulus was tested.

To assess the uniformity of auditory responses evoked by test stimuli, the set of auditory responses evoked by presentation of an ensemble of BOS songs was compared across recording locations within individuals. At any one recording location \((j)\), the BOS stimuli tested \((BOS_{1, i})\) were ranked \((r_j)\) according to the magnitude of the auditory response \((RS\) value) evoked by each stimulus. The order of this ranking was then compared across each location at which the stimulus ensemble was tested. If recording locations across HVC represent BOS auditory stimuli in the same way, the relative ranking of responses evoked by a set of test stimuli would repeat across recording sites. For example, consider an ensemble of three test stimuli \((i: BOS_{1-3})\) tested at each of four recording locations \((j_{1-4})\). At recording location \(j_1\), the auditory responses (measured as \(RS\)) evoked by these stimuli could be ranked \((r_1)\) as following:

\[
r_1 = [ RS_{BOS2} > RS_{BOS3} > RS_{BOS1} ]
\]

If the response properties of the neurons sampled at any one recording location are reflective of HVC global response properties, stimulus rankings \((r_{2-4})\) at other recording locations \((j_{2-4})\) would be the same:

**Uniform example:**

\[
\begin{align*}
r_1 &= [ RS_{BOS2} > RS_{BOS3} > RS_{BOS1} ] \\
r_2 &= [ RS_{BOS2} > RS_{BOS3} > RS_{BOS1} ] \\
r_3 &= [ RS_{BOS2} > RS_{BOS3} > RS_{BOS1} ] \\
r_4 &= [ RS_{BOS2} > RS_{BOS3} > RS_{BOS1} ]
\end{align*}
\]

However, if the subpopulation of HVC neurons sampled at a given recording location are tuned to specific song or acoustic features, uniformity of response rankings across recording locations will not obtain, for example:
Variable example: 

\[ r_1 = [ RS_{BOS2} > RS_{BOS3} > RS_{BOS1} ] \]
\[ r_2 = [ RS_{BOS1} > RS_{BOS2} > RS_{BOS3} ] \]
\[ r_3 = [ RS_{BOS1} > RS_{BOS3} > RS_{BOS2} ] \]
\[ r_4 = [ RS_{BOS2} > RS_{BOS1} > RS_{BOS3} ] \]

To measure the degree of variability in the rank achieved by a given stimulus, the frequency of rank achieved by each sound stimulus was scored. For each stimulus \((i)\) within an ensemble of stimuli tested at multiple recording locations \((j)\), a “reliability” score \((R)\) was computed, which reflects the degree to which stimulus rank scores were consistent across recording locations:

\[
\text{Reliability of stimulus } (R_i) = \frac{\sum_{j} (R_{i,j})^2}{j^2}
\]

By this measure, all reliability scores fall in the range \(0 > R \leq 1\). For example, reliability scores for the uniform example above represent perfect repeatability of stimuli \((i_1 - 3)\) rankings across recording sites \((j_1 - 4)\):

<table>
<thead>
<tr>
<th>Stimulus ((i))</th>
<th>possible ranks</th>
<th>ranks achieved</th>
<th>frequency of rank</th>
<th>(R_i)</th>
<th>ensemble mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS1</td>
<td>[ 1 2 3 ]</td>
<td>[ 3 3 3 3 ]</td>
<td>[ 0 0 4 ]</td>
<td>(0^2 + 0^2 + 4^2) / 4^2 = 1</td>
<td>= 1</td>
</tr>
<tr>
<td>BOS2</td>
<td>[ 1 2 3 ]</td>
<td>[ 1 1 1 1 ]</td>
<td>[ 4 0 0 ]</td>
<td>(4^2 + 0^2 + 0^2) / 4^2 = 1</td>
<td></td>
</tr>
<tr>
<td>BOS3</td>
<td>[ 1 2 3 ]</td>
<td>[ 2 2 2 2 ]</td>
<td>[ 0 4 0 ]</td>
<td>(0^2 + 4^2 + 0^2) / 4^2 = 1</td>
<td></td>
</tr>
</tbody>
</table>

These reliability scores indicate that each member of the test stimulus set was consistent with regard to the relative sizes of auditory responses its playback evoked. Note that stimulus reliability scores are not influenced by the particular rank achieved by a stimulus, only by rank consistency. Reliability scores for the more-variable example above reveal that test stimuli varied in the relative strength of evoked auditory responses, with BOS3 less variable than either BOS1 or BOS2:
<table>
<thead>
<tr>
<th>Stimulus (i)</th>
<th>possible ranks</th>
<th>ranks achieved</th>
<th>frequency of rank</th>
<th>$R_i$</th>
<th>ensemble mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS1</td>
<td>[1 2 3]</td>
<td>[3 1 1 2]</td>
<td>[2 1 1]</td>
<td>$\frac{2^2 + 1^2 + 1^2}{4^2} = 0.25$</td>
<td>$= 0.33$</td>
</tr>
<tr>
<td>BOS2</td>
<td>[1 2 3]</td>
<td>[1 2 3 1]</td>
<td>[2 1 1]</td>
<td>$\frac{2^2 + 1^2 + 1^2}{4^2} = 0.25$</td>
<td></td>
</tr>
<tr>
<td>BOS3</td>
<td>[1 2 3]</td>
<td>[2 3 2 3]</td>
<td>[0 2 2]</td>
<td>$\frac{0^2 + 2^2 + 2^2}{4^2} = 0.50$</td>
<td></td>
</tr>
</tbody>
</table>

The lower reliability scores obtained in this example reflect the greater variability in relative stimulus ranking across recording locations. For both awake and anesthetized data, reliability scores were computed for any ensemble of test stimuli used at multiple locations within an individual HVC. To assess the utility of reliability measures, the set of observed reliability scores was compared to those obtained after a uniform shuffling of stimulus ranks, and to those obtained from a model data set of equal size which exhibited perfect repeatability of stimulus ranks. To assess the potential for state differences in auditory processing, test stimuli reliability scores were compared between recordings collected in the awake and anesthetized states.
RESULTS

Study animals sang freely in captivity

After 1 - 2 days of acclimation, male song sparrows originally observed and recorded on territories in their natural habitats sang spontaneously in laboratory acoustic attenuation chambers. Song production varied broadly, ranging from several hundred to several thousand songs sampled per individual over multiple days of recording. All males recorded in the laboratory sang most often in the few hours after the onset of the daylight portion of the light cycle, suggesting that study individuals retained their normal diurnal singing pattern.

HVC auditory responses are elicited in both awake and anesthetized song sparrows

Auditory responses in nucleus HVC were apparent as an elevation in single- or multi-unit spike rate during auditory stimulation. In both awake (Figure 3a) and anesthetized (Figure 3b) animals, auditory responsiveness was broad, with most song stimuli eliciting an elevation in HVC activity throughout the duration of auditory playback. RS magnitude to forward-orientation self-song (BOS) playback did not differ between the two sedative mixtures used (ketamine/xylazine versus diazepam; F1,61 = 1.66, NS); data representing use of these drugs were therefore pooled in subsequent analyses.

Multiple BOS song types from a bird’s own repertoire (BOR) were used as auditory playback stimuli at each recording location. In awake animals, presentation of BOS song stimuli caused, on average, a 44% elevation (e.g., mean RS = 1.44) in neural activity over pre-stimulus levels (N = 32 recording sites in 5 birds; Figure 3c). Following the end of auditory stimulation, HVC activity returned to pre-stimulation levels (average of 1% greater than pre-stimulus levels, NS). In anesthetized animals, presentation of BOS stimuli evoked, on average, a 103% increase in HVC activity over pre-stimulus levels (N = 64 recording sites in 11 birds; mean RS = 2.03; Figure 3c). In contrast to awake animals, this excitatory response during auditory stimulation was typically followed by a period of activity suppression, to a level significantly less than that prior to auditory stimulation (sign test, M = 13.0, p < 0.005). This activity
suppression, reported previously in intracellular (Lewicki 1996) as well as extracellular recordings (Volman 1996), lasted for several seconds and was of small but consistent magnitude (mean: -6.2% after forward-orientation BOS stimuli). Overall, the magnitude of elevation of HVC activity during auditory playback was greater in anesthetized than awake animals (t-test, t_{1,94} = 2.19, p < 0.05; Figure 3c).

Increases in HVC activity during auditory stimulation were accompanied by increased variability of activity (measured as the variance over 10 msec intervals across PSTH; Cardin and Schmidt 2003). In awake animals, the 44% increase in overall activity from pre-stimulus to during-stimulus periods was associated with a 354% increase in the temporal variance of activity, while the 103% elevation of overall activity during stimulation under anesthesia was accompanied by a similar 353% increase in the variability of neural activity during auditory stimulation. The fact that presentation of sound causes the mean and variance of HVC neural activity to increase by different amounts suggests that, in both animal states, auditorially-evoked excitation in this brain region is not achieved via a uniform ‘gain control’ over spontaneous neural activity levels.

**HVC auditory responses were selective for BOS over REV sound stimuli**

Presentation of BOS auditory stimuli in normal, forward orientation reliably evoked robust auditory responses in nucleus HVC of both awake (Figure 4a) and anesthetized animals. To explore the acoustic basis for the generation of these responses, we compared the responses elicited by artificial temporal and spectral variants of normal BOS stimuli, including whole-song reversed (REV) sound stimuli. Both BOS and REV stimuli reliably evoked HVC auditory responses, of broadly variable magnitude. Across all animals and recording sites, BOS auditory responses were greater in overall magnitude (whole-song RS value) than REV responses, a difference that was significant in awake (mean ± SE: 1.47 ± 0.10 vs. 1.16 ± 0.04; F_{1,50} = 5.42, p < 0.03) but not anesthetized recordings (2.07 ± 0.25 vs. 1.67 ± 0.18; F_{1,78} = 0.67, NS; Figure 4b). Nonetheless, assessments of auditory selectivity using the d’ measure, which make use of both the magnitude and trial-to-trial variability of responses in paired comparisons, revealed consistent
selectivity of BOS over REV responses in both awake (mean $d'$ ± SE: 0.34 ± 0.08) and anesthetized (mean $d'$ ± SE: 0.41 ± 0.12) birds (Figure 4c).

To clarify the degree to which individual syllables contribute to overall auditory responsiveness in HVC, a number of experiments also included R-ORD and R-SYL variants of original BOS sound stimuli (see Methods). The magnitude ($RS$) of auditory responses evoked by these different categories of stimuli were significantly different from one another in the awake ($F_{3,80} = 3.85, p < 0.02$) but not the anesthetized ($F_{3,91} = 0.33, \text{NS}$) conditions.

In both awake and anesthetized conditions, BOS song stimuli evoked auditory responses of greater magnitude (as measured by $RS$) than did use of acoustic variants (R-ORD, R-SYLL, REV stimuli; see Methods) of BOS stimuli. In both animal states, R-ORD stimuli evoked auditory responses of greater $RS$ magnitude than did REV or R-SYLL stimuli (Figure 5). BOS and R-ORD stimuli share normal orientation of individual syllable acoustic structure (while differing in syllable sequence order), while the stimulus categories which evoked the smallest auditory responses (REV or R-SYL) present individual syllables in reverse temporal form. These findings suggest that the acoustic features of individual syllables contribute more strongly to HVC auditory responsiveness than does temporal order of song-level spectral content.

**Auditory tuning to multiple BOS: variation within HVC recording locations**

The song types which make up an individual song sparrow’s BOR vary broadly in spectral content (Figure 2b). To assess the stimulus-specific contribution to the variability seen in BOS auditory responses, $RS$ values from passive presentation of multiple BOS stimuli were compared within recording locations. Multiple BOS stimuli were presented while recording at a total of (non-simultaneous) 19 unique nucleus HVC recording locations in awake animals, and at a total of 59 recording locations in the HVC of anesthetized animals. At any given recording location, the magnitude of evoked auditory
response varied among the BOS stimuli tested, in both awake and anesthetized animals (Figure 6). Figure 6a depicts a single example of this type of variation. While each of the four BOS stimuli tested provided reliable auditory drive at this recording location (all RS > 1), the degree of excitation provided varied among the different BOS stimuli presented (compare RS for BOS1 versus BOS3), and for only one stimulus (BOS1) was this the recording location at which the sound stimulus evoked its greatest auditory response.

Above-background auditory responsiveness to one or more of the BOS sound stimuli tested was accompanied by activity suppression to a different BOS sound stimulus at < 15 % of recording locations (3 of 32 recording locations in awake animals; 10 of 63 recording locations in anesthetized HVC). However, there was a significant auditory response to at least one BOS song stimulus in both awake and anesthetized animals at every HVC recording location (Figure 6b).

As used here, RS values represent an overall, “whole-song” level of auditory responsiveness. Because many auditory responses included phasic elements (e.g., Figure 4), it was necessary to verify that this whole-song measure accurately reflected more-transient elements of neural auditory responsiveness. To do so, auditory responses were scored to identify the smallest (RSmin) and largest (RSmax) auditory responses for each recording location. Then, responses gauged to be the smallest and largest at each recording location on the basis of whole-song RS values were compared with regard to measures of auditory excitation calculated from 10 msec PSTH bin values. By multiple measures of comparison (mean, maximum and variance of neural excitation during auditory presentation, as measured over 10 msec PSTH bin values), auditory responses assessed via whole-song RS as being “largest” or smallest” within recording locations were statistically separable in both awake (paired t-test, 19 recording locations: t1,18 = 3.13 – 4.46, all p < 0.01) and anesthetized (59 recording locations: t1,58 = 6.12 – 6.62, all p < 0.0001). These results suggest that whole-song RS rank values accurately represent the relative excitation
evoked by BOS song stimuli, despite the existence of rapid modulations to the neural excitation provided by individual song stimuli.

Measures of $R_{Smin}$ and $R_{Smax}$ were then combined to provide an estimate of the neural dynamic range ($R_{Srange} = R_{Smax} / R_{Smin}$) for BOS auditory responses within individual HVC recording locations. Awake and anesthetized recording locations differed broadly in overall auditory responsiveness; recording location $R_{Smin}$, $R_{Smax}$, and $R_{Srange}$ were greater in anesthetized than in awake animals ($R_{Smin}$: $t_{1,76} = 2.97, p < 0.005$; $R_{Smax}$: $t_{1,76} = 3.11, p < 0.005$; $R_{Srange}$: $t_{1,76} = 1.66, p = 0.05$). In awake animals, HVC within-location $R_{Srange}$ averaged 1.3-fold (average $R_{Smin}$: 1.19; average $R_{Smax}$: 1.55; Figure 6b). Under anesthesia, HVC within-location BOS auditory dynamic range averaged 1.4-fold (average $R_{Smin}$: 1.53; average $R_{Smax}$: 2.19). This within-location variability in auditory responsiveness to different songs of the BOR suggests that individual neurons and/or recording sites are specifically tuned to acoustic features present in a subset of songs in the BOR.

**Auditory tuning to BOS: variation among HVC recording locations**

HVC auditory representation of BOS in the single-songed zebra finch is known to be broadly similar across the entire extent of the nucleus under anesthesia (Sutter and Margoliash 1994). This finding suggests a uniformity of auditory tuning across HVC to match the spectral characteristics of an individual’s BOS. However, for each song sparrow individual studied in these experiments, the multiple BOS songs used as test stimuli represent only a subset of the song types within each individual’s BOR. BOS song types vary widely in spectral characteristics (Figure 2b), yet all BOS sound stimuli reliably evoked auditory responses in nucleus HVC in both anesthetized and awake song sparrows. This suggests that individual song sparrow BOS stimuli may drive some HVC recording locations to a greater degree than others.
To test whether HVC auditory responsiveness varies spatially, we compared the $RS$ values obtained from passive presentation of multiple songs from an individual bird’s repertoire at multiple locations within that individual’s nucleus HVC. The ability of a given sound stimulus to elicit auditory responses in nucleus HVC varied broadly with location in the nucleus. As example, the $RS$ values generated by presentation of a single BOS stimulus at four different recording locations within nucleus HVC are presented in Figure 7a. This stimulus generated the maximum observed $RS$ (upper tick mark in each bar) at only one of the four sites tested (location #4), and was the stimulus providing the least auditory drive (lower tick mark in each bar) at none of the four recording locations.

Population data, summarized by BOS sound stimulus, are presented in Figure 7b. On average, $RS$ values obtained from presentation of a single sound stimulus at multiple locations within nucleus HVC spanned the range of 1.09 – 1.66 in awake animals, and from 1.20 – 2.94 in anesthetized animals. These results provide estimates for the auditory dynamic range among recording locations within nucleus HVC to be approximately 1.5-fold in awake recordings and 2.5-fold in anesthetized recordings, a finding which represents significant spatial variability of auditory responsiveness. These results suggests that a global form of auditory representation, as described in the zebra finch (Sutter and Margoliash 1994), does not obtain in the song sparrow. Estimates for stimulus $RS_{max}$ did not increase with the number of recording locations sampled in either awake ($F_{1,11} = 0.95$, NS) or anesthetized ($F_{1,35} = 0.03$, NS) states, suggesting an adequate sampling of among-recording location variation in $RS_{BOS}$.

**Auditory responses reflect a combination of stimulus properties and local tuning**

Both sound stimulus acoustic properties and local auditory tuning within nucleus HVC contribute to auditory responsiveness (Figure 8). The combination of location-specific and stimulus-specific variability slightly exceeds that observed in both awake ($RS_{BOS}$ observed: 1.8x; combined estimate: $1.5x * 1.3x = 1.95x$) and anesthetized ($RS_{BOS}$ observed: 3.4x; combined estimate: $2.5x * 1.4x = 3.50x$) recordings, suggesting a slight amount of covariance between stimulus and location-specific measurements. In both
the awake and anesthetized conditions, the relative contribution of sound stimulus characteristics to the variability of auditory response magnitudes is less than that contributed by the auditory tuning of the given recording location (awake: $RS_{range}$ of location = 1.5-fold, $RS_{range}$ of stimulus = 1.3-fold; anesthetized: $RS_{range}$ of location = 2.5-fold, $RS_{range}$ of stimulus = 1.4-fold; Figure 8). Thus song repertoire auditory representation within HVC is neither completely spatially distributed, nor globally uniform. Given that song types may share common syllables (Podos et al. 1992), local HVC tuning to individual syllable types, as described by Mooney (Mooney et al. 2001), may be sufficient to generate the representation we describe here.

**Spatial representation of song repertoires within HVC: repeatability of stimulus ranks**

The complexity and variability of song sparrow song types, in conjunction with the spatial variability of HVC auditory tuning, suggest a more complex auditory representation of self-song than has been described for the zebra finch. To assess the spatial organization of auditory representation within this brain region, we measured the repeatability with which stimulus ensembles (groups of 2-4 unique BOS song types) were represented across recording sites. For those recording sites at which 2 or more BOS forward-direction songs were used as auditory playback stimuli, the test stimuli were ranked according to the magnitude of evoked $RS$. The $RS$ ranks of each member of the stimulus ensemble were then compared across each recording site at which the ensemble was tested. These ranks were then used to produce reliability scores for each sound stimulus which represent the consistency of auditory responses across HVC recording sites (see Methods).

To demonstrate the utility of ranking the members of stimulus ensembles in this way, stimulus ranks achieved from presentation of two equally-sized stimulus sets (one each from awake and anesthetized HVC recordings, from different birds) are depicted in Figure 9a. Stimulus ensembles used in these examples consisted of 4 BOS song stimuli, tested at each of 5 recording locations within nucleus HVC. At each recording location, the BOS stimuli were ranked with regard to the relative magnitude of the
evoked auditory response (e.g., ranked within recording location by RS value). As depicted in Figure 9a, variation across recording locations in the relative size of the evoked auditory response to a given stimulus is substantial in both awake and anesthetized recordings. However, this variation is not structured equivalently under awake and anesthetized conditions. Figure 9b depicts the stimulus ranks and reliability scores achieved by the stimuli represented in Figure 9a. The four sound stimuli tested in the awake animal achieved reliability scores from 0.36 – 1.00, and no single sound stimulus achieved all possible rank positions within the ensemble. Reliability scores for the four sound stimuli tested under anesthesia ranged from 0.28 – 0.52, and two of the four sound stimuli achieved both the lowest and highest possible rank positions within the ensemble. Ensemble mean reliability is lower in the anesthetized (0.38) than the awake recording (0.60), suggesting that stimulus rank repeatability varies with animal state.

In awake animals, a total of 12 unique BOS stimuli were tested in 6 stimulus ensembles consisting of 2-4 stimuli each (individual BOS stimuli could be used in > 1 ensemble; ensembles comprised 17 BOS stimuli in total) across 18 HVC recording locations in 4 birds. Ensemble recordings in anesthetized animals comprised 38 BOS stimuli (36 unique) in 12 ensembles across 56 HVC recording locations in 10 birds.

To evaluate the likelihood of observed reliability scores, scores obtained during recordings in both awake and anesthetized conditions were compared to two permutations of the observed data. First, observed scores were compared to data sets of matching size which represented perfect stimulus reliability (e.g., all reliability scores = 1). Second, observed scores were compared to those obtained from a shuffled version of the observed data (reliability scores computed after random ordering of stimulus rankings within each recording location). Stimulus ensemble reliability scores differed markedly across treatments ($F_{5,48} = 21.45, p < 0.001$). The reliability scores obtained in both awake and anesthetized conditions were significantly lower than 1 (= “perfect” reliability; statistics not shown). Reliability scores from nucleus
HVC recordings in awake animals were greater than those obtained from a random shuffling of the observed data (Figure 9c; p < 0.05), while scores obtained under anesthesia did not differ from a shuffled version of scores obtained during anesthetized recordings (Figure 9c; NS).

The example described above and depicted in Figure 9a-b is representative of the overall difference in stimulus reliability seen in awake and anesthetized animals. Reliability scores summarized across all members of sound stimulus ensembles reveal that stimulus reliability is significantly greater in awake nucleus HVC than under anesthesia (Figure 9c; \( t_{1,53} = 1.92, p < 0.05 \)). High- and low-ranking stimuli did not differ in overall power spectra (not shown), which suggests that response magnitudes reflect a matching of stimulus and local neuronal tuning characteristics, rather than an overall bias for specific spectral features. These comparisons demonstrate that the spatial organization of stimulus auditory representation within HVC differs between animal states, a finding which suggests caution in the interpretation of awake animal neural phenomena on the basis of data collected under anesthesia. These results also suggest that individual song-level selectivity measures alone may be inadequate to fully describe the representation of song repertoires drawn from a large, shared pool of song syllables.
DISCUSSION

We have presented the first description of song repertoire auditory representation in the song system nucleus HVC of awake, behaving birds, and demonstrated robust and selective auditory responses to self-songs in the awake animal. Auditory responses within nucleus HVC are broadly distributed and spatially variable in form and magnitude. Multiple lines of evidence suggest that auditory responsiveness results from a correspondence of stimulus acoustic features and the auditory tuning properties of individual neurons or local subpopulations thereof. We also have demonstrated that relative auditory responsiveness to different BOS song stimuli is more spatially invariant in awake than in anesthetized nucleus HVC, a finding which suggests that the mechanisms for generation of auditory excitation in this brain region is not conserved across all animal states. Together, these data suggest that the neural strategies by which the single BOS song of the zebra finch in represented within nucleus HVC may be insufficient for the auditory representation of more complex song repertoires.

Auditory excitation within nucleus HVC

Presentation of BOS song stimuli reliably evoked excitation within nucleus HVC in both awake and anesthetized animals. As reported above (see Results), the magnitude of BOS-evoked excitation (expressed as percentage increase over spontaneous activity) was greater in anesthetized than awake animals, and yet the variance in auditory neural activity increased by an equivalent amount in both animal states. That the temporal variability of neural activity during auditory stimulation changes equivalently between states may reflect the basic temporal structure of song sparrow song stimuli, as all stimuli comprised distinct syllables separated by silent intervals up to 0.3 s in duration (Figure 2b). Alternatively, intrinsic dynamics of the HVC excitatory response may act in common across both animal states to provide this common structure to the increase in network excitation; direct tests of either possibility have not been performed.
Species differences in HVC auditory responsiveness

HVC auditory processing in song sparrows is notably distinct from that of the zebra finch in the relatively modest differences in auditory response magnitudes between awake and anesthetized conditions, and the presence of robust and stable auditory responsiveness in the awake animal. Overall, the fact that robust auditory selectivity in HVC is apparent in awake song sparrow, but not zebra finch, suggests a fundamentally different utility to the audition of song in these species. These species differ strongly in both their territorial use of song as well as the level of song sharing among “neighbors” (Arcese et al. 2003; Zann 1996); whether the neural auditory differences reported here relate to either or both of these behavioral aspects of song use is yet to be determined. Behavioral and/or ecological differences among species may explain some of the differences, as passive audition of a BOS song is an unnatural event for a zebra finch, but represents a potentially relevant signal from a song type-matching song sparrow conspecific. Passive presentation of song stimuli is insufficient, however, to fully study inter-individual recognition and signaling in this (and other) species, as visual signals, physical interaction, and non-song vocalizations also facilitate intraspecific communication. Further tests, incorporating behavioral interactions in ethologically relevant settings, will be necessary to determine if HVC auditory responses are associated with the robust behavioral response to conspecific song in this species.

Song sparrow HVC auditory responses are consistently selective for BOS song

Unlike the zebra finch (Cardin and Schmidt 2003), the degree of song sparrow HVC BOS-REV selectivity did not differ between states ($t_{1,35} = 0.47$, NS). In the zebra finch, $d'$ auditory selectivity values for awake zebra finch HVC averaged $0.05 \pm 3.5$ (Cardin and Schmidt 2003), and were broadly distributed around 0, demonstrating an overall lack of auditory selectivity in HVC of awake birds. While the overall magnitude of song sparrow HVC BOS-REV selectivity is less than that reported from studies in anesthetized zebra finch nucleus HVC (Cardin and Schmidt 2003; Lewicki and Arthur 1996; Theunissen et al. 2004; Theunissen and Doupe 1998), neural population-level BOS-REV selectivity in HVC is apparent in both awake and anesthetized song sparrows.
Overall, song sparrow HVC selectivity indices for BOS versus REV (Figure 4) are modest in comparison to anesthetized zebra finch HVC and greater than those reported for awake zebra finch HVC. Relatively few song sparrow HVC recording locations meet the threshold of \(d’ > 0.5\) or 1.0 which has been used to categorize anesthetized zebra finch recording locations as “selective” or “non-selective” (Theunissen and Doupe 1998). However, the use of a single, positive threshold value in this way is arbitrary; no use of such a threshold is indicated in the original derivation of this metric (Green and Swets 1966). As described by Janata and Margoliash (Janata and Margoliash 1999), the \(d’\) metric is symmetric about 0. Selectivity indices representing single recording locations are free to vary in either the positive or negative direction, which suggests that a population-level description of auditory selectivity is best revealed by comparing observed indices to 0. While the magnitudes of selectivity indices observed here are smaller than those commonly reported for the zebra finch, the majority of indices are \(> 0\) (Figure 4). Thus, auditory selectivity may be less extreme than reported for the zebra finch, but is apparent nonetheless. Interpretation of auditory selectivity remains difficult, as we cannot assess how the neural utility of an HVC BOS auditory response differs from that to a different song type, or to conspecific song. Similarly, we cannot, at present, relate HVC neural responses to associated behavioral responses; it is known from other systems that very modest neural sensory response differences may translate non-linearly into sharply demarcated behavioral decisions (e.g., Asaad et al. 2000; Kim and Shadlen 1999).

**Organization of BOS auditory representation within HVC**

HVC auditory responses to the single song of the zebra finch are relatively consistent across multiple recording locations within anesthetized nucleus HVC (Sutter and Margoliash 1994). However, the auditory representation of song repertoires may require something other than a global strategy of single-BOS auditory representation, as song sparrow HVC auditory responsiveness differs from that of the zebra finch in several important ways. First, auditory responses to BOR songs in song sparrows are not individually or collectively uniform across recording sites under anesthesia (Figure 7). In anesthetized song sparrow HVC, the magnitude of auditory responses to BOR stimuli vary more than three-fold
among recording locations and stimulus types (Figure 8), with both sound stimulus and local tuning contributing substantially to auditory response variation. That individual recording sites vary in the degree to which they can be driven by a given sound stimulus suggests that the neural valence of a particular stimulus is dependent upon recording location. Thus, song repertoire auditory representation in this species is global (Sutter and Margoliash 1994), or non-specific, only in the sense of broad auditory responsiveness across nucleus HVC. Each of an individual’s BOR song types reliably evokes HVC auditory responses, although the stimulus which does so best varies by recording location. No single BOR song type, or “super-stimulus”, could be identified in each animal as consistently providing the greatest degree of auditory drive.

Second, auditory response reliability indices (Figure 9) suggest an organized spatial structure to awake, but not anesthetized, HVC auditory responsiveness in this species, a finding which suggests that the network dynamics which generate auditorially-evoked excitation within HVC are not consistent across animal states. How these species differences relate to the shared-syllable organization of the multiple songs of the song sparrow repertoire remains to be determined. It is possible that syllables, rather than song-types, are discretely represented within nucleus HVC (Mooney et al. 2001), and syllable sharing among song sparrow song types itself lends consistency to auditory representation across this brain region. The mechanism by which this relationship would change between the awake and anesthetized states remains elusive, although the reduced repeatability of stimulus ensembles under anesthesia is consistent with the relatively reduced selectivity and enhanced auditory responsiveness which characterizes HVC neural recordings in this state.

**Conclusion**

Selective and acoustic-features specific auditory responses are a hallmark of the avian song system, yet their existence remains enigmatic, as they have rarely been described in awake, behaving animals. We demonstrate here that robust and selective auditory responses exist in the avian song system nucleus HVC
in awake animals, in a species with known behavioral use of song in intraspecific communication. We further demonstrate that auditory responsiveness results from a combination of local neuronal tuning with sound stimulus acoustic features, and provide evidence for a state-dependent, spatial organization to the auditory representation of song repertoires within this brain region. Current models of song learning and maintenance posit a role for the passage of auditory information from HVC into the basal ganglia-like anterior forebrain, and the present study provides strong evidence for the availability of auditory information in the nucleus HVC of awake animals. While HVC auditory responsiveness during passive song presentation may differ functionally from the real-time auditory feedback potentially available during singing, the BOS specific auditory responses described here potentially reflect the importance of such feedback for song learning and maintenance. These studies provide a gateway for further investigation of song perception in awake, behaving animals, and point to the need for careful consideration of the behavioral use of song in the design of experiments on the neural basis of song processing.
ACKNOWLEDGEMENTS

We thank Dima Rinberg for technical assistance and Alan Gelperin for research support. We also thank Josh Gold for helpful discussions, and Robin Ashmore, members of the Schmidt laboratory, and anonymous reviewers for comments. This research was supported by an NIH postdoctoral National Research Service Award (DC-00125) to P.M. Nealen. M.F. Schmidt was supported by a Sloan Foundation Award and NIH-National Institute of Deafness and Other Communication Disorders grant RO1-DC-6453-1. Present address for P.M. Nealen: Department of Biology, Indiana University of Pennsylvania, 114 Weyandt Hall, 975 Oakland Avenue, Indiana, PA 15705-1001.
REFERENCES


Cardin JA and Schmidt MF. Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. *Journal of Neurophysiology* 90: 2884-2899, 2003.


Sutter ML and Margoliash D. Global synchronous response to autogenous song in zebra finch HVc.


Volman SF. Quantitative assessment of song-selectivity in the zebra finch high vocal center. *Journal of Comparative Physiology (A) - Sensory, Neural, and Behavioral Physiology* 0178: 849-862, 1996.

FIGURE LEGENDS

Figure 1. Anatomical substrates of the passerine song system. a) Cresyl violet-stained tissue section of song sparrow forebrain, depicting the locations of nucleus HVC and its premotor target, nucleus RA. Both nuclei are readily apparent in histological examination. b) Schematic representation of the principal neural pathways related to the learning, perception, and production of passerine bird song (schematic represents only one forebrain hemisphere; these circuits are mirrored in the left and right hemispheres). Auditory information from both self and non-self vocalizations enters the song system via direct projections from the auditory thalamus into Field L and then Nucleus interfacialis (Nif). Physiological recordings were conducted here in nucleus HVC, which exhibits robust activity during both the audition and production of vocalizations. HVC receives auditory information from Nif, and provides premotor instruction to nucleus RA. HVC and RA are both interconnected with a basal ganglia circuit (X, DLM and LMAN) thought to be involved in song auditory processing and song learning.

Anatomical abbreviations:

CLM - caudolateral mesopallium
CMM - caudomedial mesopallium
DLM - nucleus dorsolateralis anterior, pars medialis
DM – dorsomedial portion of the nucleus intercollicularis
L1, L2a, L2b, L3 – subdivisions of Field L
LMAN - lateral magnocellular nucleus of the anterior nidopallium
NCM - caudal medial nidopallium
Nif - nucleus interface of the nidopallium
nXIIIs – tracheosyringeal portion of the hypoglossal nucleus
OV core, OV shell, OVm – subdivisions of the nucleus ovoidalis
RA - robust nucleus of the arcopallium
Respir. – brainstem respiratory network
Figure 1
Nealen

Page 33 of 49
Figure 2. Recording and characterization of song sparrow song. a) Schematic of sound recording scheme. Individual male song sparrows were located in the wild on natural territories, and their songs recorded over several hours. From field and laboratory recordings of focal individuals, exemplars of four song types were selected as auditory playback stimuli (“bird’s own song”, BOS1-4). Songs from other song sparrow males were recorded both within and external to focal male home populations, providing a library of conspecific (CON) song stimuli. b) Sonographic representations of part of the test stimulus set for one song sparrow male. BOS auditory playback stimuli were selected to reflect a diversity of syllable types and sequences. Syllable sharing among songs within a repertoire is common, and individual syllables occur with different repetition number and positioning among songs in the repertoire. Inset bars indicate the occurrence of syllable sharing among song types within this individual’s repertoire, and among individuals.
Figure 2

a) Schematic representation of different vocalizations recorded and their corresponding territories:

- **focal male on natural territory**
- **neighbor male on adjoining territory**
- **> 50 km**
- **conspecific male from distant population**

**Record:**
- BOS1-4
- CON

b) Waveform plots showing the frequency and time analysis for different vocalizations:

- **BOS1**
- **BOS2**
- **BOS3**
- **BOS4**
- **CON**

Frequency (kHz) and Time (sec) are indicated for each waveform.
Figure 3. Auditory response properties in nucleus HVC. Song-evoked auditory responses in song sparrow HVC were apparent in extra-cellular recordings as an elevation of spike rate during stimulus presentation. In both awake (a) and anesthetized (b) animals, presentation of sound stimuli robustly recruited HVC neurons. The top portion of each panel a) and b) depicts a multi-unit spike raster representing responses over repeated trials, which are summed to obtain the peri-stimulus time histogram (PSTH) underneath each raster. The BOS stimuli used in these presentations are depicted in waveform at panel bottoms. c) HVC population data summarized by recording site reveal that the degree of elevation of HVC activity during sound presentation, as well as the resultant post-stimulus depression of HVC activity, differ between the awake and anesthetized states. In both awake and anesthetized animals, passive playback of BOS robustly recruits HVC activity (1awake, sign test, N = 32; 2anesthesia, sign test, N = 64). The magnitude of sound-stimulus induced HVC excitation is greater under anesthesia than in awake animals (3unpaired t-test); activity suppression following stimulus offset is observed in anesthetized, but not awake, HVC (4unpaired t-test).
During After

a) Awake

b) Anesthetized

c) 

Response strength (RS) vs % change in activity:

- During auditory playback
  - Rs = 16.0, p < 0.0001
  - t = 2.19, p < 0.05

- After auditory playback
  - Rs = 13.0, p < 0.005
  - t = 1.82, p < 0.05
Figure 4. State-independent BOS-REV selectivity of song sparrow HVC auditory responsiveness.

Acoustic manipulation of BOS stimuli leads to an overall decrement in the magnitude of nucleus HVC auditory responsiveness. a) Multi-unit auditory responses obtained from nucleus HVC of an awake song sparrow during interleaved playback of forward- (BOS) and reverse-orientation (REV) versions of a single BOS sound stimulus, chosen from the bird’s own repertoire (BOR). Shown here are raster plots (upper; n = 70 trials) and summary PSTH’s (middle) representing responses to a single BOS stimulus (oscillogram; lower) presented in both forward (BOS; left) and reverse (REV; right) orientation. Response to the BOS stimulus is phasic in form, such that the overall mean level of activity during the course of the stimulus is only slightly greater than that during the REV stimulus. Note that the reliable and temporally precise responses to specific syllables in the BOS orientation are lost during REV presentation.

b) Comparison of auditory response strength measures to BOS and REV stimuli in awake and anesthetized nucleus HVC. The response strength ($R_S$) evoked by BOS stimuli was greater than that evoked by REV stimuli in awake, but not anesthetized, animals. The number of recording locations sampled is displayed in each bar.

c) $d'$ selectivity indices (Green and Swets 1966) depicting the degree to which auditory responsiveness was biased toward BOS over REV sound stimuli. In both awake and anesthetized animals, the distribution of selectivity indices was significantly greater than zero (mean of each distribution indicated), while selectivity indices did not differ between the awake and anesthetized states ($F_{1,35} = 0.22$, NS). Inset numbers depict the number of recording locations sampled for each state.
Figure 4
Nealen
Figure 5. Manipulation of BOS stimuli decreases HVC auditory responsiveness. Acoustic manipulation of normal song stimuli caused a decrement in the magnitude of evoked auditory responses. Song stimuli subject to syllable reversal and re-ordering were tested at recording locations across HVC in both awake and anesthetized animals. Plotted here for each stimulus category are the mean response strength ($RS$) values achieved by presentation of the various stimuli [forward-orientation song (BOS), syllable reverse-order (R-ORD), reverse-orientation of syllables (R-SYLL) and reverse song (REV)] at each recording location in both awake (○) and anesthetized animals (●). Distribution means are indicated by a horizontal bar in each distribution. Inset numbers depict the number of recording locations at which each category of stimulus was tested.
Figure 5
Nealen
Figure 6. BOS song types vary in the degree of excitation they provide within individual recording locations in HVC. At any given recording location, multiple BOS song types from the BOR were used as playback stimuli. a) Example from a single recording location in awake HVC. Each of the four different BOS song stimuli tested at this location elicited above-background auditory responses (grey bars), although the stimuli differed in the degree to which they excited nucleus HVC during sound presentation. Inset bars depict the maximum and minimum $RS$ values achieved during presentation of an individual stimulus across all of the HVC locations at which it was tested. In this example, only one of the song stimuli (BOS1) achieves maximum $RS$ at this particular recording location, while none of the four stimuli achieve their minimum $RS$ at this recording location. b and c) Population data for all HVC recording locations tested in awake (b) and anesthetized (c) animals. Each column in the plots represents a unique HVC recording location; each symbol within a column represents the $RS$ value achieved from presentation of a distinct BOS song type at that location. The single-location example depicted in panel (a) is highlighted in grey in the middle of panel (b). A total of 32 unique HVC recording locations were tested in 5 awake animals, and 63 different HVC recording locations were tested in 12 anesthetized animals. For those locations at which 2 or more different BOS song stimuli were tested, the $RS_{\text{min}}$ and $RS_{\text{max}}$ were scored for that location; population average $RS_{\text{min}}$ and $RS_{\text{max}}$ are plotted as horizontal lines. The data collected from each individual bird is grouped together here for presentation; animal numbers and inset lines at panel bottoms span the recording locations sampled from each individual bird.
4 BOS stimuli tested at 1 HVC location

b) Awake

RS

BOS1 BOS2 BOS3 BOS4

RS_{max} = 1.55

RS_{range} = 1.3x

RS_{min} = 1.19

Unique HVC recording locations

b) Anesthetized

RS

BOS1 BOS2 BOS3 BOS4

RS_{max} = 2.19

RS_{range} = 1.4x

RS_{min} = 1.53

Figure 6
Nealen
Figure 7. Auditory responsiveness to a given BOS song stimulus varies across HVC recording locations. The magnitude of auditory response evoked by a BOS sound stimulus varied among the HVC recording locations tested. Individual sound stimuli were presented while recording at a variable number of locations within nucleus HVC.  

(a) In this example, a single BOS song type was used as a playback stimulus at four different HVC recording locations in an awake bird. The excitatory drive provided by this stimulus varied among the locations tested. Inset bars depict the maximum and minimum RS values achieved during presentation of any stimulus at each of the four recording locations. In this example, this single BOS song stimulus provided the most auditory drive of any stimulus tested at recording location 4 only, and provided the least auditory drive at none of the four locations tested.  

(b and c) Population data for all unique song sparrow BOS songs used as test stimuli in awake (b) and anesthetized (c) animals. Each column in the plots represents a unique BOS song stimulus; each symbol within a column represents the RS value achieved from presentation of that BOS stimulus at a given HVC recording location. The single-location example depicted in panel (a) is highlighted in grey in the middle of panel (b). A total of 17 unique BOS sound stimuli were tested in 5 awake animals, and 41 BOS sound stimuli were tested in 11 anesthetized animals. For those stimuli which were used as test stimuli at 2 or more HVC recording locations, the $R_{S_{\text{min}}}$ and $R_{S_{\text{max}}}$ were scored for that stimulus; population average $R_{S_{\text{min}}}$ and $R_{S_{\text{max}}}$ are plotted as horizontal lines. The data collected from each individual bird is grouped together here for presentation; animal numbers and inset lines at panel bottoms span recordings collected from the same individual.
a) 1 BOS stimulus tested at 4 HVC locations

b) Awake

RS

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1.41 & 1.17 & \text{-} & \text{-} \\
\end{array}
\]

RS \text{ max} = 1.66

RS \text{ range} = 1.5x

RS \text{ min} = 1.09

b) Anesthetized

RS

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
0.8 & 1.0 & 1.2 & 1.4 & 1.6 & 1.8 & 2.0 & 2.2 & 2.4 & 2.6 & 2.8 \\
\end{array}
\]

RS \text{ max} = 2.94

RS \text{ range} = 2.5x

RS \text{ min} = 1.20
Figure 8. Stimulus and recording location contributions to auditory responsiveness in song sparrow HVC.

a) Across all song types and recording locations, the multiple BOS song types of a BOR yielded a range of HVC auditory responsiveness ($RS_{range}$) of approximately 1.8x in awake song sparrows (average $RS_{min}$ of 0.97 and $RS_{max}$ of 1.73) and approximately 3.4x in anesthetized animals (average $RS_{min}$ of 1.00 and $RS_{max}$ of 3.53). Assessment of recording location- and stimulus-specific contributions to this observed variability (see Results) suggests that recording location contributes more to the observed variability in $RS_{BOS}$ than do stimulus features, in both awake (54% versus 46% of total $RS_{BOS}$ variability) and anesthetized (64% versus 36% of total $RS_{BOS}$ variability) birds. b) Variability in $RS_{BOS}$ among recording locations suggests a spatial segregation of auditory tuning across HVC, while variability of $RS_{BOS}$ is indicative of the specificity of auditory tuning within any one location. Auditory representation of song repertoires within song sparrow nucleus HVC is achieved by a combination of tuning schemes, with spatial variability within HVC being the predominantly structuring agent. In contrast, the global scheme of auditory responsiveness reported for the zebra finch (Sutter and Margoliash 1994) suggests that auditory representation in this species may be achieved via a common auditory tuning across all of the nucleus HVC.
a) Observed variation in $RS_{pos}$:
- Awake: 1.8x
- Anesth: 3.4x

Relative contributions of:
- HVC recording location
  - Awake: 1.5x (54%)
  - Anesth: 2.5x (84%)
- Stimulus acoustic features
  - Awake: 1.3x (46%)
  - Anesth: 1.4x (36%)

b) Spatially discrete HVC tuning to repertoire components
- RS variance entirely due to local auditory tuning
- Local tuning and stimulus features contribute equally to RS variance

Homogenous auditory tuning across all of HVC
- RS variance entirely due to stimulus acoustic features
Figure 9. State-dependent repeatability of stimulus ensemble responses across nucleus HVC. The reliability with which stimulus ensembles are represented within nucleus HVC is greater in the awake than in the anesthetized animal. a) Upper plot: In this example, 4 different BOS songs were used as auditory test stimuli at 5 different recording locations in nucleus HVC of an awake male song sparrow. Within each recording location, the song stimuli were ranked (from 1 – 4) with regard to the magnitude of their evoked auditory response (as measured by the whole-song RS value). As revealed in these rankings, variability among recording locations within nucleus HVC in the relative representation of sound stimuli is considerable. Lower plot: relative rankings achieved from auditory tests of a set of 4 different BOS song stimuli at 5 recording locations in the nucleus HVC of a different, but anesthetized male song sparrow. As above, variability across recording locations in the relative magnitude of the auditory response evoked by a given sound stimulus is considerable. b) Reliability scores for the sound stimuli represented in panel a. The rank achieved for each sound stimulus is tallied across recording locations, and an overall “reliability” score for the rank of each stimulus is computed as shown. These scores reveal that, despite the overt similarity in the disparity of rankings achieved in awake and anesthetized states (panel a), reliability of the stimuli depicted in panel (a) is greater in the awake (mean reliability: 0.60) than in the anesthetized (mean reliability: 0.38) animal. c) Summary data for all stimulus ensembles tested. In simultaneous comparisons, reliability scores obtained in both awake and anesthetized animals were significantly less than 1 (= “perfect” reliability), but only those scores from awake animals were greater than chance (see Results).
HVC recording location (j)

<table>
<thead>
<tr>
<th>Response rank (R_{i,j})</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

BOS sound stimulus (i)

- Awake
  - BOS 1
  - BOS 2
  - BOS 3
  - BOS 4
- Anesthetized
  - BOS 1
  - BOS 2
  - BOS 3
  - BOS 4

b) BOS sound stimulus (i) | Response ranks (R_{i,j}) achieved | Reliability score $\sum_{j} (R_{i,j})^2$ | Ensemble mean
--- | --- | --- | ---
BOS 1 | [0 3 0 2] | 0.52 | 0.60
BOS 2 | [0 0 3 2] | 0.52 |
BOS 3 | [5 0 0 0] | 1.00 |
BOS 4 | [0 2 2 1] | 0.36 |

- Awake
- Anesthetized

Figure 9

Nealen