Calcium imaging in canary (serinus canaria) HVC reveals latent states supporting behavioral sequencing with long range history dependence

Yarden Cohen (yardenc@bu.edu)
Department of Biology, Boston University, 24 Cummington Mall
Boston, MA 02215 USA

Jun Shen (junshen@bu.edu)
Department of Psychological and Brain Sciences, Boston University, 610 Commonwealth Avenue
Boston, MA 02215 USA

Dawit Semu (dawits@bu.edu)
Department of Biology, Boston University, 24 Cummington Mall
Boston, MA 02215 USA

Timothy M. Otchy (totchy@bu.edu)
Department of Biology, Boston University, 24 Cummington Mall
Boston, MA 02215 USA

Timothy J. Gardner (timothyg@ bu.edu)
Department of Biology, Boston University, 24 Cummington Mall
Boston, MA 02215 USA

Abstract:
History dependent behavior is a key readout of neural processing. In skills, like speech or dance, motor sequences follow syntactic rules in which transitions between motor elements rely on past actions. Canary songs are defined by syllable repeats, called phrases, whose syntax exhibits long range order. The phrase sequence neural underpinnings must either rely on fixed action patterns or maintain historic context to influence ongoing transitions. To discriminate such mechanisms, we recorded Ca^{2+} signals from the premotor nucleus HVC in freely behaving canaries. We find that song history is reflected in identified ROIs up to 4 phrases apart, spanning up to 3 seconds and 40 syllables and that some ROIs exhibit mixed history selectivity. Moreover, we find that signals, reflecting sequence history information are more frequent during phrase transitions that are history dependent compared to history insensitive ones. These findings suggest that the network dynamics reflects historic context relevant to flexible transitions. Additionally, we find ROIs whose signals last several seconds and span 3-4 phrases. These signals are rarely modulated by syllable or phrase boundaries and initiate mostly during stereotyped sequences, suggesting distinct network dynamics during stereotyped and variable behavior.

Keywords: canary motor syntax; HVC imaging; hidden states

Background

Natural behavior provides an uncontrived readout of neural computations. Analyses of neural recordings, accompanying behavior, often require considering additional variables. These factors can be intrinsic, such as electrical and neuromodulator states affecting motion-related spiking (Schiemann et al., 2015), or extrinsic, such as stimuli that drive behaviorally relevant neural representations (Panzeri, Harvey, Piasini, Latham, & Fellin, 2017).

Here we pursue the idea that the behavior's dependence on its own history offers valuable constraints on the neural processing. These constraints can be mechanical or statistical limitations, recently used in relating motor plans to the previous movement and current limb position to reveal its premotor representation (Glaser, Perich, Ramkumar, Miller, & Kording, 2018). Alternatively, in behavior that can be reliably divided into basic elements (c.f. Wiltschko et al., 2015), the syntax of motor sequences provides a powerful tool for studying the neural basis of long range history-dependent behavior.

To pursue this approach, we take advantage of a uniquely rich and inherently segmented behavior - the song repertoire of the domestic canary (serinus canaria). All songbirds have a learned and naturally recurring behavior, song, whose temporal structure is largely governed by the premotor nucleus HVC (Hahnloser, Kozhevnikov, & Fee, 2002; Long & Fee, 2008; Nottebohm, Stokes, & Leonard, 1976; Wang, Herbst, Keller, & Hahnloser, 2008). Canary song repertoires are defined by units of repeated syllables, known as phrases (Figure 1a), whose syntax is controlled separately from the syllables' identity (Gardner, Naef, & Nottebohm, 2005) and which exhibit long range order (Markowitz, Ivie, Kligler, & Gardner, 2013). Specifically, certain phrase transitions have history dependence, originating 2-3 phrases upstream (Figure 1b). Accordingly, to support the generation of phrase sequences, the premotor neural activity must either rely on a complex memorized repertoire, receive directives from other brain regions, or maintain historic context to influence ongoing transitions.

To investigate such mechanisms, we recorded and analyzed $[Ca^{2+}]$ dynamics from canary HVC projection neurons. We discover two novel types of neural correlates –

latent, phrase-locked network states that reflect behaviorally relevant historic information and slowly decaying signals, much longer than the syllable and phrase time scales, that mostly accompany stereotyped phrase sequences but rarely initiate in variable parts of the songs. Together, these findings suggest new neural dynamics motifs to support complex, history-dependent cognitive functions.
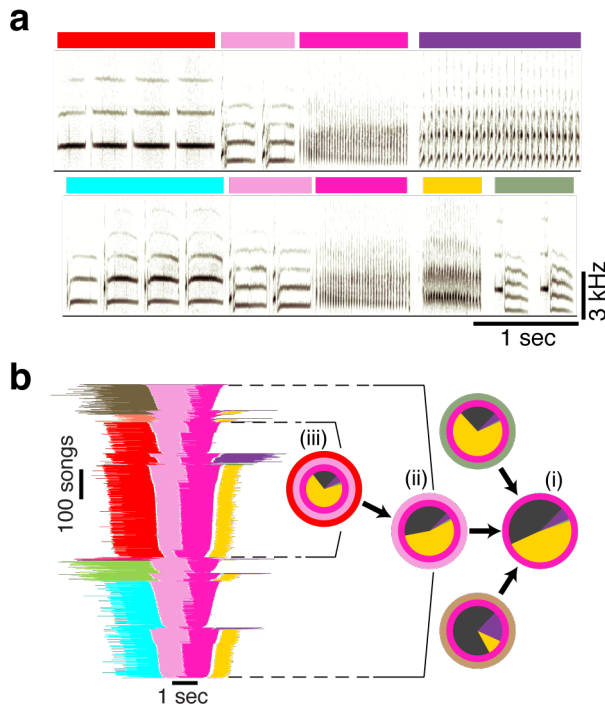
a

b

Figure 1: Canary song syntax has phrase transitions with long range history dependence. **a.** Canary phrases (color coded above the sonograms) are repeats of basic elements called syllables. **b.** Summary of phrase sequences that contain the pair, marked by light and bright pink in panel a and aligned to their center. Lines represent phrase identity and duration (color, length). Sequences are ordered by the identity of the 1st phrase, the last phrase and then the center phrases' duration. History dependence of the transition probabilities from the pink phrase (color coded pie charts, e.g. in (i) for the 1st order transitions) is determined sequentially by adding past phrases (2nd and 3rd order coded by concentric circles in (ii),(iii), corresponding to subsets marked by dashes lines) that significantly affect those probabilities.

Results

The domestic canaries' springtime songs have stable syllables and phrase boundaries – allowing for considerable automation of their segmentation and annotation and working on large datasets (>5000 songs from 3 birds were pre-processed with an in-house developed machine learning based classification algorithm). The birds' repertoire included 24-37 different syllables with typical durations between 10-350 msec. The average number of syllable repeats per phrase type ranges from 1 to 38 with individual phrases exceeding 10 seconds and 120 syllables. Transitions between phrases can be highly stereotyped, where one phrase type always follows another or diverge, allowing a certain phrase to precede a set of other phrases (Figure 1).

To investigate neuronal signals that reflect flexible behavior with long-range dependencies we used head-mounted fluorescence microscopes and microphones to image the genetically-encoded calcium indicator GCaMP6f in HVC projection neurons of freely-behaving adult male canaries (n=3, age > 1yr). With this technique (adapted from Liberti et al., 2016), we cast a wide net, observe all neurons in the field of view, including sparsely participating ones, and characterize the behavior correlates of annotated regions of interest (ROIs)

HVC projection neurons reflect long order historic information

We find signals that are time-locked to phrases but do not simply reflect the current behavior. Figure 2a shows the Ca^{2+} signal raster from a ROI, locked to the phrase types, marked in pink and related to the identity of the 2nd upstream phrase (a 2nd order correlation). This sequence preference is quantified by integrating the ROI-averaged signal (Figure 2b, 1-way ANOVA, p < 1e-7).

Importantly, despite keeping the sequence of phrases fixed, the phrase durations and the number of syllables in each phrase vary. In a previous study it was shown that phrase durations also correlate to transition probabilities and, in principle, could communicate the sequence information. Here, the relation to the 2nd upstream phrase identity, in Figure 2a, remains significant when taking these confounds into account (1-way ANOVA of residuals, p<0.03). Namely, these signals do not tightly follow the motor sequence but communicate the identity of a prior state, separated in time by multiple syllables.

The global time-in-song can also be a major contributor to seemingly long-range correlations. Specifically, processes such as monotonous neuromodulator buildup or ongoing adaptation to auditory feedback are potential mechanisms, relating song time to transition probabilities. When including the time-in-song as a confounding variable, as well as the relative timing of sequence phrase edges (in total 4 variables for 2nd order relations) 29% (25/85 from 3 birds) of ≥2nd order relations and 60% (104/172 from 3 birds) of the 1st order relations remain significant.

Latent network states reflect relevant historic information in variable phrase sequences

If the order of phrases is indeed controlled during the song, then the identity of historic sequence elements needs to be available for biasing downstream transition probabilities. Oftentimes, 2-3 phrases, lasting ≥ 2 seconds and involving 10s of syllables, separate between the transition and the relevant upstream event. Still, throughout a fixed sequence of four phrases, we find ROIs that carry information about an upstream phrase identity (Figure 2c).
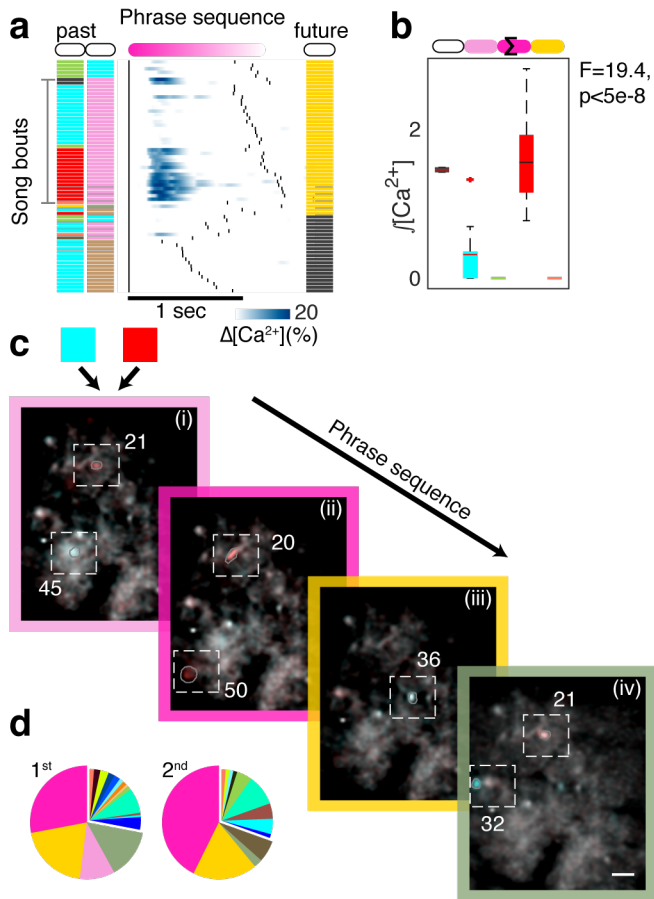
Figure 2: Calcium signals from ROIs in HVC correlate to identity of nonadjacent phrases across variable sequences. **a.** ROI related to 2nd upstream phrase identity. Ca^{2+} traces are aligned to the onset of one phrase type (Pink, marked by Σ in panel b). Trials (y-axis) are arranged by the following phrase identity (right, color coded), then by the phrase sequence history (left, color patches), and by duration of the pink phrase. Ticks mark phrase onsets. **b.** Distributions of Ca^{2+} integrals (y-axis, summation in pink phrase) for various phrase sequences that vary only in the 2nd upstream phrase in panel a (x-axis). F-number and p-value show the significance of separation by history (1-way ANOVA). **c.** A sequence of four phrases (i-iv, color coded) is preceded by two upstream phrase types (red or cyan). Average maximum projection images during each phrase in the sequence (i-iv) are colored by history (red or cyan) and overlaid. White dashed boxes mark ROIs that reflect historic information (color bias). **d.** Fraction of significant ANOVA tests relating signals in each phrase type (colors) and sequence elements. Pie charts separate 1st order (transitions) and higher order ($\geq 2^{nd}$). Pie slices of larger diameter mark phrase types that participate in transitions with long range history dependence (Figure 1).

In examining the types of phrases, during which sequence correlations occur, we find that the phrase types, participating in behavioral high-order transitions (in Figure 1), are overrepresented in holding sequence correlates of both 1st or higher orders (~70%, Figure 2d, i.e. exactly where they're needed). Additionally, we find both ROIs that correlate to a single transition type and ROIs with mixed history selectivity, showing elevated calcium signals in a subset of upstream phrase sequences.

Stereotyped song progression is accompanied with sustained signals, lasting several phrases

The phrase-type-locked Ca^{2+} signals, described so far, decayed within the phrase time or shortly afterwards. Considering the indicator's dynamics, these signals are in line with HVC projection neurons bursting locked to specific syllable types or specific transitions (Fujimoto, Hasegawa, & Watanabe, 2011). However, we also find ROIs with markedly longer signals (Figure 3a). These signals ramp relatively quickly, within a single phrase, but decay very slowly – spanning 2-3 phrases and lasting 2-4 seconds. Focusing on signal decay periods that exceed the expected GCaMP6f time constant, we find that some of these ROIs initiate slowly-decaying signals in more than a single phrase type and that slowly-decaying signals are usually stronger, relative to the shorter transients (as in Figure 2). Additionally, these slow signals usually do not fluctuate with syllable and phrase boundaries (Figure 3a).

Importantly, these signals are found throughout the data acquisition period (indicating stability up to 4 months after the implant), mostly initiate in phrases that reliably follow the previous phrase, typically appearing earlier in the song (Figure 3b,c), and, in some cases, reflect information about the phrase sequence.
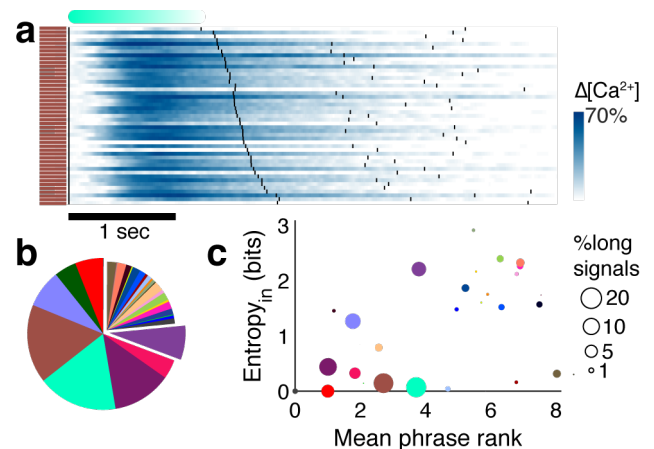

Figure 3: Long Ca^{2+} signals span multiple phrases in stereotyped sequences. **a.** ROI with long Ca^{2+} signal. Traces are aligned to the onset of one phrase type (turquoise, always follows brown). Trials (y-axis) are arranged by phrase duration. Ticks mark phrase onsets. **b.** Fraction of slowly decaying Ca^{2+} transients that appear in each phrase type (colors). Slices with larger (smaller) diameter are $\geq 4\%$ ($\leq 2\%$). **c.** The mean position (rank, x-axis) of each phrase type (colors, matching panel b) is compared to its degree of stereotypy (how reliably it follows the preceding phrase type,

input entropy, y-axis). Marker size matches the % in panel b (scale code in legend).

Discussion

Motor sequences with long range order dependencies are key motifs in history dependent behaviors, like language, whose supporting neural mechanisms are largely unknown.

We observe that the prior state or past syllable information is revealed in calcium activity during a fixed sequence of four canary phrases – showing that Ca^{2+} signals do not just reflect the current state or current transition. (e.g. in Fujimoto et al., 2011). These properties, as well as the existence of neurons with calcium activity, locked to the same phrase type in a subset of phrase sequence histories, may be a signature of new forms of phrase-level hidden states in HVC that, with further investigation, will allow us to refine models of syntax control for species that sing complex songs.

Additionally, we found neurons whose Ca^{2+} signals' decay time spans several phrases. These signals are rarely modulated by phrase or syllable boundaries and mostly appear in early and stereotyped parts of the sequence. We do not yet know what cell types express the properties described here, and a small fraction of neurons infected with Lentivirus, used for calcium indicator expression, are likely newborn cells that have not yet acquired their adult firing behaviors. In HVC of zebra finches, the vast majority of neurons infected with the Lentivirus used here are projection neurons (Markowitz et al., 2015). If the novel firing patterns described here belong to HVC_X or HVC_{RA} projection neurons (projecting to Area X and RA, the main two targets of HVC), this activity departs from previously reported phasic firing patterns of HVC_{RA} and HVC_X neurons. Future experiments will determine if we are observing a new class of cell, the occasional newborn neuron, or intriguingly, a new higher order form of song coding in projection neurons in canaries. The calcium activity reported here is among the first neural data recorded from singing canaries – a species whose complex song may shed new light on the neural circuit principles underlying the broad variations in song among the 4,500 species of songbirds that learn to sing by imitation.

That long order syntax generation is reflected by hidden network states and signals with qualitatively different time scales suggest that behavioral sequencing could be specified by sustained network dynamics in HVC. Perhaps, these network motifs reflect a rudimentary form of working memory that bridges past and present to create the long-range rules that govern canary song. Investigating the underlying mechanisms may reveal neural substrates that also exist in other motor and cognitive functions such as inference and decision making.

Acknowledgments

References

Fujimoto, H., Hasegawa, T., & Watanabe, D. (2011). Neural Coding of Syntactic Structure in Learned Vocalizations in the Songbird. *The Journal of Neuroscience*, *31*(27), 10023–10033.

Gardner, T. J., Naef, F., & Nottebohm, F. (2005). Freedom and Rules: The Acquisition and Reprogramming of a Bird's Learned Song. *Science*, *308*(5724), 1046–1049.

Glaser, J. I., Perich, M. G., Ramkumar, P., Miller, L. E., & Kording, K. P. (2018). Population coding of conditional probability distributions in dorsal premotor cortex. *Nature Communications*, *9*(1), 1788.

Hahnloser, R. H. R., Kozhevnikov, A. A., & Fee, M. S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*, *419*(6902), 65–70.

Liberti Iii, W. A., Markowitz, J. E., Perkins, L. N., Liberti, D. C., Leman, D. P., Guitchounts, G., … Gardner, T. J. (2016). Unstable neurons underlie a stable learned behavior. *Nature Neuroscience*, *19*(12), 1665–1671.

Long, M. A., & Fee, M. S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*, *456*(7219), 189–194.

Markowitz, J. E., Iii, W. A. L., Guitchounts, G., Velho, T., Lois, C., & Gardner, T. J. (2015). Mesoscopic Patterns of Neural Activity Support Songbird Cortical Sequences. *PLOS Biol*, *13*(6), e1002158.

Markowitz, J. E., Ivie, E., Kligler, L., & Gardner, T. J. (2013). Long-range Order in Canary Song. *PLOS Comput Biol*, *9*(5), e1003052.

Nottebohm, F., Stokes, T. M., & Leonard, C. M. (1976). Central control of song in the canary, Serinus canarius. *The Journal of Comparative Neurology*, *165*(4), 457–486.

Panzeri, S., Harvey, C. D., Piasini, E., Latham, P. E., & Fellin, T. (2017). Cracking the Neural Code for Sensory Perception by Combining Statistics, Intervention, and Behavior. *Neuron*, *93*(3), 491–507.

Schiemann, J., Puggioni, P., Dacre, J., Pelko, M., Domanski, A., van Rossum, M. C. W., & Duguid, I. (2015). Cellular Mechanisms Underlying Behavioral State-Dependent Bidirectional Modulation of Motor Cortex Output. *Cell Reports*, *11*(8), 1319–1330.

Wang, C. Z. H., Herbst, J. A., Keller, G. B., & Hahnloser, R. H. R. (2008). Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biology*, *6*(10), e250.

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., … Datta, S. R. (2015). Mapping Sub-Second Structure in Mouse Behavior. *Neuron*, *88*(6), 1121–1135.