

Central Contributions to Acoustic Variation in Birdsong

Samuel J. Sober,* Melville J. Wohlgemuth,* and Michael S. Brainard

Department of Physiology, W. M. Keck Center for Integrative Neuroscience, San Francisco, California 94143-0444

Birdsong is a learned behavior remarkable for its high degree of stereotypy. Nevertheless, adult birds display substantial rendition-by-rendition variation in the structure of individual song elements or “syllables.” Previous work suggests that some of this variation is actively generated by the avian basal ganglia circuitry for purposes of motor exploration. However, it is unknown whether and how natural variations in premotor activity drive variations in syllable structure. Here, we recorded from the premotor nucleus robust nucleus of the arcopallium (RA) in Bengalese finches and measured whether neural activity covaried with syllable structure across multiple renditions of individual syllables. We found that variations in premotor activity were significantly correlated with variations in the acoustic features (pitch, amplitude, and spectral entropy) of syllables in approximately a quarter of all cases. In these cases, individual neural recordings predicted $8.5 \pm 0.3\%$ (mean \pm SE) of the behavioral variation, and in some cases accounted for 25% or more of trial-by-trial variations in acoustic output. The prevalence and strength of neuron–behavior correlations indicate that each acoustic feature is controlled by a large ensemble of neurons that vary their activity in a coordinated manner. Additionally, we found that correlations with pitch (but not other features) were predominantly positive in sign, supporting a model of pitch production based on the anatomy and physiology of the vocal motor apparatus. Collectively, our results indicate that trial-by-trial variations in spectral structure are indeed under central neural control at the level of RA, consistent with the idea that such variation reflects motor exploration.

Key words: basal ganglia; birdsong; correlated variability; extracellular recordings; premotor; vocalization

Introduction

The acquisition of any complex skill, whether learning to speak or to throw a curveball, is associated with a gradual decrease in motor variability. Our initial attempts are variable and inaccurate, apparent in the babbling speech or wild pitches familiar to anyone who has observed motor learning in action. With practice, however, we learn to reliably produce the desired output.

Even well-practiced movements, however, retain some variability. Such variation might reflect unreliability at the neuromuscular junction or in the muscles themselves, although the brain might repeat the same activity pattern during each movement. Alternately, variability might result from variations in the motor command. In primates, variations in neural activity predict variations in eye and arm movements (Churchland et al., 2006b; Medina and Lisberger, 2007). These results suggest that some of the “residual” variation in well-learned skills is driven by the brain.

Birdsong is an excellent model system for studying the control of behavioral variation. After exposure to the song of an adult

“tutor,” developing song becomes highly stereotyped, or “crystallized” (Arnold, 1975), as illustrated by the example recordings in Figure 1*a*. Crystallized song, however, displays significant variation across multiple renditions of the same syllable (Kao et al., 2005; Olveczky et al., 2005; Sakata et al., 2008). We refer to such cross-rendition variation in syllable structure as “trial-by-trial” variation. Figure 1*b* shows the trial-by-trial variation of two acoustic parameters (pitch and amplitude) for one syllable of crystallized song.

A significant component of this variation likely originates in the brain. The anterior forebrain pathway (AFP), via its output, the lateral magnocellular nucleus of the anterior nidopallium (IMAN), sends input to premotor circuits in the robust nucleus of the arcopallium (RA) (see Fig. 1*c*). Several lines of evidence suggest that IMAN drives behavioral variation by injecting neural variation into RA: lesions or inactivation of IMAN dramatically reduce variation, stimulation in IMAN affects syllable pitch and amplitude, and the level of variation in IMAN activity correlates with the level of variation in behavior (Kao et al., 2005; Olveczky et al., 2005; Kao and Brainard, 2006).

It is unclear, however, how variation generated by the AFP is distributed across RA, or indeed whether trial-by-trial neural variation in RA has any behavioral consequences at all. Neural variability in RA is remarkably low (Chi and Margoliash, 2001), and RA activity is far more precise than that of primate cortex. If RA and upstream areas are to drive trial-by-trial changes in behavior, then small variations in RA activity must in turn drive variations in song. Neural variation might be distributed across RA in several different ways. Variation might be restricted to a small subpopulation of neurons, which exert powerful control over trial-by-trial variations in behavior (see Fig. 2*a*). Alternately,

Received May 29, 2008; revised Aug. 7, 2008; accepted Sept. 3, 2008.

This work was supported by a Helen Hay Whitney postdoctoral fellowship (S.J.S.) and by a National Institute on Deafness and Other Communication Disorders R01 award, a National Institute of Mental Health Conte Center for Neuroscience Research award, and a McKnight Foundation Scholars award (M.S.B.). We thank Allison Doupe, Stephen Lisberger, and Philip Sabes for helpful comments on this manuscript, and Jonathan Wong and Rajarshi Mazumder for technical assistance.

*S.J.S. and M.J.W. contributed equally to this work.

Correspondence should be addressed to Samuel J. Sober, Department of Physiology, W. M. Keck Center for Integrative Neuroscience, Box 0444, 513 Parnassus Avenue, San Francisco, CA 94143-0444. E-mail: sam@phy.ucsf.edu.

DOI:10.1523/JNEUROSCI.2448-08.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/2810370-10\$15.00/0

the activity of RA neurons might vary independently, influencing motor output with the sum of many independent fluctuations (see Fig. 2*b*). In a third model, trial-by-trial variations across RA might be correlated such that variations in behavior result from coordinated changes in firing across many neurons (see Fig. 2*c*).

We hypothesized that a component of trial-by-trial acoustic variability results from trial-by-trial variations in RA activity. To test this, we recorded from RA neurons in adult Bengalese finches and asked whether variations in spiking activity across multiple renditions of individual syllables could account for variations in the syllables' pitch, amplitude, and spectral entropy. Furthermore, by examining the prevalence, signs, and strengths of neural–behavioral correlations, we investigated how variation is distributed across neurons and compared how pitch, amplitude, and spectral entropy are encoded in RA.

Materials and Methods

Adult (>100 d old) Bengalese finches (*Lonchura striata* var. *domestica*) were bred in our colony and housed with their parents until at least 60 d of age. After electrode implantation, birds were isolated and housed individually in sound-attenuating chambers (Acoustic Systems) with food and water provided *ad libitum*. Unless otherwise specified, all recordings presented here are from undirected song (i.e., no female was present). All procedures were performed in accordance with established animal care protocols approved by the University of California, San Francisco Institutional Animal Care and Use Committee.

Electrophysiological data collection. Birds were anesthetized (induction with 20 mg/kg ketamine and 1.5 mg/kg midazolam, maintained with 0.5–2.0% isoflurane) and a lightweight microdrive (Hessler and Doupe, 1999) was positioned stereotactically over RA in one hemisphere (10 implants over right RA, 3 over left RA) and secured to the skull with epoxy. Each microdrive carried a custom-made array of three to five high-impedance microelectrodes (Microprobe WE1.5QT35.0A3) with all electrode tips grouped within 300 μ m of each other. After recovery from surgery, birds resumed singing within 1–3 d. After singing resumed, electrode arrays were lowered into RA. Extracellular spike waveforms as large as 4.5 mV were recorded during and between bouts of singing. RA recording sites were identified by the presence of characteristic changes in activity associated with the production of songs and calls and by *post hoc* histological confirmation of trajectory of each electrode array. In a subset of birds (which contributed ~31% of all recorded units), we were able to estimate the dorsal–ventral position of the array on each recording. We observed no significant differences between dorsal and ventral RA with respect to the prevalence or sign of correlations between neural activity and acoustic output.

We used a quantitative technique (supplemental text, Quantifying unit isolation, available at www.jneurosci.org as supplemental material) to measure the isolation of spike waveforms. Briefly, we performed principal component analysis (PCA) on recorded voltage waveforms, examined their projections along the first two components, and quantified the extent of overlap between waveform clusters. Recordings yielding clusters with overlaps of <0.01 were classified as single units, and recordings with larger overlaps were classified as multiunit clusters, reflecting the potential contribution of several neurons to each recording. This technique yielded isolation estimates that agreed well with both qualitative assessments of isolation and estimates based on spike refractory periods. As described below, data from single-unit and multiunit recordings yielded nearly identical results. In total, we collected 145 RA recordings (25 single-unit, 120 multiunit) from 13 birds. Unless the level of isolation is specified, the term “unit” in this study refers to either a single unit or to one multiunit cluster.

Spiking statistics. In quantifying the firing statistics of single units ($n = 25$) in the Bengalese finch, we performed analyses similar to those in a previous study of zebra finch RA (Leonardo and Fee, 2005) to facilitate comparison between the two species. The instantaneous firing rate was defined as the inverse of the interspike interval (ISI) and computed for all single units using only data recorded during singing. Based on the bi-

modal distribution of instantaneous firing rate shown in Figure 3*c*, we selected a 50 Hz threshold dividing bursting intervals (with higher firing rates) from nonbursting intervals (with lower rates). Using this criterion, we assigned every ISI either to an ongoing burst or to an interburst pause based on its instantaneous firing rate. Instantaneous firing rate distributions from individual single units resembled the pooled distribution shown in Figure 3*c*, with thresholds distributed at $\sim 54.0 \pm 19.9$ Hz (mean \pm SD).

Only single units were used in the analyses shown in Figure 3*c*. However, the distributions of burst durations from single-unit and multiunit recordings were highly overlapping (29 ± 25 ms for single-unit, 33 ± 37 ms for multiunit, mean \pm SD). Although the true number of neurons contributing to a multiunit signal is difficult to ascertain, these results suggest that many of our multiunit recordings reflect the activity of a relatively small number of neurons.

Acoustic features. To characterize the relationship between neural and behavioral variation, we measured the acoustic properties of each syllable rendition as well as the premotor spiking activity before each syllable. Owing to the complex acoustic structure of song, we must take care that the acoustic parameters being quantified reflect important dimensions of behavioral variation, because failing to do so could cause us to underestimate the contributions of RA activity to behavioral variation. Previous work on the song system has identified fundamental frequency (pitch), amplitude, and spectral entropy as potentially important axes of behavioral variation because they are refined during song learning, vary from trial to trial in the adult, and can be perturbed by electrical stimulation of the song system during singing (Tchernichovski et al., 2001; Fee et al., 2004; Kao et al., 2005). To address this issue quantitatively, we performed a separate analysis in which we used PCA to ask which dimensions of spectral variation capture the greatest fraction of the total spectral variability of each syllable. As described in detail in the supplemental text (Quantitative analysis of spectral variability and supplemental Figs. 2–4, available at www.jneurosci.org as supplemental material), the PCA-based analysis revealed that, in most cases, variations in pitch, amplitude, or entropy indeed captured the greatest fraction of the total behavioral variation. These acoustic features therefore dominate behavioral variability and as such constitute a reasonable choice of behavioral parameters that might reveal the influence of trial-by-trial variations in RA firing.

For each syllable, we defined a measurement time relative to syllable onset that corresponded to a well-defined spectral feature (Fig. 1*a*, band of spectral power at ~ 5.3 kHz in syllable B). Syllable onsets were defined based on amplitude threshold crossings. Pitch was defined as the fundamental frequency at the measurement time and was quantified by finding peaks in spectral power. Amplitude was defined as the value of the smoothed rectified amplitude trace at the measurement time. Spectral entropy was defined as the entropy of spectral power at the measurement time within one octave centered on the peak power. Entropy was quantified according to the equation $E = -\sum(p \log_{10} p)$, where p is the probability distribution of spectral power.

Premotor neural activity. Premotor activity was quantified by measuring the number of spikes occurring in a “premotor window” before the time at which the acoustic properties of each syllable were measured. The timing of this window was chosen to reflect the latency at which RA activity influences the acoustic structure of song. Stimulation studies have shown that disrupting ongoing activity in RA during song perturbs motor output, although different groups have produced varying assessments of the nature and latency of these effects (Vu et al., 1994; Ashmore et al., 2005). Stimulation with a single pulse, however, has been shown to disrupt the pitch of an ongoing syllable at a short (~ 15 ms) latency without altering the sequence of syllables being produced (Fee et al., 2004). To allow for some uncertainty about the premotor latency (and for the possibility that different acoustic parameters may have different latencies), we measured premotor neural activity in a 40-ms-wide window that ended at the time when acoustic parameters were measured. Premotor neural activity was measured by counting the number of spikes in this window. We validated this approach to quantifying neural activity by comparing several models of premotor encoding in which spiking activity is quantified on different timescales (supplemental text, Testing

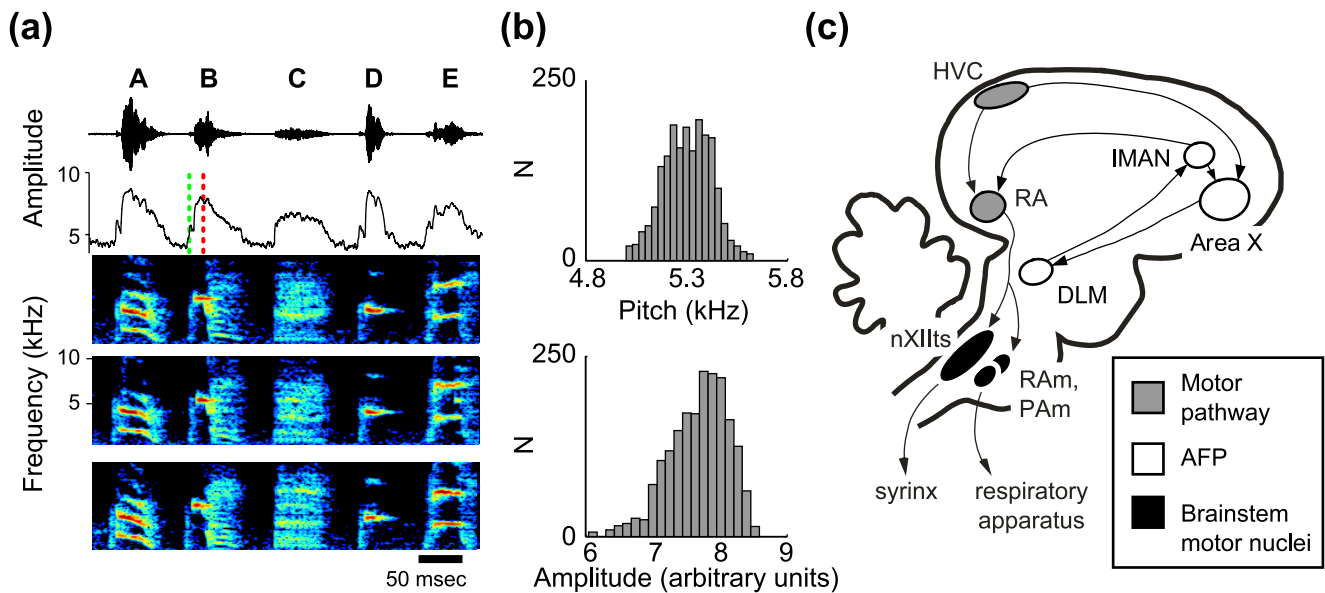


Figure 1. Acoustic variation in Bengalese finch song. *a*, Top, Raw sound amplitude waveform of five syllables (labeled A–E) from the song of an adult Bengalese finch. Middle, Smoothed rectified sound amplitude waveform. Bottom, Spectrograms of three different iterations of this motif. Spectrograms show the power at each frequency (color scale) as a function of time. The topmost spectrogram corresponds to the example sound waveform. Acoustic features such as pitch and amplitude were measured at a fixed time (red dashed line for syllable B) relative to syllable onset (green dashed line). *b*, Distributions of the pitch (top) and amplitude (bottom) of syllable B across 1919 renditions of the syllable. *c*, The song system is composed of a direct motor pathway consisting of nuclei HVC (used as a proper name) and RA and an AFP containing area X, the medial portion of the dorsolateral thalamus (DLM) and IMAN. RA sends projections to motor neurons in the tracheosyringeal portion of the 12th motor nucleus (nXIIts), which innervates the muscles of the syrinx, and to motor nuclei retroambigualis (RAM) and paraambigualis (PAm), which innervate the respiratory musculature (Vicario and Nottebohm, 1988; Wild, 1993; Reinke and Wild, 1998).

the timescale of premotor encoding, available at www.jneurosci.org as supplemental material).

Correlation analyses. To examine the relationship between premotor activity and acoustic output, we computed the linear correlation between each acoustic feature and the number of spikes in the premotor window. Before computing correlations, we discarded outliers with acoustic feature measurements lying >4 SDs from the mean. Inspection of audio recordings revealed that these outliers usually resulted from noise artifacts unrelated to vocal production. Additionally, we performed a partial correlation analysis in which the relationship between each acoustic parameter and neural activity was considered while controlling for correlations between neural activity and the other two acoustic parameters.

Proportion of cases correlated. We define the prevalence with which neural activity is correlated with a given acoustic feature in terms of the proportion of cases with significant correlations. One “case” is defined as one unit (that is, one single unit or one multiunit cluster) being active before one syllable. A unit is defined as active before a syllable if it fires on average at least one spike in the 40 ms premotor window, corresponding to a mean rate of ≥ 25 Hz. For a given acoustic parameter, therefore, one unit will contribute multiple cases if it is active before more than one syllable. For each acoustic parameter, we found the proportion correlated by dividing the number of cases in which the acoustic parameter was significantly correlated (at $p < 0.05$) with neural activity by the total number of cases.

Multiple comparisons. We expect that even if no relationship between premotor neural activity and song existed, some correlations would be significant by chance ($\sim 5\%$ of all correlations, corresponding to our significance criterion of $p < 0.05$). We used a permutation technique to quantify whether the observed proportion of correlations between neural activity and acoustic output was significantly greater than chance. Briefly, we created an artificial dataset in which all correlations of interest (those between premotor activity and acoustic features) are broken, but all other correlations (such as those between different acoustic measures and between neural activity on consecutive syllables) are preserved. We then performed correlation tests on the artificial dataset and noted the proportion of cases with significant correlations. By performing this procedure 1000 times, we estimated the distribution of proportions of significant correlations under the null hypothesis, and then asked whether

the proportion of significant correlations in the real dataset exceeded the 95th percentile of this distribution. For a detailed explanation of this technique and a discussion of related issues, see the supplemental text, Multiple comparisons (available at www.jneurosci.org as supplemental material).

Fraction of units active. To quantify the fraction of the population active at a given time during song, we considered only birds from which we had recorded at least 10 units ($n = 7$). The mean neural activity of each unit was quantified in 1 ms bins across the duration of a frequently occurring syllable sequence (motif). Repeating this analysis for each recorded unit allowed us to infer the mean activity across the population. The fraction active at a given time was defined as the percentage of recorded units with mean rate > 25 Hz. For the plot shown in Figure 4b, the fraction active is averaged over a sliding 5 ms time bin.

Time warping. Piecewise linear time warping (Leonardo, 2004) was used to create Figure 4a. Briefly, spike times were aligned to the mean durations of syllables and intersyllable pauses by “stretching” them linearly. These small adjustments allow for easy comparison of neural activity across trials and units by eliminating small (typically 1–6% in our data) variations in song tempo. Note however that time warping was used only for display purposes and was not applied as part of any of the analyses described above, in which the 40 ms premotor window was applied on a syllable-by-syllable basis.

Results

Our recordings revealed that neurons in Bengalese finch RA fire distinct patterns of activity for distinct syllables (as previously observed in the zebra finch), consistent with the accepted role of RA in controlling syllable structure (Fig. 2) (Yu and Margoliash, 1996; Leonardo and Fee, 2005). Close inspection of firing patterns, however, revealed trial-by-trial variations in the number of spikes in each burst (Fig. 3b). The analyses described below investigate whether this neural variation results in trial-by-trial variation in the acoustic structure of individual syllables.

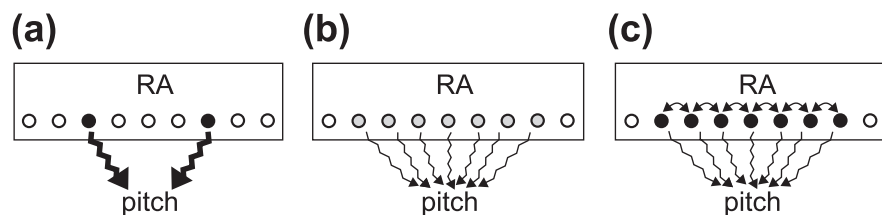


Figure 2. Three models of how RA might drive trial-by-trial variation in song. Circles represent populations of RA projection neurons, and wavy arrows represent their contribution to variations in syllable structure (pitch in this case). *a*, A small number of independently varying RA neurons each drive a substantial amount of pitch variation (heavy arrows). In this case, a small proportion of RA neurons would exhibit correlations with pitch, and these correlations would be quite strong (indicated by filled black circles), reflecting the substantial influence of each neuron. *b*, Large numbers of independently varying RA neurons each make small contributions to pitch variation (thin arrows), with pitch variations resulting from the sum of independent modulations in firing. In this case, the activity of many neurons would be correlated with pitch, but these correlations would be weak (indicated by the lightly shaded circles), because each neuron is responsible for only a small fraction of behavioral variation. *c*, Large numbers of RA neurons generate correlated activity (curved arrows), driving pitch modulations with coordinated changes in activity. In this case, many neurons might exhibit correlations with each acoustic feature, and the strength of these correlations could be quite high (black circles), because variations in the firing of any one cell are correlated with variations in the entire population. In Discussion, we consider the implications of our results in discriminating between these models.

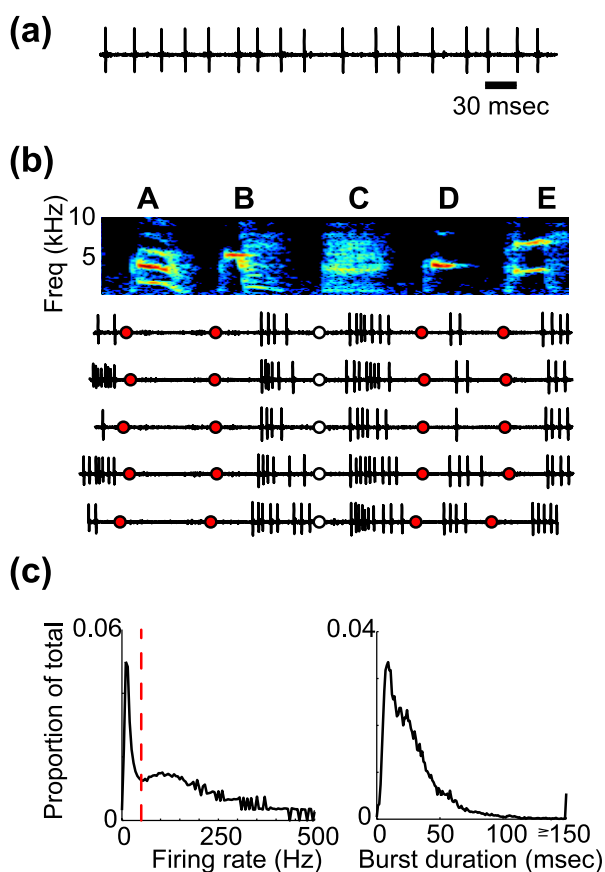


Figure 3. Neural activity in Bengalese finch RA. *a*, Extracellular recording of an RA single unit in a nonsinging, awake bird. *b*, Firing of the same unit during singing. Neural recordings are shown for five repetitions of the song motif “ABCDE,” shown in the spectrogram at top. The five neural traces are aligned to the onset of syllable C (white dots). Onsets of the other syllables are shown as well (red dots). Data in *a* and *b* are plotted with the same time scale. *c*, Left, Instantaneous firing rates (1/ISI) during song for all single units. A 50 Hz rate threshold (dashed red line) was used to segregate bursting from nonbursting epochs to compute the distribution of burst durations during singing (right).

Neural activity in Bengalese finch RA

Neurons in RA exhibited characteristic patterns of activity before, during, and after song. The majority of RA units displayed regular tonic activity of 20–50 Hz during rest (Fig. 3*a*) and fired syllable-

locked bursts during singing (Fig. 3*b*). After song offset, the resting tonic activity of most units was transiently inhibited (supplemental Fig. 7, available at www.jneurosci.org as supplemental material). Additionally, a small subset of recordings (one single-unit, three multiunit) had very low or no spiking activity when the bird was at rest, displayed bursty spiking activity during song, and had narrower spike widths than the rest of the population. Based on these criteria (Spiro et al., 1999; Leonardo and Fee, 2005), these four units (representing 3% of the total dataset) were classified as putative interneurons and were excluded from additional analysis (supplemental text, Putative interneurons vs putative projection neurons, available at www.jneurosci.org as supplemental material). The results that follow describe the properties of putative projection neurons.

The qualitative impression that RA neurons fire in bursts during singing was confirmed by the bimodal distribution of instantaneous firing rates (Fig. 3*c*, left), defined here as the inverse of ISIs. The peaks of the firing rate distribution (at ~12 and 110 Hz) correspond to periods between and within bursts, respectively. The trough between these peaks, which was centered at 50 Hz (Fig. 3*c*, left, red line), suggested a criterion for assigning ISIs either to an ongoing burst (if they were <1/50 Hz = 20 ms long) or to an interburst pause. Using this criterion, we computed the distribution of burst durations shown at right in Figure 3*c*.

Serially recording from many units in the same bird allowed us to investigate how premotor activity is distributed across the population of RA neurons. Figure 4*a* shows examples of spiking activity recorded from 25 units in a single bird. Spike times (colored tick marks) are aligned relative to the mean duration of syllables (gray boxes) and intersyllable pauses using piecewise linear time warping. Both single-unit and multiunit recordings displayed bursts of spikes throughout the song. As described in Materials and Methods, we computed the proportion of units active at each time during song. Figure 4*b* shows that this proportion varied considerably over time. Averaging across time and combining data across birds, we found that 58 ± 19% (mean ± SD) of the recorded population was active at any given time during singing.

Because of the temporally sparse nature of RA activity, not all units were active preceding all syllables. For example, unit 2 in Figure 4*a* (indicated by an asterisk at left, see also Fig. 3*b*) was not active before syllable B. When comparing neural variability to acoustic variability, we therefore restricted our analysis to cases in which the recorded unit was active before the syllable in question.

Prevalence and strength of neuron–behavior correlations

We quantified premotor spiking activity and three acoustic measures (pitch, amplitude, and spectral entropy) (see Materials and Methods) for each recorded unit and syllable. Previous studies have shown these three acoustic parameters to be under the control of the song system (Tchernichovski et al., 2001; Fee, 2002; Fee et al., 2004; Kao et al., 2005), and our PCA-based analysis shows that in most cases these features dominate the trial-by-trial spectral variation of each syllable (supplemental text, Quantitative analysis of spectral variability, available at www.jneurosci.org as supplemental material). Premotor neural activity was quantified

by counting the number of spikes in a 40 ms premotor window (Fig. 5*a*). Figure 5*b* shows the distribution of pitches of syllable E in bird 1. The insets in Figure 5, *b* and *c*, quantify pitch and neural variation using the coefficient of variation (CV; equal to SD/mean) and Fano factor (variance/mean). The fundamental question we address in this study is whether some of the observed behavioral variation can be explained by the variation in premotor neural activity illustrated in Figure 5*a* and quantified in Figure 5*c*. To do this, we measured the correlation between pitch and premotor neural activity (Fig. 5*d*). The example shown here yielded a highly significant ($p < 10^{-13}$) positive correlation with an r^2 value of 0.15 (indicating that premotor neural activity could account for ~15% of the behavioral variation).

We repeated this analysis for each acoustic parameter and for each case in which a recorded unit was active before a given syllable. The distributions of CVs and Fano factors across all syllables and neurons are summarized in Table 1. Figure 6 shows the distribution of significant correlations in bird 1 for each syllable, recorded unit, and acoustic feature (for similar representations of data from the other 12 birds in our study, see supplemental Fig. 9, available at www.jneurosci.org as supplemental material). In the example shown in Figure 6*a*, there were 91 cases (the total number of dots) in which units were active before a syllable. Of these 91 cases, significant correlations were found in 17.6% (16 of 91) of cases. These are indicated here as green dots for cases with significant positive correlations and red dots for cases with significant negative correlations. Because we expect that some comparisons would yield significant correlations by chance, we used a resampling-based technique (see Materials and Methods) (supplemental text, Multiple comparisons, available at www.jneurosci.org as supplemental material) to ask whether the proportion of significant correlations was itself significantly greater than chance. Applying this test to the example of pitch in bird 1, we found that the proportion of significant correlations was highly significant ($p < 0.001$).

When data from all birds were combined, the proportion of correlated cases was significantly different from chance for all measured acoustic parameters ($p < 0.001$ in all three cases). As shown in Figure 7*a*, premotor activity was correlated with pitch and amplitude in 26.1 and 26.6% of cases, respectively. Both of these parameters were correlated with premotor neural activity significantly more frequently than was entropy (20.8% of cases).

In principle, a given unit might have a consistent relationship to an acoustic feature (such as pitch) across all syllables for which it was active. This was not the case. Firing rates were typically correlated with an acoustic parameter in some contexts but not others. For example, unit 8 (Fig. 6*a*, black arrow) was positively correlated with pitch during syllables B and E, but not during

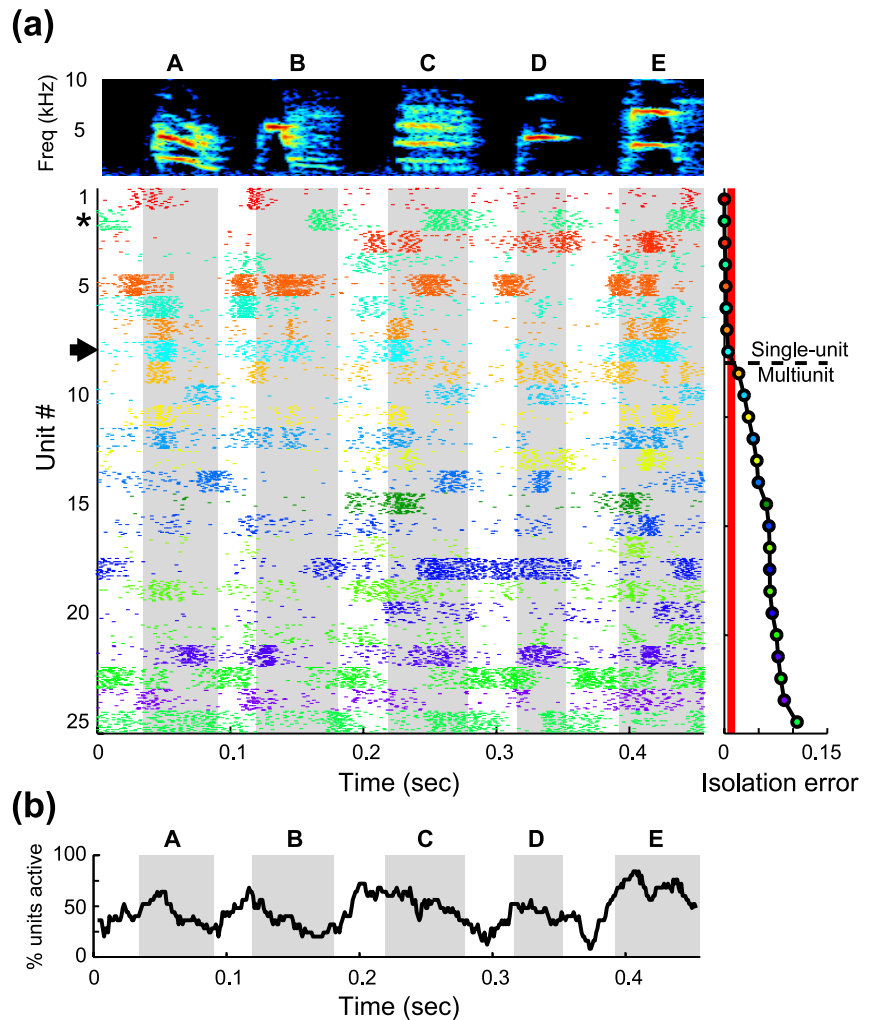


Figure 4. *a*, Population activity and unit isolation estimates of 25 units recorded from bird 1. Each tick mark represents one spike; each row represents the activity during one iteration of the song motif “ABCDE,” and 20 iterations from each unit are shown. Colors differentiate recordings from the 25 units. Mean syllable durations are shown as gray boxes. The unit shown in Figure 3, *a* and *b*, is unit 2 (asterisk at left), and the unit shown in Figure 5*a* is unit 8 (arrow at left). The plot at the far right shows the estimated isolation error (see Materials and Methods). Units with isolation errors of < 0.01 (red line) were classified as single units. *b*, Fraction of units active (mean rate, ≥ 25 Hz) as a function of time.

other syllables. Also, neural activity was sometimes correlated with different acoustic parameters in different contexts. This pattern can be seen in unit 9 (Fig. 6*a*, gray arrow), which was positively correlated with pitch during syllable E and negatively correlated with amplitude during syllable F. Finally, for a single acoustic parameter, some units displayed correlations of opposite signs during different syllables. For example, unit 19 (Fig. 6*a*, white arrow) was positively correlated with pitch during syllable B but negatively correlated during syllable C. Across the entire dataset, the sign of a significant correlation between premotor activity and an acoustic parameter during one syllable was not predictive of the sign of significant correlations during other syllables. For a fuller discussion of these issues, see the supplemental text, Sparse distribution of significant correlations (available at www.jneurosci.org as supplemental material).

The r^2 values for significant correlations showed that appreciable amounts of behavioral variability could be predicted from the activity of individual units. Figure 7, *b* and *c*, show the probability density functions and cumulative distributions of r^2 values from significantly correlated cases. For all acoustic parameters, r^2

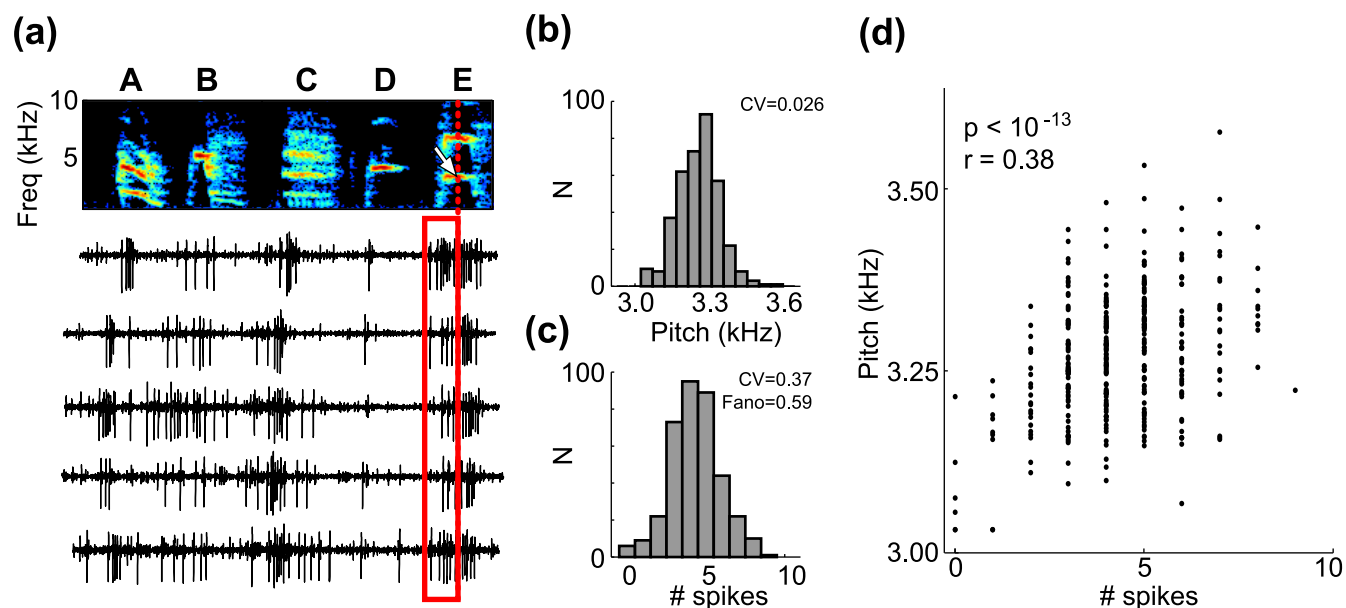


Figure 5. Correlation between neural and behavioral (pitch) variation. **a**, Spectrogram of a song motif from bird 1. Below the spectrogram are five neural traces showing the activity of unit 8 from this bird (Fig. 4*a*, arrow). Premotor neural activity for this unit was quantified in a 40 ms window (red box) preceding the time of pitch measurement (dashed red line). The white arrow indicates the measured pitch of syllable E. Neural traces are aligned on the onset of syllable E. **b**, Distribution of pitches measured from syllable E while recording from the unit shown in **a**. **c**, Distribution of the number of spikes within the premotor window preceding syllable E. **d**, Plot showing the relationship between pitch and the number of spikes in the premotor window for each rendition of syllable E ($n = 371$), along with the correlation coefficient (r) and p value of the correlation.

Table 1. Quantification of the degree of neural and behavioral variation

Measure	Mean \pm SD
CV	
Spike number, single units	0.46 ± 0.23 (across recordings sites and syllables)
Spike number, multiunits	0.60 ± 0.21
Pitch	0.053 ± 0.052 (across syllables)
Amplitude	0.056 ± 0.026
Entropy	0.22 ± 0.14
Fano factor	
Spike number, single units	0.66 ± 0.55
Spike number, multiunits	0.98 ± 0.41

Measures of spiking variation were calculated from the distribution of spike counts in the 40 ms premotor window. A separate CV (SD/mean) and Fano factor (variance/mean) were computed for each case in which a unit was active (mean rate, ≥ 25 Hz) (see Materials and Methods) before a syllable ($n = 705$ cases, mean \pm SD of each distribution is reported). Measures of behavioral variation were calculated by pooling measurements of pitch, amplitude, or entropy from all recorded renditions of each syllable ($n = 83$ syllables).

values were distributed densely below 0.15 and had a long tail of larger values extending beyond 0.25. Mean r^2 values for pitch, amplitude, and entropy were 0.08, 0.09, and 0.07, respectively. The prevalence (proportion of cases correlated) and explanatory power (r^2 values) of correlations with pitch, amplitude, and entropy show that variations in the activity of individual units can predict a substantial amount of trial-by-trial behavioral variation across a range of acoustic features. Together, these data indicate that a component of motor variation is centrally generated at the level of RA.

Additionally, we collected a small amount of data during female-directed song to compare the overall level of neural variation across social contexts (supplemental text, RA activity in directed versus undirected song, available at www.jneurosci.org as supplemental material). Previous studies have established that directed song is less variable than undirected song (Sossinka and Bohner, 1980; Kao et al., 2005; Kao and Brainard, 2006; Sakata et al., 2008), suggesting that trial-by-trial variability in RA activity might similarly be reduced in the directed condition. As shown in supplemental Figure 10, neural variability was indeed signifi-

cantly lower during directed song. Although preliminary, these results suggest that not only do trial-by-trial variations in RA activity drive variations in song (the main finding in our study), but also that modulations in the overall level of RA variability can account for social context-dependent changes in song.

The analyses presented here examine correlations between acoustic features and the amount of neural activity in an immediately preceding 40 ms premotor window. This window was chosen based on previous studies to reflect the likely causal delay between RA activity and the acoustic structure of song (see Materials and Methods). However, in principle, correlations between activity and behavior could extend across neighboring syllables. For example, if the pitches of successive syllables were serially correlated, then RA activity correlated with the pitch of one syllable would also be expected to correlate with the pitch of neighboring syllables. We found that serial correlations in behavior were indeed common in Bengalese finch song, and correspondingly, we sometimes found significant correlations between premotor activity measured for one syllable and the acoustic features of neighboring syllables. However, the prevalence of significant neuron–behavior correlations was sharply peaked for the neural activity occurring in the 40 ms window immediately preceding measured acoustic features (supplemental text, Correlations extend across time, available at www.jneurosci.org as supplemental material), indicating that neural activity in the premotor window used in this study had significantly more predictive power than neural activity preceding adjacent syllables.

Despite differences in unit isolation, single-unit and multiunit recordings were correlated with acoustic parameters in approximately equal proportions and yielded correlations with similar explanatory power. Across the entire dataset, the proportions of cases from single-unit and multiunit recordings with significant correlations (27.9 and 24.1%, respectively) were not significantly different (Z test for proportions, $p = 0.24$). Also, distributions of

r^2 values for the two classes of recordings were not significantly different for any of the three acoustic parameters (two-tailed t tests, smallest $p = 0.44$) or when all r^2 values were pooled ($p = 0.60$).

Positive correlations with pitch

Inspection of data from both single-unit and multiunit recordings revealed a strong asymmetry in correlations with pitch but not with other acoustic parameters. In the example shown in Figure 6*a*, positive correlations with pitch (green dots) were present in greater numbers than negative correlations (red dots). In five of five birds (including this one) where there was a significant difference in the number of positive and negative correlations with pitch, positive correlations outnumbered negative ones. When data were combined across all birds, this asymmetry was significant for data from both single-unit and multiunit recordings (Fig. 8*a*). This asymmetry in the sign of correlations with pitch supports a model of pitch production based on recent anatomical and physiological studies (see Discussion).

In the analyses presented thus far, the relationship between neural activity and each of three acoustic parameters was assessed in separate correlation tests. However, examination of our behavioral data revealed that, in many cases, the measured acoustic features were significantly correlated with each other (data not shown). Some of the neuron–behavior correlations revealed in our initial analysis might therefore arise because of correlations between behavioral measures. For example, a correlation between neural activity and syllable amplitude might result from the combined effects of a correlation between neural activity and pitch and a correlation between pitch and amplitude. To disambiguate the correlations between premotor activity and each acoustic parameter, we performed a partial correlation analysis, in which the relationship between neural activity and each behavioral measure was assessed while controlling for the influence of correlations with the other two acoustic parameters (see Materials and Methods). This alternate analysis yielded nearly identical results as the primary analysis (Fig. 8*b*).

Discussion

Our results show that trial-by-trial variations in RA activity predict a significant component of acoustic variation in song syllables. We found correlations between premotor activity and all three acoustic parameters examined, although correlations with pitch and amplitude were found significantly more often than correlations with spectral entropy. Additionally, correlations with pitch had a positive sign in a significant majority of cases. Together, these results provide strong evidence that trial-by-trial variations in syllable structure result in part from variations in the motor command.

By exploiting trial-by-trial variability at particular times dur-

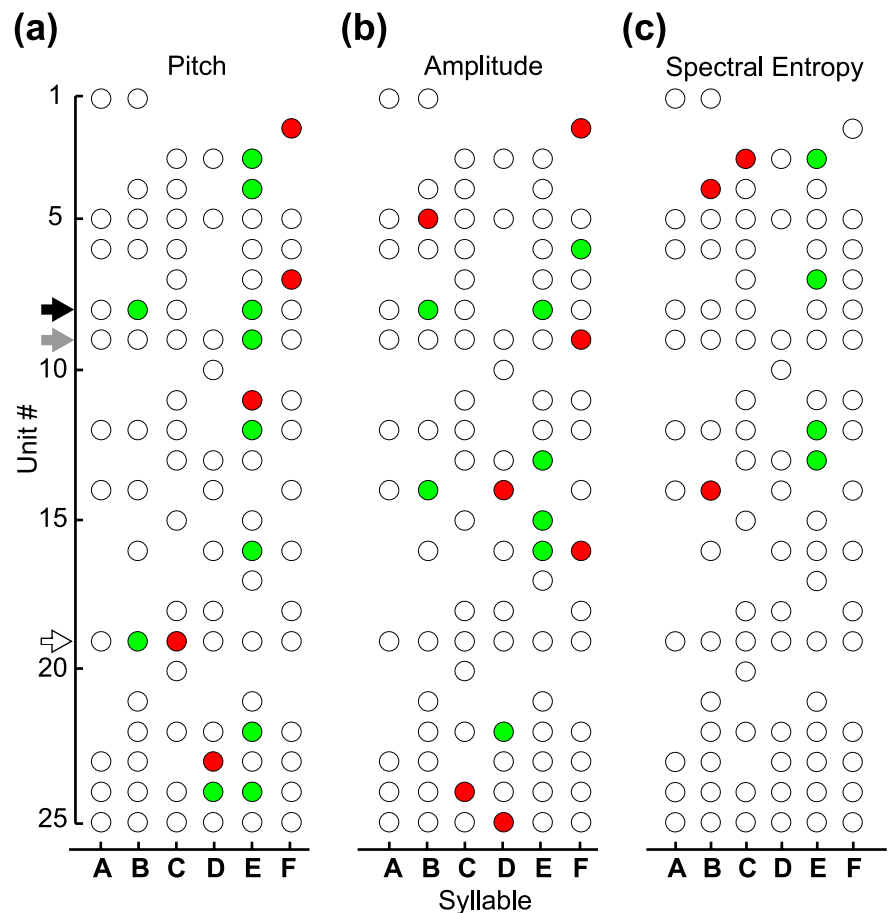


Figure 6. Significant correlations of premotor neural activity with acoustic structure across a population of RA units. *a*, Correlations with pitch for all units recorded in bird 1. Each row represents data from one single-unit or multiunit recording (see Fig. 4), and each column represents firing in the premotor window before one of the six syllables produced by the bird (as indicated at the bottom: A–F). A dot indicates that the unit in question was active (mean rate ≥ 25 Hz in the premotor window preceding a syllable). Dot color indicates whether neural activity was positively correlated (green), negatively correlated (red), or uncorrelated (white) with pitch. Black, gray, and white arrows indicate units 8, 9, and 19, respectively, as referred to in text. *b*, *c*, Same conventions as *a*, but dot color signifies correlations between premotor neural activity and amplitude (*b*) or spectral entropy (*c*).

ing song, our analysis provides the first description of covariation between RA activity and syllable structure. Using a contrasting approach, Leonardo and Fee (2005) found that mean population activity was on average uncorrelated with mean spectral output across different times in song, demonstrating that similar acoustic patterns can be produced by unrelated ensembles of RA cells. Our results are complementary to these previous findings. We show that within each ensemble of active neurons, trial-by-trial variations in activity can account for variations in behavior.

Comparison of our results with recent studies in primates suggests similarities between the neural control of birdsong and primate reaching movements. The activity of single neurons in motor, premotor, and parietal cortex is often correlated with multiple parameters describing reach kinematics (Fu et al., 1995; Buneo et al., 2002; Wang et al., 2007), just as RA neurons often appear to encode multiple acoustic parameters. Furthermore, although correlations between cortical activity and kinematic parameters (such as hand position or velocity) vary widely in strength, the r^2 values we report in RA fall within the range reported for several areas of motor and premotor cortex (Carmena et al., 2005; Stark et al., 2007). These similarities suggest that the generation of variable motor commands using populations of

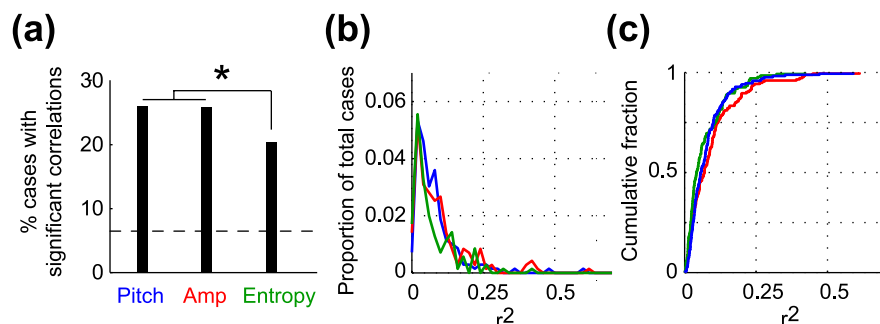


Figure 7. Prevalence and explanatory power of neuron–behavior correlations. **a**, Bar plot of the proportion of cases in which neural activity was significantly ($p < 0.05$) correlated with each of the three acoustic parameters measured. Note that one case corresponds to a unit’s being active in the premotor window before a single syllable, and that one unit can therefore contribute several cases to each acoustic feature. Black lines and asterisk above bars indicate that the proportions for pitch and amplitude were both significantly greater than the proportion for entropy ($p < 0.05$, Z test for proportions). The dashed line shows the significance threshold for the proportion of cases correlated exceeding chance at $p < 0.05$, as determined by a permutation test (supplemental text, available at www.jneurosci.org as supplemental material). All three acoustic parameters were correlated with premotor activity significantly more often than chance. **b**, **c**, Probability density functions (**b**) and cumulative distributions (**c**) of r^2 values for the three acoustic parameters, corresponding to the fraction of behavioral variability in each case that can be accounted for by neural activity. Color conventions are the same as in **a**.

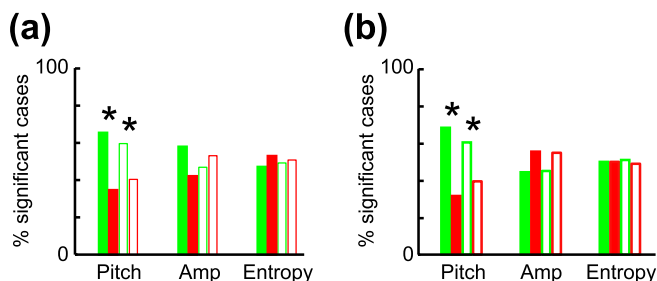


Figure 8. Correlation signs. Bars show the fractions of significant correlations with positive (green) and negative (red) slopes for single-unit (filled) and multiunit (empty) recordings. Asterisks indicate a ratio of positive to negative correlations significantly different from equality ($p < 0.05$, binomial test). **a**, Results from the primary analysis. **b**, Results from the partial correlation analysis (see Materials and Methods).

neurons each moderately correlated with several task parameters might be a general principle of skilled motor control.

Our results provide an initial characterization of RA in the Bengalese finch. Consistent with recordings in zebra finch RA (Yu and Margoliash, 1996; Leonardo and Fee, 2005), the neurons described here are tonically active at rest, fire syllable-locked bursts during song, and are transiently inhibited after song offset. In contrast, RA neurons in the Bengalese finch have a lower peak firing rate during bursts and fire bursts of greater duration (Fig. 3c) than their counterparts in the zebra finch (Leonardo and Fee, 2005).

Our analysis describes correlations between RA activity and acoustic output. However, the strength of such a correlation does not necessarily reflect the causal influence of a neuron. Neurons in RA might covary such that an increase in one cell’s firing is often accompanied by an increase in the firing of other cells, which make their own contributions to acoustic output. The measured correlation between neural activity and a behavioral parameter therefore depends both on the ability of the neuron to drive changes in behavior and on its correlation with other RA neurons.

At one extreme, neural activity could vary independently across RA neurons. In this case, the correlation between each neuron’s premotor activity and pitch (for example) would accu-

rately reflect the contribution of that cell to the total behavioral variability. If a small pool of independently varying neurons controlled pitch, these correlations would be strong, reflecting the small number of neurons governing behavior (Fig. 2a). Conversely, if pitch were controlled by a large number of independent neurons, correlations between activity and pitch would be weak (Fig. 2b).

At the other extreme, neural activity could be strongly correlated across many RA neurons. In this case, the measured correlation between any one cell and pitch would include the contributions of the entire correlated ensemble. These correlations could be quite strong, because they reflect the contributions of many neurons (Fig. 2c).

We can distinguish between these possibilities (Fig. 2a–c) by estimating the number of neurons that control acoustic

variation. In the zebra finch, the right and left RA each contain ~8000 neurons that project to brainstem motor nuclei (Gurney, 1981). Assuming a similar figure for Bengalese finch RA, which is of comparable volume to zebra finch RA (Tobari et al., 2005), we can estimate the number of neurons controlling pitch during each syllable. Of 16,000 total projection neurons, our data indicate that ~60% are active at any given time, and that of these ~25% make significant contributions to the control of pitch. Assuming that our recordings represent a uniform sampling, we can therefore estimate that approximately $16,000 \times 0.60 \times 0.25 = 2400$ RA projection neurons control pitch at any given time. (A similar figure is obtained for the number of neurons controlling amplitude, and a smaller number for spectral entropy, reflecting the smaller proportion of cases with significant correlations.) If each of these neurons contributed equally to pitch, then each would contribute $1/2400$ of the total behavioral variation (in the absence of downstream motor noise). If the activity of all neurons were independent, the measured correlation between each unit’s activity and pitch would have an r^2 value of $1/2400 = 0.000417$. Alternately, if RA neurons were strongly correlated, recording from any one neuron could have as much predictive power as recording from the entire population, and r^2 values would be far higher than those expected from an independent population.

The measured distribution of r^2 values (Fig. 7b,c) suggests that covariation between RA neurons is common. We found r^2 values far larger than expected from a population of independent neurons, with a mean r^2 value (0.08 for pitch) nearly 200 times larger than the value predicted by the independent-activity model (0.000417) shown in Figure 2b. Put another way, only 13 independent RA neurons with r^2 values at the mean of our observed distribution could in principle account for 100% of the behavioral variation. Because the number of neurons correlated with each acoustic parameter is far larger than this (~2400 neurons), some of the explanatory power of the measured correlations must arise from covariation between RA neurons (Fig. 2c), ruling out a model in which a small number of independent neurons drive behavioral variation (Fig. 2a). Covariation across RA might rely on networks of inhibitory interneurons that coordinate the activity of spatially separated projection neurons (Spiro et al., 1999). Although our calculations are based on rough estimates of neu-

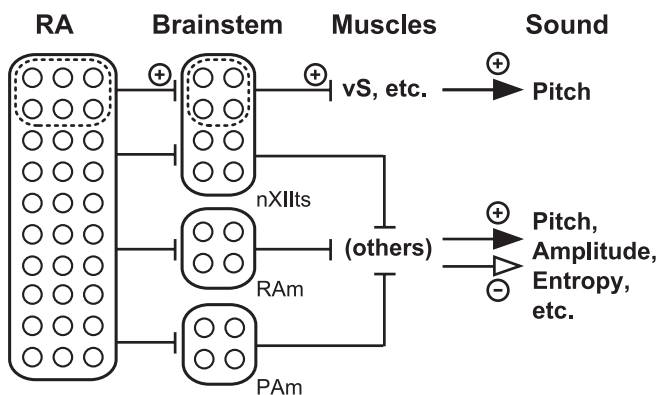


Figure 9. Acoustic control in the descending motor pathway. In our schematic, neurons are represented as white circles, synaptic connections are represented as lines connecting brain regions and muscles, and causal influences on acoustic structure are represented as arrows. RA projection neurons excite brainstem motor neurons in the tracheosyringeal portion of the 12th motor nucleus (nXIIIts), nucleus retroambiguus (RA_m), and nucleus paraambiguus (PA_m), which in turn activate the muscles controlling the syrinx and respiratory apparatus (Sturdy et al., 2003). Recordings from syrinx muscles reveal strong positive correlations between muscle activity and pitch, and one muscle in particular, the musculus syringealis ventralis (vS), has been shown in the brown thrasher to have an exceptionally strong association with pitch production (Goller and Suthers, 1996). Models of syrinx function suggest that increased muscular tension raises pitch by putting more tension on the vibrating structures of the syrinx, thereby increasing their vibrational frequency when air is blown through the syrinx (Gardner et al., 2001; Laje et al., 2002; Suthers and Zollinger, 2004). This string of excitatory relationships, between RA and nXIIIts, nXIIIts and muscle contraction in the syrinx, and between muscle contraction and pitch, likely results in a net excitatory relationship between RA activity and the pitch of song. The excess of positive correlations with pitch in our data might therefore be attributable to a subpopulation of RA projection neurons (far left, dashed box) that activate the motor neurons innervating the vS muscle of the syrinx (middle, dashed box) or other muscles for which activation drives increases in pitch. The approximately equal mixture of positive and negative correlations with other acoustic features might be attributable to the mix of positive (+) and negative (–) influences of other syrinx and respiratory muscles. Relative neuron numbers in the four nuclei are not shown to scale.

ron number and the prevalence of significant correlations, the difference between the empirical r^2 values and those expected from independent neurons is large enough to allow robust conclusions.

The prevalence and strengths of neuron–behavior correlations therefore point to a model of motor variation in which “cooperating” (that is, covarying) assemblies of a few thousand neurons produce trial-by-trial modulations of song (Fig. 2c). To the extent that acoustic variations are driven by the AFP (Kao et al., 2005; Olveczky et al., 2005), our results also suggest that IMAN drives coherent modulation of a pool of RA neurons rather than injecting independent noise across RA.

Although the activity of a single unit could account for as much as 40% of the variation in an acoustic parameter (Fig. 7b,c), the fraction of behavioral variation controlled by the entire RA population is unknown. Although RA is the sole output nucleus of the motor pathway, the brainstem motor nuclei controlling song additionally receive inputs from other parts of the brain (Wild, 2004), which may also contribute to premotor variation. Furthermore, peripheral motor noise presumably contributes to song variability as well. Note that if RA drives <100% of the behavioral variation, the r^2 value predicted by the independent-activity model would be even lower than 0.000417.

The observed predominance of positive correlations with pitch (Fig. 8) is consistent with the functional anatomy of the descending motor system. As schematized in Figure 9, it is likely that increases in RA activity ultimately result in a net increase in

the pitch of song. The observed surplus of positive correlations may therefore reflect a subpopulation of RA cells responsible for activating (via the brainstem) muscles that drive increases in pitch.

Our data suggest that birds modulate song by distributing variation across a few thousand neurons, thereby allowing them to explore the sensory consequences of varying the motor command. This motor exploration might be guided by differential reinforcement signals related to overall song quality (Tumer and Brainard, 2007). Alternately, by listening to these variations, adult birds could monitor the relationship between small changes in neural activity and small changes in acoustic structure. Knowing this relationship constitutes a local (that is, local to a single syllable) model of motor production. Maintenance of such a model might be necessary for the animal to adapt to changes in the strength of motor effectors as the bird ages or to changes in synaptic strength or connectivity over time. Song deteriorates dramatically when auditory feedback is removed in adulthood (Nordeen and Nordeen, 1992; Okanoya and Yamaguchi, 1997). Such deterioration might result from the inability of the bird to hear the consequences of motor exploration and thus maintain motor performance in adulthood.

References

- Arnold AP (1975) The effects of castration and androgen replacement on song, courtship, and aggression in zebra finches (*Poephila guttata*). *J Exp Zool* 191:309–326.
- Ashmore RC, Wild JM, Schmidt MF (2005) Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J Neurosci* 25:8543–8554.
- Buneo CA, Jarvis MR, Batista AP, Andersen RA (2002) Direct visuomotor transformations for reaching. *Nature* 416:632–636.
- Carmena JM, Lebedev MA, Henriquez CS, Nicolelis MA (2005) Stable ensemble performance with single-neuron variability during reaching movements in primates. *J Neurosci* 25:10712–10716.
- Chi Z, Margoliash D (2001) Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* 32:899–910.
- Churchland MM, Santhanam G, Shenoy KV (2006b) Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. *J Neurophysiol* 96:3130–3146.
- Fee MS (2002) Measurement of the linear and nonlinear mechanical properties of the oscine syrinx: implications for function. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 188:829–839.
- Fee MS, Kozhevnikov AA, Hahnloser RH (2004) Neural mechanisms of vocal sequence generation in the songbird. *Ann N Y Acad Sci* 1016:153–170.
- Fu QG, Flament D, Coltz JD, Ebner TJ (1995) Temporal encoding of movement kinematics in the discharge of primate primary motor and premotor neurons. *J Neurophysiol* 73:836–854.
- Gardner T, Cecchi G, Magnasco M, Laje R, Mindlin GB (2001) Simple motor gestures for birdsongs. *Phys Rev Lett* 87:208101.
- Goller F, Suthers RA (1996) Role of syrinx muscles in controlling the phonology of bird song. *J Neurophysiol* 76:287–300.
- Gurney ME (1981) Hormonal control of cell form and number in the zebra finch song system. *J Neurosci* 1:658–673.
- Hessler NA, Doupe AJ (1999) Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. *J Neurosci* 19:10461–10481.
- Kao MH, Brainard MS (2006) Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J Neurophysiol* 96:1441–1455.
- Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433:638–643.
- Laje R, Gardner TJ, Mindlin GB (2002) Neuromuscular control of vocalizations in birdsong: a model. *Phys Rev E Stat Nonlin Soft Matter Phys* 65:051921.
- Leonardo A (2004) Experimental test of the birdsong error-correction model. *Proc Natl Acad Sci U S A* 101:16935–16940.

- Leonardo A, Fee MS (2005) Ensemble coding of vocal control in birdsong. *J Neurosci* 25:652–661.
- Medina JF, Lisberger SG (2007) Variation, signal, and noise in cerebellar sensory-motor processing for smooth-pursuit eye movements. *J Neurosci* 27:6832–6842.
- Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol* 57:58–66.
- Okanoya K, Yamaguchi A (1997) Adult Bengalese finches (*Lonchura striata* var. *domestica*) require real-time auditory feedback to produce normal song syntax. *J Neurobiol* 33:343–356.
- Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3:e153.
- Reinke H, Wild JM (1998) Identification and connections of inspiratory premotor neurons in songbirds and budgerigar. *J Comp Neurol* 391:147–163.
- Sakata JT, Hampton CM, Brainard MS (2008) Social modulation of sequence and syllable variability in adult birdsong. *J Neurophysiol* 99:1700–1711.
- Sossinka R, Bohner J (1980) Song types in the zebra finch *Poephila guttata castanotis*. *Zeitschrift für Tierpsychologie* 53:123–132.
- Spiro JE, Dalva MB, Mooney R (1999) Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *J Neurophysiol* 81:3007–3020.
- Stark E, Drori R, Asher I, Ben-Shaul Y, Abeles M (2007) Distinct movement parameters are represented by different neurons in the motor cortex. *Eur J Neurosci* 26:1055–1066.
- Sturdy CB, Wild JM, Mooney R (2003) Respiratory and telencephalic modulation of vocal motor neurons in the zebra finch. *J Neurosci* 23:1072–1086.
- Suthers RA, Zollinger SA (2004) Producing song: the vocal apparatus. *Ann N Y Acad Sci* 1016:109–129.
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291:2564–2569.
- Tobari Y, Nakamura KZ, Okanoya K (2005) Sex differences in the telencephalic song control circuitry in Bengalese finches (*Lonchura striata* var. *domestica*). *Zool Sci* 22:1089–1094.
- Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of “crystallized” adult birdsong. *Nature* 450:1240–1244.
- Vicario DS, Nottebohm F (1988) Organization of the zebra finch song control system: I. Representation of syringeal muscles in the hypoglossal nucleus. *J Comp Neurol* 271:346–354.
- Vu ET, Mazurek ME, Kuo YC (1994) Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 14:6924–6934.
- Wang W, Chan SS, Heldman DA, Moran DW (2007) Motor cortical representation of position and velocity during reaching. *J Neurophysiol* 97:4258–4270.
- Wild JM (1993) The avian nucleus retroambigualis: a nucleus for breathing, singing and calling. *Brain Res* 606:319–324.
- Wild JM (2004) Functional neuroanatomy of the sensorimotor control of singing. *Ann N Y Acad Sci* 1016:438–462.
- Yu AC, Margoliash D (1996) Temporal hierarchical control of singing in birds. *Science* 273:1871–1875.