Published in final edited form as:

J Comp Physiol A Neuroethol Sens Neural Behav Physiol. 2015 December ; 201(12): 1103–1114. doi: 10.1007/s00359-015-1037-0.

Temporal and rate code analysis of responses to low-frequency components in the bird's own song by song system neurons

Makoto Fukushima¹, Peter L. Rauske², and Daniel Margoliash^{1,2}

¹Department of Psychology, University of Chicago, Chicago IL 60637 U.S.A.

²Department of Organismal Biology and Anatomy, University of Chicago, Chicago IL 60637 U.S.A.

Abstract

Auditory feedback (AF) plays a critical role in vocal learning. Previous studies in songbirds suggest that low-frequency (< ~1 kHz) components may be salient cues in AF. We explored this with auditory stimuli including the bird's own song (BOS) and BOS variants with increased relative power at low frequencies (LBOS). We recorded single units from BOS-selective neurons in two forebrain nuclei (HVC and Area X) in anesthetized zebra finches. Song-evoked responses were analyzed based on both rate (spike counts) and temporal coding of spike trains. The BOS and LBOS tended to evoke similar spike-count responses in substantially overlapping populations of neurons in both HVC and Area X. Analysis of spike patterns demonstrated temporal coding information that discriminated among the BOS and LBOS stimuli significantly better than spike counts in the majority of HVC (94%) and Area X (85%) neurons. HVC neurons contained more and a broader range of temporal coding information to discriminate among the stimuli, than Area X neurons. These results are consistent with a potential role of temporal coding in differentiating in the spectral components of the BOS in HVC and Area X neurons.

Keywords

zebra finch; auditory feedback; temporal coding; HVC

Introduction

Avian vocal learning and adult song perception are model systems for attributes of human speech acquisition, production, and perception (Doupe and Kuhl 1999; Gentner et al. 2006; Fehér et al. 2009; Margoliash and Schmidt 2010; Lipkind et al. 2013). Auditory feedback (AF) is required for speech development (Oller and Eilers 1988) and song development in songbirds (Konishi 2004). AF is necessary to maintain adult vocalizations in humans, although variations occur across individuals (Borden 1979; Waldstein 1990). AF is also necessary for adult song maintenance in some, or perhaps many, songbird species

(Nottebohm et al. 1986; Nordeen and Nordeen 1992). Previous studies in songbirds have investigated the behavior and neurobiology underlying AF processing, demonstrating the behavioral salience of the low-frequency component of AF. Experimental induction of high frequency hearing loss (above 1.5 kHz) had no effect on adult song maintenance (Woolley and Rubel 1999) over the time period when complete deafening would have resulted in a degraded song (Woolley and Rubel 1997; Okanoya and Yamaguchi 1997). Although low-frequency power below 1 kHz is relatively suppressed in airborne song recordings, sounds generated by the syrinx include more power at low frequencies (Fee et al. 1998; Goller and Daley 2001; Jensen et al. 2007). Birds may have access to such low frequencies through bone conduction (Fukushima and Margoliash 2015).

Song system neurons express highly selective auditory responses to the bird's own song (BOS) in anesthetized or sleeping birds (Margoliash 1983; Doupe and Konishi 1991; Margoliash and Fortune 1992; Dave et al. 1998). This BOS-selective response develops during song sensorimotor learning and closely tracks the song of the developing bird (Volman 1993; Doupe 1997; Nick and Konishi 2005). As the low-frequency power of the BOS might be a salient cue in AF, here we tested the hypothesis that the low-frequency component of the BOS could evoke auditory responses in song system neurons, similar to that of the original BOS. Therefore, we used a version of the BOS (low-frequency BOS, LBOS) with a power spectrum that was skewed to decrease power in the high-frequency (>1 kHz) range, thus increasing the relative power of the BOS at frequencies below 1 kHz. We then recorded auditory evoked spiking activity during playback of the BOS and LBOS from neurons in the sensorimotor nucleus HVC, and the basal ganglia component of the song system, Area X, in anesthetized zebra finches. We evaluated the temporal coding properties of the neurons, and compared differences in properties of HVC and Area X neurons, both of which are largely unexplored features of the BOS response.

Materials and Methods

Animals

All animal procedures were approved by the Institutional Animal Care and Use Committee at the University of Chicago. Adult zebra finches (*Taeniopygia guttata*) were obtained from either a commercial vendor (Magnolia Bird Farms, Anaheim, CA, USA) or our breeding colony at the University of Chicago.

Stimuli

Songs were recorded from birds isolated in a small sound-attenuation chamber (AC-1; Industrial Acoustics Corporation, NY, USA) with an omnidirectional microphone (AT803B, Audio Technica) and digitally saved to a computer disk (sampling rate 20 kHz; band-pass filtering 200 Hz – 10 kHz). A representative song exemplar consisting of two to four motifs – repeated sequences of song syllables – was chosen from the recordings as the BOS stimulus (duration, 1482 – 2772 ms). We then constructed versions of LBOS from each BOS stimulus by amplifying the power in the BOS syllables between 200 Hz – 1000 Hz without any shift in phase using the following cosine curve:

$$K(f) = \begin{cases} A [\cos(k(f - f_0)) + 1] + 1 & (f_1 < f < f_2) \\ 1 & (\text{otherwise}) \end{cases}$$

where $f_0 = (f_2 + f_1)/2$, $k = 2\pi(f_2 - f_1)$, $f_1 = 200$ Hz, and $f_2 = 1000$ Hz (Fig. 1a). The power spectrum of the LBOS was obtained by multiplying $K(f)$ with the power spectrum of the BOS, i.e., $LBOS(f) K(f)^2 P_{BOS}(f)$. Two levels of amplification were used: $A = 80$ or 160 , resulting in $LBOS_{A80}$ and $LBOS_{A160}$ (see Fig. 1c). All stimuli were scaled to a 70 dB root-mean-squared amplitude, i.e., all stimuli had the same total energy but different ratios of power between the low and high frequency ranges (see Fig. 1b). Thus, the LBOS was effectively a low-pass version of the BOS, as the relative power in the high frequency range was attenuated. To the human observer, the resultant LBOS stimuli had a very different “wooden” or “deadened” sound quality compared to the original BOS, consistent with the attenuation of high frequencies. A time-reversed BOS (revBOS) was also presented to all cells, and a time reversed $LBOS_{A160}$ was presented to a subset of the cells. We presented these stimuli in pseudorandom order with an inter-stimulus interval of 10 s. Each stimulus was presented 20 times to each cell for a majority of recorded cells in HVC (85%) and all recorded cells in Area X. We presented it 10 times for the rest of recorded cells in HVC.

Electrophysiological recordings were conducted in an AC-3 (Industrial Acoustics Corporation, NY) sound attenuation booth. For stimulus presentation, we used a power amplifier (D60, Crown Audio, Elkhart, Indiana) with a calibrated mid-woofer speaker (AC-130F1, Aurum Cantus, Penglai City, China). The frequency response of the speaker in the booth was ± 5 dB, as measured with a 1-inch microphone (Brüel & Kjær type 4145) at the position of the bird’s head. A tracking generator (Hewlett Packard 3581C) was used to exclude background noise while measuring the microphone output produced by frequency sweeps from the speaker. The background noise level was 42 dB SPL. The sound intensity was calibrated with a Brüel & Kjær type 4230 calibrator.

Electrophysiological recording and analysis

Several days before an electrophysiological recording session, we implanted a stainless-steel pin after making an opening in the top layer of the bird’s skull caudal to the bifurcation of the midsagittal sinus. For the implantation surgery, the bird was deprived of food and water for one hour and then anesthetized with an intramuscular injection of 50 ml of Equithesin. On the recording day, the bird was food and water deprived for one hour and then anesthetized with three doses of 20% urethane (5 μ l/g, divided into three doses over 1 hour) administered intramuscularly over one hour. The bird was comfortably positioned on a cushion, and the head was immobilized by fastening the implanted pin to a frame. A small window in the bottom layer of the skull was opened to allow access to HVC or Area X. All neurons were initially targeted with stereotaxic coordinates and later confirmed with small electrolytic lesions (5 μ A / 5 s) made at the recording site or nearby fiducial locations. At the end of each experiment, the bird was deeply anesthetized with an overdose of Nembutal and perfused transcardially with heparinized saline followed by formalin. The brains were stored for several days in formalin and then infused with 30% sucrose in formalin before being cut

into 50 μm sections with a cryostat and stained with cresyl violet. The sections were examined to verify the electrode locations, and only cells within the HVC or Area X were included in the following analyses.

Recordings were made with either a silicon microelectrode array (Neuronexus Technologies) or custom-built Pt/Ir solder glass-insulated electrodes (0.003-inch wire; AM Systems Inc., WA). The signals from the electrodes were amplified, filtered (300 Hz to 5 kHz bandpass; Grass Technologies, RI), digitized (DaqBoard/3000; IOtech Inc., OH) at 20 kHz with 16-bit resolution, and saved to a computer disk by a data acquisition program written in C (by Amish S. Dave) running on a Linux operating system. MATLAB® (The Mathworks Inc., MA) was used for offline analysis of the neural data; single units were identified from raw voltage traces using a spike sorting program based on a spike classification algorithm (Fee et al. 1996) included in Chronux, a MATLAB toolbox (<http://chronux.org>). We analyzed well-isolated single units with inter-spike interval histograms that showed few events (less than 1%) in the range of < 1 ms and an inter-spike-interval (ISI) histogram that smoothly converged to zero for small ISIs. During the experiments, the cells were isolated based on spontaneous firing (which was especially salient for Area X neurons) or responses to search sound stimuli. Only cells with significantly different mean firing rates for auditory responses and spontaneous activity (t-test, $p < 0.05$) were included in the analysis. The spontaneous firing rate was determined by counting the number of spikes produced during the one second immediately before each stimulus presentation.

Metric-space analysis. Metric space analysis was conducted with functions included in the Spike Train Analysis Toolkit (Goldberg et al. 2009). Metric-space analysis for spike trains has been applied broadly including the analysis of HVC neurons in canaries (Huetz et al. 2006) and is extensively described elsewhere (Victor and Purpura 1997; Di Lorenzo and Victor 2003). The analysis uses a metric that measures the “distance” between spike trains. The distance is defined as the “cost” to transform a given spike train into another using three types of manipulation: adding, deleting, or shifting a spike. Addition and deletion are assigned identical costs ($=1$), and shifting a spike by an amount of time (Δt [sec]) costs $q|\Delta t|$, where q is a parameter of temporal precision that has units of 1/sec; therefore, the cost is dimensionless quantity. Using the three elementary manipulations, there are multiple ways to transform one spike train into another, and the minimal cost is defined as the distance between the two spike trains that can be found using an efficient algorithm (Aronov 2003). At $q = 0$, there is no cost associated with shifting a spike in time, and thus the distance is determined only by the difference in the total number of spikes (i.e., the spike rate). For $q > 0$, the distance quantifies the difference in the timing of spikes, as a non-zero cost is associated with shifting a spike. A higher q penalizes smaller shifts for a spike, and thus the distance quantifies the difference in the finer temporal structure of spike trains.

At a given temporal precision, we computed the pairwise distance for all of the spike train pairs from the three stimulus types (BOS, LBOS_{A80}, LBOS_{A160}). Next, we calculated the average distance from a given spike train to all other spike trains evoked by each of the three stimuli. We then classified the spike train as belonging to the stimulus type to which the spike train showed the shortest average distance. This classification provides a prediction

regarding which stimulus would evoke the spike train based on the spike distance. We then used mutual information to quantify, on average, how much the prediction reduced the uncertainty in associating stimuli with spike trains. For each neuron, the mutual information was computed with temporal precision from 0 to 1000 [1/sec] with logarithmic steps (52 points). The chance level of the mutual information was estimated by randomly assigning spike trains to a stimulus. A distribution of 20 such randomly shuffled data sets was used to set the lower level of significance as the sum of the mean plus two standard deviations (Victor and Purpura 1996; Victor and Purpura 1997; Di Lorenzo and Victor 2003; Huetz et al. 2006) (Fig. 5). The mutual information at $q = 0$ is called H_{count} as it provides information only regarding differences in spike counts. H_{max} is the maximum amount of information that can be obtained, and its associated temporal precision is called q_{max} . We used the mutual information to discriminate among the three stimulus types, and thus the possible maximum mutual information is equal to $\log_2 3 = 1.585$ [bits].

Results

We analyzed “BOS selective” neurons that showed significant selectivity to the BOS over the revBOS (Margoliash 1983; Margoliash and Fortune 1992). Selectivity was quantified with the d-prime measure for the two stimuli (Solis and Doupe 1997; Theunissen and Doupe 1998), and cells with a d-prime value larger than 0.5 were defined as BOS selective. Among the recorded neurons, 52 of 54 HVC neurons from 11 birds and 30 of 47 Area X neurons from 3 birds satisfied this criterion. We did not identify the cell types of neurons, but the majority of the cells showed tonic activity during the BOS presentation and thus were putatively classified as interneurons (Hahnloser et al. 2002; Rauske et al. 2003).

The firing patterns of HVC neurons were qualitatively very different from those of Area X neurons (Doupe 1997; Rauske et al. 2003). In response to the BOS, most HVC neurons exhibited phasic peaks of excitation over a low rate of spontaneous firing (4.3 ± 6.1 spikes/s), whereas Area X neurons exhibited small peaks of excitation or suppression against a background of very high spontaneous firing (51.9 ± 20.7 spikes/s). This high spontaneous firing rate suggests that recorded neurons are putative globus pallidus-like (GPi) neurons (Woolley et al. 2014). HVC neurons tended to show similar patterns of excitation to sequential motifs of songs whereas Area X neurons tended to show much broader and more generalized patterns of activation throughout the songs (Fig. 2a and b, left panels).

Mean firing rate differences between the BOS and LBOS stimulus, and across HVC and Area X

We first analyzed differences in the mean firing rate (FR) during the entire stimulus presentation period and compared responses to the BOS with the responses to the two types of LBOS for each cell (Fig. 3). The majority of cells (54% – 70% depending on the nucleus and comparison) showed no significant difference between the BOS and LBOS stimuli (Table 1). Thus, the majority of HVC and Area X cells did not alter their FRs based on the relative power of low and high frequencies in the BOS stimuli.

For the substantial minority of cells that showed significant differences in FRs between the BOS and LBOS stimuli, the pattern of differences varied across the two nuclei. The LBOS_{A80} evoked significantly higher mean FRs (one-tailed t-test, $p < 0.025$) than the BOS in 17% (9/52) of BOS-selective HVC neurons, while the BOS evoked significantly higher FR than LBOS_{A80} in 23% (12/52) of BOS-selective HVC neurons (Fig. 3a). In contrast, in Area X, 30% (9/30) of neurons showed significantly higher FRs in response to the LBOS_{A80} than to the BOS, but none of the neurons showed a higher FR in response to the BOS than to the LBOS_{A80} (Fig. 3b). Similar results were obtained for the LBOS_{A160}, with the same percentage of cells showing significantly higher FRs to the LBOS_{A160} than the BOS (30%; 9/30). Except for one or two cells (depending on the comparison), the same cells showed higher FR in response to the LBOS_{A80} and the LBOS_{A160}.

In contrast, for cells that preferred the BOS to the LBOS stimuli the percentage of cells that showed a significantly higher FR in response to the BOS than to the LBOS_{A80} or LBOS₁₆₀ was approximately the same in HVC, with a slight tendency for more neurons to prefer the BOS over the LBOS_{A160} than the BOS over the LBOS_{A80} (Table 1). Zero (0/30, LBOS_{A80}) or only 7% (2/30, LBOS₁₆₀) of Area X cells, however, preferred the BOS over the LBOS (Table 1).

Overall, comparing FR responses of HVC and Area X neurons, these results are summarized as a tendency for an increase in the percentage of cells preferring the LBOS over BOS with a tendency for loss of cells preferring the BOS over LBOS.

This sensitivity to the LBOS over the BOS was not explained, in either nucleus, by the strength of BOS selectivity of those cells. We evaluated this by comparing the response to BOS relative to revBOS using the d-prime value (Fig. 4a). We did not find a significant correlation between the d-prime values calculated for the BOS relative to the revBOS and those for the BOS relative to the LBOS (A=80 or 160) in either HVC ($r=0.079$, $p=0.57$ for A=80; $r=0.852$, $p=0.55$ for A=160) or Area X ($r=0.01$, $p=0.96$ for A=80; $r=0.143$, $p=0.44$ for A=160) (Fig. 4b).

We also evaluated whether the FR was significantly different among the three stimuli (BOS, LBOS_{A80} and LBOS_{A160}) at the population level. We performed a two-way ANOVA with the stimulus type as a fixed effect and the bird as a random effect. No significant main effect was found for the stimulus type in either HVC ($F_{(2, 143)}=0.09$, $p=0.91$) or Area X ($F_{(2, 85)}=0.7$, $p=0.50$), consistent with our prior analysis showing that the majority of cells did not show individual differences in spike rates that could distinguish among the stimuli.

Temporal coding conveys additional information for discriminating between the BOS and LBOS

The prior analyses demonstrate that the majority of HVC and Area X cells exhibit equal firing rates in response to the BOS and LBOS. These analyses, however, do not account for potential information in the temporal structure of spiking activity. Cells may produce different temporal patterns of spikes in response to the BOS and LBOS, which may carry information about the differences in the spectral content of those stimuli. To test this

hypothesis, we quantified the difference in spiking patterns at different time scales by applying metric-space analysis (Victor and Purpura 1996; Victor and Purpura 1997).

We computed the mutual information to discriminate among the three stimuli (BOS, LBOS_{A80}, LBOS_{A160}) for each neuron. Most of the neurons (51/52, HVC; 27/30, Area X) showed significant amounts of information (see Materials and Methods) at one or more values of the cost parameter for temporal precision (q) that were larger than 0. In these neurons, the temporal code yielded a larger amount of information than the spike count. To determine how much more information could be obtained from the temporal code in each neuron's response to the three stimuli, we compared the maximum mutual information (H_{max}) with the information based on the difference in spike counts (H_{count}). For cells with temporal information, the mutual information peaked at a non-zero q , and the H_{max} was different from the H_{count} (e.g., Fig. 5). The great majority of HVC cells (94%; 48/51) and Area X cells (85%, 23/27) showed the maximum mutual information value at a non-zero q (Fig. 6a, non-diagonal values), suggesting that the temporal code carried additional information beyond the spike count in the responses of these cells.

Comparing the distributions of q_{max} , H_{max} , and H_{count} across the two nuclei also yielded insight into differences in neural coding between HVC and Area X. The medians of the q_{max} values were 39.8 [1/sec] (i.e., $1/q_{max} = 25.1$ ms) for HVC, and 15.8 [1/sec] (i.e., $1/q_{max} = 63.2$ ms) for Area X, which are not significantly different (Wilcoxon rank sum test, $p=0.145$). (Making this calculation using the geometrical means of non-zero q_{max} yielded similar results: 24.7 [1/s] for HVC and 17.872 [1/s] for area X, with non-significant ($p=0.518$) differences between these geometrical means, as evaluated by a t-test with log q_{max} .) Moreover, the distributions of q_{max} in the two nuclei were not significantly different (Kolmogorov-Smirnov test, $p = 0.13$, Fig. 6b). Thus, neurons in both nuclei maximally discriminated the difference in the low-frequency components on a similar time scale. The population mean of the H_{max} for HVC, however, was significantly larger than that for Area X (0.45 bits HVC, 0.20 bits Area X, one-sided t-test, $p<0.0001$, $t=5.99$) (Fig. 6c, left), while the means of the H_{count} were not significantly different between HVC and Area X (0.145 bits HVC, 0.11 bits Area X, t-test, $p = 0.088$, $t=1.36$) (Fig. 6c, right). Thus, more mutual information is contained in the HVC temporal code than in the Area X temporal code when the LBOS and BOS are compared, whereas there are similar amounts of information in both nuclei in terms of spike count differences. Qualitatively, a much broader distribution in H_{max} was observed for HVC neurons, and over half the population had higher H_{max} values than any Area X neuron.

In summary, the results indicate that: 1) temporal coding in both nuclei carries more information than firing rates with respect to the relative power distribution of the BOS and the LBOS stimuli, 2) the optimal discrimination of the stimuli is obtained with similar temporal precision in both nuclei, and 3) HVC neurons are, on average, better than Area X neurons in discriminating the stimuli using the temporal code.

Discussion

We investigated the response of song system neurons to BOS and LBOS. Although the power spectrum of the LBOS had a spectral tilt that was very different from that of the BOS, the firing rates (spike counts) evoked by the LBOS stimulus were similar to those evoked by the BOS in both HVC and Area X neurons, although almost no Area X neurons preferred BOS over LBOS based on firing rates. In contrast, the metric-space analysis revealed that the great majority of the cells in both nuclei encode the largest amount of mutual information when discriminating between the BOS and LBOS using spike patterns. Thus, our data are characterized by relative invariance across BOS variants in the firing rate of the neurons, and sensitivity across those neurons in the temporal coding of BOS variants.

Low-frequency information in the song system

Since the earliest auditory neurophysiological experiments on the song system, airborne (microphone) recordings of the BOS have been used to stimulate song system neurons. Numerous electrophysiological experiments in multiple songbird species have established that the BOS is an effective stimulus throughout the forebrain song system (McCasland and Konishi 1981; Margoliash 1983; Williams and Nottebohm 1985; Doupe and Konishi 1991; Margoliash and Fortune 1992; Janata and Margoliash 1999; Nealen and Schmidt 2006; Prather et al. 2008). Song system neurons actively adapt their BOS-selective auditory responses as birds develop their songs during sensorimotor learning (Volman 1993; Doupe 1997; Nick and Konishi 2005). In most cases, the selectivity of neurons has been assessed by comparing responses to conspecific songs or other songs of the bird's repertoire, albeit rarely in an ethologically defined context (e.g. Margoliash and Konishi 1985; Prather et al. 2009). Song system neurons can exhibit highly non-linear spectral and temporal receptive field properties (Margoliash and Fortune 1992), and thus even extensive testing does not exclude the possibility of a correlated feature of the BOS (such as the low-frequency components) that has yet to be described. Novel regions of extensively studied auditory receptive fields have also been described in mammalian systems (Ohlemiller et al. 1996).

Previous studies have systematically explored the acoustic features of the song to which song system neurons are sensitive. HVC neurons in white-crowned sparrows tuned to the absolute frequency of song components were more sensitive to spectral than to amplitude modulation components of the song (Margoliash 1983; Margoliash and Konishi 1985). Another study in zebra finches explored the highly BOS-selective responses of song system neurons with systematically degraded versions of the BOS. This result demonstrated that the mean firing rate has greater sensitivity to temporal than to spectral cues of the BOS (Theunissen and Doupe 1998). Similar to these results, we found that the mean firing rates were not significantly different among BOS and LBOS stimuli in HVC. This small difference was somewhat unexpected given the approximately 15 dB reduction in power at frequencies above 1 kHz between the LBOS and BOS (Fig. 1). Such a response property might result from the intensity invariant response of neurons in early auditory nuclei (Billimoria et al. 2008; Nagel and Doupe 2008). Our subsequent analyses revealed that most of HVC neurons convey additional information to discriminate the BOS and the LBOS via

patterns of spike timing. Interestingly, a recent study using metric-space analysis showed that the temporal code in the robust nucleus of the arcopallium (RA) conveys much information about trial-to-trial variation of song spectral content during singing (Tang et al. 2014).

Overall, for both the rate and temporal codes, we found that neurons in Area X had a reduced ability to discriminate between the BOS and the LBOS than those in HVC. These results might suggest that BOS-selective HVC neurons have a greater ability to discriminate among song stimuli than BOS-selective neurons in the basal ganglia nucleus Area X, which is implicated in song plasticity (Bottjer and Arnold 1984; Farries et al. 2005). However, we recorded neurons with high spontaneous firing rates in Area X, which suggests that these are internal globus pallidus-like (GPi) neurons (Woolley et al. 2014). In addition, another type of cells is found in Area X, the spiny neurons (SNs). These neurons have very low spontaneous firing rates, and their firing is very well timed and corresponds to a particular syllable in the song motif, similar to projection neurons in HVC (Woolley et al. 2014). Thus, SNs neurons may be more sensitive to the frequency components of the BOS than GPi neurons. Furthermore, we did not identify the classes of HVC cells that we recorded from, although probably most were interneurons with few if any being Area X projecting HVC neurons. Thus, the locus where the BOS selectivity distinction first arises in the basal ganglia pathway remains to be determined.

Low-frequency components can be a salient cue in AF during singing

Prior studies have estimated the auditory thresholds of zebra finches to airborne tone bursts using behavioral responses (Okanoya and Dooling 1987) or evoked potentials from the auditory brainstem (Zevin et al. 2004). These results suggest that birds have a relatively weak sensitivity to low frequency tones below 1 kHz. Greater power at low frequencies is a feature of songs recorded at the syrinx (Fee et al. 1998; Goller et al. 2004) compared to airborne songs, however. This suggests that low frequencies could be perceived during singing by coupling to the middle ear via substrate vibration through the body. Recent studies using accelerometers to measure songs conducted through the body (Anisimov et al. 2014) and the cranium (Fukushima and Margoliash 2015) from singing zebra finches may support this possibility. In particular, the recordings from the cranium produce very high quality copies of the songs the birds sang, and with an increase in power at low frequencies. These vibrations are likely to be coupled to the middle ear, providing a copy of the song that emphasizes the low-frequency components. The transfer function from bone vibration during singing to cochlear stimulation is unknown, and birds will also experience airborne signals that may interact with bone-conducted signals. Measurements of cochlear microphonics or auditory brainstem responses may be valuable approaches to assess what signals birds perceive during singing.

Internally propagated AF components could act as an additional channel to robustly detect self-generated sounds robustly in the presence of contaminating signals such as sounds made by other birds or siblings. A similar strategy has been shown in bat echolocation (Suga et al. 1979; O'Neill and Suga 1982). The neural mechanism of auditory feedback processing is a major outstanding issue in birdsong learning and adult song maintenance. Our results may

motivate the exploration for a role of the low-frequency signals in the song system during vocal learning.

Acknowledgements:

We thank Daniel Baleckaitis for technical assistance in histology.

Grants: Supported by NIH grant MH59831 to D.M.

References

- Anisimov VN, Herbst JA, Abramchuk AN, Latanov AV, Hahnloser RHR, Vyssotski AL. Reconstruction of vocal interactions in a group of small songbirds. *Nat Methods*. 2014; 11:1135–1137. doi: 10.1038/nmeth.3114. [PubMed: 25262206]
- Aronov D. Fast algorithm for the metric-space analysis of simultaneous responses of multiple single neurons. *J Neurosci Methods*. 2003; 124:175–179. [PubMed: 12706847]
- Billimoria CP, Kraus BJ, Narayan R, Maddox RK, Sen K. Invariance and sensitivity to intensity in neural discrimination of natural sounds. *J Neurosci*. 2008; 28:6304–6308. doi: 10.1523/JNEUROSCI.0961-08.2008. [PubMed: 18562600]
- Botzler SW, Arnold AP. The role of feedback from the vocal organ. I. Maintenance of stereotypical vocalizations by adult zebra finches. *The Journal of Neuroscience*. 1984; 4:2387–2396. [PubMed: 6541246]
- Borden GJ. An interpretation of research of feedback interruption in speech. *Brain Lang*. 1979; 7:307–319. [PubMed: 455050]
- Dave AS, Yu AC, Margoliash D. Behavioral state modulation of auditory activity in a vocal motor system. *Science*. 1998; 282:2250–2254. [PubMed: 9856946]
- Di Lorenzo PM, Victor JD. Taste response variability and temporal coding in the nucleus of the solitary tract of the rat. *J Neurophysiol*. 2003; 90:1418–1431. doi: 10.1152/jn.00177.2003. [PubMed: 12966173]
- Doupe AJ. Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J Neurosci*. 1997; 17:1147–1167. [PubMed: 8994068]
- Doupe AJ, Konishi M. Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad USA*. 1991; 88:11339–11343.
- Doupe AJ, Kuhl PK. Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci*. 1999; 22:567–631. doi: 10.1146/annurev.neuro.22.1.567. [PubMed: 10202549]
- Farries MA, Ding L, Perkel DJ. Evidence for "direct" and "indirect" pathways through the song system basal ganglia. *J Comp Neurol*. 2005; 484:93–104. doi: 10.1002/cne.20464. [PubMed: 15717304]
- Fee MS, Mitra PP, Kleinfeld D. Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *J Neurosci Methods*. 1996; 69:175–188. doi: 10.1016/S0165-0270(96)00050-7. [PubMed: 8946321]
- Fee MS, Shraiman B, Pesaran B, Mitra PP. The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature*. 1998; 395:67–71. doi: 10.1038/25725. [PubMed: 12071206]
- Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O. De novo establishment of wild-type song culture in the zebra finch. *Nature*. 2009; 459:564–568. doi: 10.1038/nature07994. [PubMed: 19412161]
- Fukushima M, Margoliash D. The effects of delayed auditory feedback revealed by bone conduction microphone in adult zebra finches. *Sci Rep*. 2015; 5:8800. doi: 10.1038/srep08800. [PubMed: 25739659]
- Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC. Recursive syntactic pattern learning by songbirds. *Nature*. 2006; 440:1204–1207. doi: 10.1038/nature04675. [PubMed: 16641998]

- Goldberg DH, Victor JD, Gardner EP, Gardner D. Spike train analysis toolkit: enabling wider application of information-theoretic techniques to neurophysiology. *Neuroinformatics*. 2009; 7:165–178. doi: 10.1007/s12021-009-9049-y. [PubMed: 19475519]
- Goller F, Daley MA. Novel motor gestures for phonation during inspiration enhance the acoustic complexity of birdsong. *Proc Biol Sci*. 2001; 268:2301–2305. doi: 10.1098/rspb.2001.1805. [PubMed: 11703869]
- Goller F, Mallinckrodt MJ, Torti SD. Beak gape dynamics during song in the zebra finch. *J Neurobiol*. 2004; 59:289–303. doi: 10.1002/(ISSN)1097-4695. [PubMed: 15146546]
- Hahnloser RHR, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*. 2002; 419:65–70. doi: 10.1038/nature00974. [PubMed: 12214232]
- Huetz C, Del Negro C, Lebas N, Tarroux P, Edeline J-M. Contribution of spike timing to the information transmitted by HVC neurons. *Eur J Neurosci*. 2006; 24:1091–1108. doi: 10.1111/ejn.2006.24.issue-4. [PubMed: 16930435]
- Janata P, Margoliash D. Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *J Neurosci*. 1999; 19:5108–5118. [PubMed: 10366643]
- Jensen KK, Cooper BG, Larsen ON, Goller F. Songbirds use pulse tone register in two voices to generate low-frequency sound. *Proc Biol Sci*. 2007; 274:2703–2710. doi: 10.1098/rspb.2007.0781. [PubMed: 17725979]
- Konishi M. The role of auditory feedback in birdsong. *Ann N Y Acad Sci*. 2004; 1016:463–475. doi: 10.1196/annals.1298.010. [PubMed: 15313790]
- Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Fehér O, Ravbar P, Okanoya K, Tchernichovski O. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*. 2013; 498:104–108. doi: 10.1038/nature12173. [PubMed: 23719373]
- Margoliash D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J Neurosci*. 1983; 3:1039–1057. [PubMed: 6842281]
- Margoliash D. Offline learning and the role of autogenous speech: New suggestions from birdsong research. *Speech Communication*. 2003; 41:165–178. doi: 10.1016/S0167-6393(02)00101-2.
- Margoliash D, Fortune ES. Temporal and harmonic combination-sensitive neurons in the zebra finch's HVC. *J Neurosci*. 1992; 12:4309–4326. [PubMed: 1432096]
- Margoliash D, Konishi M. Auditory representation of autogenous song in the song system of white-crowned sparrows. *Proc Natl Acad USA*. 1985; 82:5997–6000.
- Margoliash D, Schmidt MF. Sleep, off-line processing, and vocal learning. *Brain Lang*. 2010; 115:45–58. doi: 10.1016/j.bandl.2009.09.005. [PubMed: 19906416]
- McCasland JS, Konishi M. Interaction between auditory and motor activities in an avian song control nucleus. *Proc Natl Acad USA*. 1981; 78:7815–7819.
- Nagel KI, Doupe AJ. Organizing principles of spectro-temporal encoding in the avian primary auditory area field L. *Neuron*. 2008; 58:938–955. doi: 10.1016/j.neuron.2008.04.028. [PubMed: 18579083]
- Nealen PM, Schmidt MF. Distributed and selective auditory representation of song repertoires in the avian song system. *J Neurophysiol*. 2006; 96:3433–3447. doi: 10.1152/jn.01130.2005. [PubMed: 16885516]
- Nick TA, Konishi M. Neural auditory selectivity develops in parallel with song. *J Neurobiol*. 2005; 62:469–481. doi: 10.1002/neu.20115. [PubMed: 15616963]
- Nordeen KW, Nordeen EJ. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol*. 1992; 57:58–66. [PubMed: 1567334]
- Nottebohm F, Nottebohm ME, Crane L. Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song-control nuclei. *Behav Neural Biol*. 1986; 46:445–471. [PubMed: 3814048]
- Ohlemiller KK, Kanwal JS, Suga N. Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. *Neuroreport*. 1996; 7:1749–1755. [PubMed: 8905657]

- Okanoya K, Dooling RJ. Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. *J Comp Psychol.* 1987; 101:7–15. [PubMed: 3568610]
- Okanoya K, Yamaguchi A. Adult Bengalese finches (*Lonchura striata* var. *domestica*) require real-time auditory feedback to produce normal song syntax. *J Neurobiol.* 1997; 33:343–356. [PubMed: 9322153]
- Oller DK, Eilers RE. The role of audition in infant babbling. *Child Dev.* 1988; 59:441–449. [PubMed: 3359864]
- Prather JF, Nowicki S, Anderson RC, et al. Neural correlates of categorical perception in learned vocal communication. *Nat Neurosci.* 2009; 12:221–228. doi: 10.1038/nn.2246. [PubMed: 19136972]
- Prather JF, Peters S, Nowicki S, Mooney R. Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature.* 2008; 451:305–310. doi: 10.1038/nature06492. [PubMed: 18202651]
- Rauske PL, Shea SD, Margoliash D. State and neuronal class-dependent reconfiguration in the avian song system. *J Neurophysiol.* 2003; 89:1688–1701. doi: 10.1152/jn.00655.2002. [PubMed: 12626633]
- Solis MM, Doupe AJ. Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. *J Neurosci.* 1997; 17:6447–6462. [PubMed: 9236252]
- Tang C, Chehayeb D, Srivastava K, et al. Millisecond-Scale Motor Encoding in a Cortical Vocal Area. *PLoS Biol.* 2014; 12:e1002018. doi: 10.1371/journal.pbio.1002018.s004. [PubMed: 25490022]
- Theunissen FE, Doupe AJ. Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVC of male zebra finches. *J Neurosci.* 1998; 18:3786–3802. [PubMed: 9570809]
- Victor JD, Purpura KP. Metric-space analysis of spike trains: Theory, algorithms and application. *Network.* 1997; 8:127–164.
- Victor JD, Purpura KP. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J Neurophysiol.* 1996; 76:1310–1326. [PubMed: 8871238]
- Volman SF. Development of neural selectivity for birdsong during vocal learning. *J Neurosci.* 1993; 13:4737–4747. [PubMed: 8229196]
- Waldstein RS. Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *J Acoust Soc Am.* 1990; 88:2099–2114. [PubMed: 2269726]
- Williams H, Nottebohm F. Auditory responses in avian vocal motor neurons: a motor theory for song perception in birds. *Science.* 1985; 229:279–282. [PubMed: 4012321]
- Woolley SC, Rajan R, Joshua M, Doupe AJ. Emergence of Context-Dependent Variability across a Basal Ganglia Network. *Neuron.* 2014; 82:208–223. doi: 10.1016/j.neuron.2014.01.039. [PubMed: 24698276]
- Woolley SMN, Rubel EW. Bengalese finches *Lonchura Striata domestica* depend upon auditory feedback for the maintenance of adult song. *J Neurosci.* 1997; 17:6380–6390. [PubMed: 9236246]
- Woolley SMN, Rubel EW. High-frequency auditory feedback is not required for adult song maintenance in Bengalese finches. *J Neurosci.* 1999; 19:358–371. [PubMed: 9870965]
- Zevin JD, Seidenberg MS, Bottjer SW. Limits on reacquisition of song in adult zebra finches exposed to white noise. *J Neurosci.* 2004; 24:5849–5862. doi: 10.1523/JNEUROSCI.1891-04.2004. [PubMed: 15229232]

Fig. 1. Examples of low-frequency BOS (LBOS) stimuli

a. The cosine curves (see Materials and Methods) used to amplify the low-frequency ranged between 200 and 1000 Hz; $k(f)$ is the amplification factor. The maximum amplification was 320 times ($A=160$) or 160 times ($A=80$) at 600 Hz. **b.** Examples of the power spectrums of the BOS and LBOS ($A=80, 160$) are shown. The total energy (the area under the curve) is constant for all three stimuli. **c.** An example spectrogram and waveforms for the BOS (top), LBOS_{A80} (middle), and LBOS_{A160} (bottom) are shown. The same maximum value is used in the color map to graphically display the three spectrographs, emphasizing the relatively greater power at low frequencies for the LBOS and the relatively greater power at higher frequencies for the BOS.

Fig. 2. Examples of single cell responses to the BOS and LBOS₈₀ in HVC (a) and Area X (b) Each panel shows (from the top to the bottom) the sound waveform, spectrogram, peristimulus histogram (PSTH), and raster plot of the spike trains. The HVC and Area X neurons have low and high spontaneous firing rates, respectively, and the HVC responses are more phasic than those of Area X neurons. The response profiles to the LBOS are similar to the BOS, especially for HVC neurons.

Fig. 3. Mean firing rates during presentation of the LBOS (FR_{LBOS}) and BOS (FR_{BOS})
 For each point, the baseline firing rate of the cell was subtracted. The X-axis shows the response to the LBOS, and the Y-axis shows the response to the BOS. Red dots indicate cells with significantly higher FRs in response to the BOS than the LBOS, and black dots indicate cells with higher FR response to the LBOS than the BOS (one-sided t-test, $p < 0.025$). Blue circles indicate cells that did not show any significant difference in FR for the LBOS and BOS. *Left:* BOS vs. $LBOS_{A80}$, *Right:* BOS vs. $LBOS_{A160}$. **a.** HVC neurons ($n = 52$), **b.** Area X neurons ($n = 30$).

Fig. 4. d-prime values of the responses to BOS and BOS variants

a. The cumulative probability distributions of the d-prime value for the discriminability of the BOS compared to BOS variant stimuli (revBOS, Red; LBOS₈₀, Black; and LBOS₁₆₀, Blue) are shown. A positive d-prime value indicates that the BOS evoked a higher response rate than the BOS variant stimuli. Left: HVC, Right Area X. Notice the broader distribution of d-prime values for HVC compared with Area X. **b.** Scatter plots showing the d-prime value for each cell. Each circle corresponds to a neuron. The x-axis indicates the d-prime value for BOS compared to revBOS, while the y-axis indicates the d-prime value for BOS compared to LBOS (black, A=80, blue, A=160). Left: HVC, Right: Area X. Notice that in each nucleus, the distributions overlap.

Fig. 5. Examples of the mutual information as the function of q , the parameter for temporal precision

a. HVC, b. Area X. In each plot, the black line with circles indicates the information from the original (actual) data, and the red line with squares indicates the information from the shuffled data (see Materials and Methods). The mean and standard deviation (SD) of the shuffled data were estimated from 20 randomly shuffled data points. The error bar corresponds to ± 2 SD. The information from the original data reached its maximum (H_{max}) at $q = q_{max}$. The information from the spike count corresponds to the information at $q = 0$ (H_{count}).

Fig. 6. Mutual information carried by HVC and Area X cells used to discriminate the three types of BOS

a. For each cell, the H_{count} (the mutual information from spike counts) and H_{max} (the maximum information) values were computed and plotted on the X and Y axes, respectively. Only cells with significant information (see Materials and Methods) were included (HVC, $n=51$, Area X, $n=27$). **b.** In the distribution of q_{max} , the temporal precision produces the maximum amount of information. **c.** Population mean ($\pm$ SEM) of mutual information. *Left:* population means of the H_{max} for HVC and Area X. *Right:* population means of H_{count} for HVC and Area X. For each plot, a t-test was used to evaluate the significance of the differences in the mean values, and the p-value is shown if the difference was significant.

Table 1

Summary of pair wise comparisons of mean firing rates during presentation of the LBOS and BOS.

Nucleus (total)	Type of comparison	Higher FR for LBOS	Higher FR for BOS	No significant difference
HVC (n=52)	BOS vs. LBOS _{A80}	17% (9 cells)	23% (12 cells)	60% (31 cells)
	BOS vs. LBOS _{A160}	17% (9 cells)	29% (15 cells)	54% (28 cells)
Area X (n=30)	BOS vs. LBOS _{A80}	30% (9 cells)	0% (0 cells)	70% (21 cells)
	BOS vs. LBOS _{A160}	30% (9 cells)	7% (2 cells)	63% (19 cells)

For each cell, a one-tailed t-test was used to compare the FRs evoked by the LBOS and BOS. The significance level for the p-value is 0.025.