The Neural Basis of Birdsong Syntax

Dezhe Z. Jin *
Department of Physics, The Pennsylvania State University,
University Park, PA 16802, USA

August 28, 2013

Abstract

Many complex actions such as speaking and playing music consist of sequences of action units strung together according to syntactic rules. How such rules are implemented in the brain is a critical problem for understanding the neural basis of complex behaviors. Songbirds are ideal model systems for tackling this problem. Birdsong consists of sequences of stereotypical syllables. The syntax of the syllable sequences ranges from simple to complex in different species. Many species such as Bengalese finch sing songs with variable sequences with probabilistic transitions between the syllables. Neural circuits underlying the singing behavior are well characterized, and are accessible to experimental manipulations and detailed computational modeling. Experimental and modeling works have established that syllables are encoded in unidirectional chain networks of projection neurons in the songbird premotor nucleus HVC (used as a proper name). Spike propagation along a chain network drives downstream neurons and produces a specific syllable. Computational models suggest that the song syntax can be embedded in the connection patterns of the syllable encoding chains. Allowed transitions from a syllable to other syllables are encoded by connecting the chain networks associated with the syllables into branched patterns. Through a winner-take-all mechanism enforced through the local inhibitory circuit and noise, a single chain is selected to propagate the spikes at a branching point, producing a probabilistic syllable transition. Probabilistic state transition models inspired by the network models can accurately describe the statistical properties of observed Bengalese finch song sequences, supporting the network mechanism of birdsong syntax.

Introduction

Sequences of actions are fundamental to many animal and human behaviors, including locomotion, vocal communications, language, music, dancing, martial arts, and logical reasoning. In a seminal paper in 1951 (Lashley (1951)), Lashley observed that sequential behaviors often

*Email djin@phys.psu.edu, Published in *Progress in Cognitive Science: From Cellular Mechanisms to Computational Theories*, Edited by Zhong-lin Lu and Yuejia Luo, Peking University Press (2013).

follow some "action syntax", similar to language. Using error patterns in speech and typing, he argued that such syntax must be generated within the central nervous system without moment to moment feedback from the periphery. Numerous experimental and theoretical works that followed Lashley, typically on human and non-human primates performing learned serial movements (Rhodes et al. (2004); Rosenbaum et al. (2007); Tanji (2001); Averbeck et al. (2002); Ohbayashi et al. (2003)), provided insights into the neural mechanism of action syntax. However, most of these works are based on simple, artificially designed, easily learnable sequences with a few elementary actions. How complex action sequences are generated and controlled in the nervous system remains largely unexplored

Typically, human subjects are required to memorize a number of short sequences consisting of 4 to 10 items and reproduce them verbally or through typing (Baddeley (1968); Page & Norris (1998); Sternberg et al. (1978); Verwey & Dronkert (1996); Salthouse (1986)), or produce short sequences guided by external cues (Rosenbaum et al. (1984); Verwey (2001); Nissen & Bullemer (1987); Keele et al. (2003)). By examining the types and rates of errors, as well as timing of the action elements in various conditions, a number of cognitive models of how sequences are stored in memory and retrieved during production have been proposed (Lashley (1951); Rhodes et al. (2004); Shallice (1972); Estes (1972); Grossberg (1978); Rumelhart & McClelland (1982); Rosenbaum et al. (1984); Keele et al. (2003)). The models argued for hierarchical organization of sequential memory, parallel activation of action elements before the initiation of the sequence, formation of chunks when subsequences are well practiced, and division of implicit and explicit learning of serial movements. These behavioral studies have been adapted to train non-human primates to perform serial movements with eyes, hands or arms while activities of single units are recorded (Mushiake et al. (1991); Tanji (2001); Nakamura et al. (1998); Clower & Alexander (1998)). These studies have revealed sequence and position selective activities in frontal lobe cortical areas, including the supplementary motor area and pre-supplementary motor area. Interestingly, neural activity in the prefrontal cortex while monkeys engaged in drawings of 3 to 4 segments supports the parallel activation model (Averbeck et al. (2002)). Brain areas engaged in learning and execution of serial movements are also studied using imaging techniques (Hikosaka et al. (1996); Sakai et al. (1998); Destrebecqz et al. (2003)), lesion studies (Miyachi et al. (1997); Lu et al. (1998)), and patients with neurological disorders (Ferraro et al. (1993); Vicari et al. (2003)). Although these studies give valuable insights into mechanisms of sequential actions, the simplicity of the behavioral sequences prevents a complete understanding of the neural basis underlying complex action sequences. Moreover, due to the complexity of the primate brain, the neural circuitry involved, even at the level of brain areas, has yet to be completely mapped out; this further hinders the progress in understanding the neural mechanisms.

The songbird is an excellent model system for studying the neural mechanisms of production and learning of action syntax. Birdsong is a learned vocalization that has many parallels with human language (Thorpe (1958); Marler (1970); Immelmann (1969); Price (1979); Doupe & Kuhl (1999); Williams (2004)); songs of many species, such as Bengalese finch, have complex syntactical structures that are rudimentarily similar to human speech (Okanoya (2004)). Figure 1a shows the spectrogram of the song of a Bengalese finch. The song consists of discrete units - bursts of sounds separated by silent intervals. These discrete units are called syllables. The song in Fig.1a has seven different syllables which are labeled with letters on top of the spectrogram. The syllable sequence is complex, and can be described by a transition rule that allows a syllable to be followed by another one chosen with a certain probability from a restricted set of syllables

(Dietrich (1980); Okanoya & Yamaguchi (1997); Woolley & Rubel (1997); Honda & Okanoya (1999); Okanoya (2004); Sakata & Brainard (2006)) (Fig.1b). Such transition rules are similar in an elementary way to how words can be strung together with restricted flexibility, following grammatical rules, in language (Jurafsky & Martin (2000)). Like grammar, these transition rules allow generation of an unlimited number of distinctive syllable sequences.

Progress in experimental techniques and theoretical work on songbirds provides a unique opportunity to significantly advance our understanding of the neural mechanism of complex action sequences. Since the 1970s, many aspects of the neural control of birdsong have been elucidated. The brain structures responsible for production and learning of song, collectively known as the song control system, have been identified with anatomical and lesion studies (Nottebohm et al. (1976, 1982)) (Fig.2). *In vitro* and *in vivo* intracellular recordings revealed properties of neurons and their connectivity in a number of key nuclei in the song system (Dutar et al. (1998); Kubota & Taniguchi (1998); Spiro et al. (1999); Mooney (2000); Sturdy et al. (2003); Wild et al. (2005); Mooney & Prather (2005); Kubke et al. (2005); Solis & Perkel (2005)). *In vivo* recordings demonstrated the spike patterns of neurons in singing and sleeping birds (Yu & Margoliash (1996); Vu et al. (1998); Dave & Margoliash (2000); Hahnloser et al. (2002); Schmidt (2003); Leonardo & Fee (2005); Hahnloser et al. (2006)). These advances make the study of birdsong a fertile ground for developing biologically constrained and detailed computational models that are also predictive, and for the productive interplay between the experiments and the models.

Simple organisms often provide opportunities to precisely understand neural mechanisms of motor control. Over the years, electrophysiological investigations on rhythmic behaviors such as swimming, walking, and digestion in animals such as lamprey, frog, and lobster have yielded important concepts, including the notion of the central pattern generator (CPG) (Marder (2000)), and afferent and sensory feedback regulation of pattern generation (Pearson (1993, 2000)). Research on songbirds, a much simpler system compared to primates, should similarly yield concrete insights on the neural mechanisms of complex action sequences.

The song system

The song control system of songbird consists of a set of brain nuclei linked to form a mostly feedforward excitatory pathway (Nottebohm et al. (1976, 1982); Vicario & Nottebohm (1988); Wild (1997)) (Fig.2). The premotor nucleus HVC plays a key role in the song system. HVC projects to RA (the robust nucleus of the arcopallium). RA, in turn, projects to a hypoglossal motor nucleus containing motor neurons innervating the syrinx - the vocal organ of birds. HVC, RA and the hypoglossal motor nucleus are necessary for song generation and form a motor pathway in the song control system.

HVC is also indirectly connected to RA via the anterior forebrain pathway (AFP), which is crucial for song learning (Bottjer et al. (1984); Sohrabji et al. (1990); Williams & Mehta (1999); Brainard & Doupe (2000); Olveczky et al. (2005)). The AFP includes area X, which is a homologue of mammalian basal ganglia, the thalamic nucleus DLM, and ultimately LMAN (the lateral magnocellular nucleus of the nidopallium). This pathway is homologous to the mammalian basal ganglia-cortical loop (Luo et al. (2001); Farries & Perkel (2002); Farries (2004)).

Besides being a key premotor area, HVC is a site of sensorimotor integration: It gets audi-

tory input from NIF (the nucleus interfacialis of the nidopallium) (Fortune & Margoliash (1995); Vates et al. (1996); Janata & Margoliash (1999); Cardin et al. (2005)) and CM (caudal mesopallium) (Bauer et al. (2008)); it also gets input from UVA (the nucleus uvaeformis), a thalamic nucleus (Nottebohm et al. (1982); Williams & Vicario (1993)). HVC has a rich internal structure (Dutar et al. (1998); Kubota & Taniguchi (1998); Mooney (2000); Wild et al. (2005); Mooney & Prather (2005)). There are at least three types of neurons: HVC(RA) neurons, which project to RA; HVC(X) neurons, which project to area X; and inhibitory interneurons, which do not project out of HVC. There are extensive connections between these neurons (Mooney & Prather (2005)).

The roles of the various nuclei can be partially inferred from the organization of the song control system. Motor neurons in the motor areas directly innervate muscles in the syrinx and in the abdomen and lung that control airflow through the syrinx (Wild (1997); Suthers & Zollinger (2004); Goller & Cooper (2004)). RA projects to the motor areas, and shapes the spiking patterns of the motor neurons. Similarly, HVC shapes spike patterns in RA; and UVA, CM, and NIF influence patterns in HVC. Anatomy alone, however, is insufficient to explain why HVC is needed in addition to RA for patterning spike activity of the motor neurons. Experiments in RA and HVC of singing zebra finch helped to resolve this issue (Yu & Margoliash (1996); Hahnloser et al. (2002); Schmidt (2003); Leonardo & Fee (2005)).

Song syllables and synfire chains

The syllables in birdsongs do not change much each time they are produced. The neural mechanism of such stereotypy is best studied in zebra finch. The song of a zebra finch consists of several repetitions of a motif, which is a fixed sequence of a few syllables. Recordings in RA and HVC in singing zebra finches revealed that neurons in RA and in HVC have different spiking patterns (Yu & Margoliash (1996); Hahnloser et al. (2002); Leonardo & Fee (2005)). RA neurons that project to motor areas spike reliably with precise timing relative to acoustic features of the motif (Yu & Margoliash (1996); Leonardo & Fee (2005)). An RA neuron bursts about 10 times during a motif. Different combinations of RA neurons burst at different times, driving different spike patterns in the motor neurons to generate varying acoustic features. In contrast, HVC(RA) neurons have very sparse spiking patterns (Hahnloser et al. (2002)). An HVC(RA) neuron bursts only once during a song motif. The burst typically lasts about 6 ms, emitting 3-7 spikes, and has a precise timing relative to the motif. Different HVC(RA) neurons burst at different times. Thus, HVC(RA) neurons burst sequentially and form a sequencer that drives different combinations of RA neurons at different times (Hahnloser et al. (2002); Fee et al. (2004); Kozhevnikov & Fee (2006)). An RA neuron bursts several times during the motif since it is driven by different HVC(RA) neurons that burst at different times. These experiments show that, whereas RA neurons encode moment-to-moment patterns directly involved in producing the acoustic features of the song, HVC(RA) neurons encode the sequence of these features (Fee et al. (2004); Kozhevnikov & Fee (2006)). The stereotypy of song syllables are encoded in the precisely timed sequential spiking of HVC(RA) neurons.

How is the HVC(RA) activity pattern generated? An experiment that cooled HVC of zebra finch during singing conclusively demonstrated that HVC(RA) bursts are not driven by external inputs that are precisely timed themselves; rather, they are generated intrinsically within HVC

(Long & Fee (2008)). A simple intrinsic model is the synfire chain networks, in which successive groups of HVC(RA) neurons are connected unidirectionally (Fig.3a) (Doya & Sejnowski (1999); Li & Greenside (2006); Jin et al. (2007); Long & Fee (2008)). Burst spikes initiated in the first group propagate down the chain through the excitatory synapses, producing sequential spiking of HVC(RA) neurons. Chaining is the simplest network structure that can generate sparse, sequential firing patterns, and has been suggested as a mechanism for sequential order in general (Fig.3a) (James (1890); Hebb (1949); Abeles (1982); Amari (1972); Kleinfeld (1986); Sompolinsky & Kanter (1986); Abeles (1991); Diesmann et al. (1999)). Neurons in the same groups tend to spike in synchrony (Abeles (1982, 1991); Diesmann et al. (1999)). The microcircuit of HVC (Mooney & Prather (2005)) supports the possibility that HVC(RA) neurons are organized into a chain connectivity. HVC(RA) neurons have excitatory connections with each other. The inhibitory interneurons are excited by HVC(RA) neurons, and also send inhibitory connections to HVC(RA) neurons, providing a local feedback inhibition to HVC(RA) neurons. Such connectivity suggests that the burst sequence in HVC can be generated through the excitatory connections between HVC(RA) neurons and regulated by the feedback inhibition through the interneurons.

The synfire chain model is further supported by a modeling work (Jin et al. (2007)) and related experiments (Long et al. (2010)). By constructing biologically realistic neuron models and exploring network parameters, we demonstrated that stereotypical, short bursts consisting of a few spikes at about 600 Hz, as seen in HVC(RA) neurons in zebra finch, can be robustly generated in a synfire chain network (Fig.3). Our work predicted that the robustness and stability of the dynamics require that HVC(RA) neurons have intrinsic conductance that contributes to bursting, and such conductance might be located at the dendrites (Fig.3). No intrinsic bursting was found in previous experiments that injected step currents to the somata of HVC(RA) neurons (Dutar et al. (1998); Kubota & Taniguchi (1998); Mooney (2000); Wild et al. (2005); Mooney & Prather (2005)). Subsequent intracellular recordings of HVC(RA) neurons in slices and in sleeping birds identified calcium spikes in HVC(RA) neurons, most likely in the dendrites, as the source of the stereotypical bursts in HVC(RA) neurons (Long et al. (2010)), directly supporting our predictions.

The strongest evidence supporting the synfire chain model comes from intracellular recordings of HVC(RA) neurons in singing zebra finches (Long et al. (2010)). The model predicts that the membrane potential of an HVC(RA) neuron remains flat throughout the song motif, except right before and during the burst spiking of the neuron, when it receives strong excitations from HVC(RA) neurons in the preceding group. This is exactly what was observed in the experiments (Long et al. (2010)).

Birdsong syntax

While zebra finches sing songs with a simple syntax of fixed syllable sequences, many other songbird species such as Bengalese finch and canary sing complex songs with variable syllable sequences. Figure 1 shows an example of Bengalese finch song. A syllable can be followed by multiple syllables, and often repeats itself a variable number of times. In addition to such a flexibility, the syntax contains restrictions - not all syllables are allowed to follow a given syllable. The songs of songbird species are diverse (Brenowitz et al. (1997)). Our knowledge of birdsong

syntax is quite limited at present. The best approach is to study the syntax one species at a time.

Bengalese finch is an ideal species to study variable birdsong syntax. Biologically, Bengalese finch and zebra finch are close relatives (Okanoya (2004)). Therefore, the knowledge accumulated over the years on zebra finch can serve as a reliable guide for constructing computational models and designing experiments. Moreover, experimental techniques previously used on zebra finch can be easily transferred to study Bengalese finch. Like zebra finch, Bengalese finch is suitable for laboratory studies (Dietrich (1980); Okanoya (2004)).

Branched chain networks

A simple model of complex birdsong syntax is based on the notion that syllables are encoded by synfire chains in HVC (Jin (2009); Chang & Jin (2009)). The syllable encoding chains (or "syllable chains") are connected in branched patterns such that the end of a chain can be connected to the beginnings of multiple chains. Spike activity propagate along the chains. At the branching points, one of the connected chains is selected to propagate the activity. The selection is due to the winner-take-all competition mediated by feedback inhibition through HVC interneurons. An example of how HVC(RA) neurons are connected to encode the syntactical rule of A to B or C is shown in Fig.4. This model suggests that birdsong syntax is encoded in branching chain networks in HVC (Jin (2009)).

A key feature of the model is the competition at the branching points. It is crucial that no more than one syllable chain is selected for spike propagation, and the selection is probabilistic. Also, the activity should not stop at the branching points. These requirements can be achieved if the mutual inhibition between the chains is strong, and the activity is supported by constant external inputs to HVC(RA) neurons (Jin (2009); Chang & Jin (2009)).

Figure 6 shows a simulation of a branching chain network consisting of four syllable chains corresponding to syllables A, B, C and D (adapted from (Jin (2009))). The mutual inhibition between the syllable chains is implemented by randomly connecting HVC(RA) neurons and HVC(I) neurons (Fig.5a). This is based on the experimental observation that HVC(I) neurons receive convergent inputs from HVC(RA) neurons and send back divergent output; moreover, excitations from HVC(RA) to the interneurons are strong (Mooney & Prather (2005)). HVC(I) neurons spike spontaneously at about 10Hz due to noise injection. The transition rule (Fig.5b) allows branches and cycles, similar to Bengalese finch song (Fig.1). Each of the four syllable chains consists of 1200 HVC(RA) neurons divided into 20 groups. The HVC(RA) neurons are driven by Poisson spike trains, simulating tonic drives from nucleus upstream of HVC. The syllable chains are connected according to the transition diagram in the syntax (Fig.5c). For example, the neurons at the tail of the syllable chain A are connected back to the neurons at the head of the same chain, and also to the neurons at the head of the syllable chain B. Thus, the connections from the end of the syllable chain A bifurcate according to the transition diagram. Connections between other syllable chains are similarly constructed. At the start of the dynamics, the neurons in the first group of chain A are induced to burst. As can be seen in the raster plot (Fig.5d), the spikes propagate in chain A. Once the spikes arrive to the end of the chain, neurons in the first groups of both chain A and chain B are excited. However, due to the mutual inhibition, spikes continue to propagate only in chain B (this selection of chain

A over chain B is determined by noisy fluctuations). This corresponds to the transition from A to B. Following such spiking dynamics, the network successfully generates the syllable sequence "ABBDCDCDCDABBCDCDCDABCDCDABBBB", which obeys the specified transition rule. HVC(I) neurons spike continuously throughout the dynamics (Fig.5e). In different runs of the network dynamics, other sequences allowed by the transition rule are generated. As demonstrated by this example, the syntax of Bengalese finch song can be generated within HVC. The model is robust against changes of the excitation and inhibition strength (Jin (2009); Chang & Jin (2009)).

The external tonic input is important for sustaining the spike propagation in the network, especially at the branching points. The alternative model without the tonic input can only work in a small parameter regime and thus is not robust. This is because the inhibition, which is required for making sure that only one syllable chain is active at any given time, makes the spike propagation prone to extinction at the branching points. Although making the excitatory connections strong can help the survival of the propagation, it tends to promote simultaneous activation of more than one syllable chain, which is undesirable. The tonic input eliminates the possibility of extinction, and thus enhances the robustness of the branching chain network model.

Statistical model of birdsong syntax

The branched chain networks can be mapped to state transition models of birdsong syntax. Each syllable chain is considered as a state. Each state is associated with a syllable, and state transitions generate syllable sequences. The transitions between the states are probabilistic. The simplest state transition model is the Markov model: each state is associated with a unique syllable, and the transition probabilities do not depend on the history of how the current state is arrived at. There are also two special states, the start state and the end state. The state transition dynamics starts from the state state and ends at the end state. Constructing the Markov model from observed syllable sequences is simple. The number of states equals the number of the syllables plus two. The transition probabilities between the states can be computed from the observed frequencies of syllable transitions (Jin & Kozhevnikov (2011)).

To evaluate the state transition model, one can generate syllable sequences following the state transitions, and then compare the statistical properties of the generated sequences to the observed sequences. Syllable repetitions are common in songs of many species such as Bengalese finch. The distributions of the repeat numbers of the syllables is a convenient statistical property to compare. In the Markov model, a repeat distribution is described by $P(n) = p^{n-1}(1-p)$, where p is the probability of self transition of the state, and n is the number of times the syllable is repeated. For all p, this is an exponentially decreasing function of n. Analysis of the songs of a Bengalese finch showed that this prediction does fit the repeat distributions for some syllables but fail for others (Fig. 6a) (Jin & Kozhevnikov (2011)). Some repeat distributions are closer to be Gaussian than exponential. The Markov model is clearly inadequate. This conclusion is also supported by two other measures of the statistical properties of the syllable sequences. An N-gram is defined as a fixed sequence of the syllables of length N. Comparisons of the probabilities distributions of up to 7-grams show that the Markov model fails to capture the probabilities of observing fixed syllable sequences in the songs of the Bengalese finch (Fig.6c)

(Jin & Kozhevnikov (2011)). The probability of observing a given syllable or the end of the sequence at a specific position after the start of the syllable sequence also shows significant deviations from the predictions of the Markov model (Fig.6e) (Jin & Kozhevnikov (2011)).

One possible way of modifying the Markov model is to introduce adaptation. In the branched chain network model, repetitions of syllables is accomplished by reactivating the same chains multiple times. Biophysically, repeated reactivations weaken the synaptic strengths and neural responsiveness (Koch (2004)), which should decrease the probabilities of selecting the same syllable chains with repetitions (Jin (2009)). This suggests that the state transition probabilities may not be fixed. Instead, they could be modified with the state transition dynamics. In particular, the self transition probabilities could decrease as the states are revisited consecutively. A simple implementation of this observation is to decrease the self transition probability to $\alpha^{n-1}p$ if the state is visited for n consecutive times. Here $0 < \alpha < 1$ is an adaptation parameter. With adaptation, the repeat distribution as function of n is given by $P_n = \alpha^{(n-2)(n-1)/2}p^{n-1}(1-\alpha^{n-1}p)$. Depending on the values of p and α, the distribution can have decreasing or close to Gaussian profiles. Adaptation of the transition probabilities allows more varied repeat distributions and can explain some of the Gaussian-like repeat distributions in Bengalese finch songs (Jin & Kozhevnikov (2011)). However, modifying the Markov model with the adaptation of the transition probabilities does not completely dissolve the discrepancy between the model-predicted and observed statistical properties of the syllable sequences (Jin & Kozhevnikov (2011)).

Another way of modifying the Markov model is to introduce many-to-one associations from the states to the syllables (Jin (2009)). Multiple states can be associated with the same syllable. This makes the number of states exceed the number of syllables, and the model becomes the partially observable Markov model (POMM) (Callut & Dupont (2004); Jin (2009)). Because multiple states can produce the same syllables, Markovian state transitions can produce non-Markovian syllable sequences (Jin (2009)). The POMM can be derived from the observed syllable sequences using a state-merging method (Jin & Kozhevnikov (2011)). Although the POMM can be successful in capturing the statistical properties of the syllable sequences, the number of states required can be quite large (Jin & Kozhevnikov (2011)).

Combining the POMM with the adaptation of the transition probabilities (POMMA) leads to an accurate model without adding too many states. The POMMA for the Bengalese finch songs is shown in Fig.7 (Jin & Kozhevnikov (2011)). The model has 14 states for the syllables. Syllable C and D required the most complex representations with three states each. The model accurately captures the statistical properties of the observed syllable sequences (Jin & Kozhevnikov (2011)), as shown in Fig.6b for the repeat number distributions; Fig.6d the N-gram distributions; and Fig.6f for the probabilities of observing syllables at given positions from the start. The number of states for repeating syllables can be further reduced with more complex model for adapting the repeat probabilities (Wittenbach & Jin, to be published).

Conclusion

Birdsongs provide a unique opportunity to understand the neural basis of sequential actions. The relative simplicity of the song system and advances in experimental techniques make it possible to construct biologically constrained computational models that can provide novel predictions for further experiments. This combined experimental and computational approach has shown that

syllables are encoded by synfire chains in HVC. A computational model shows that connecting these chains into branched chain networks produces variable syllable sequences, and suggests that HVC is the site of neural control of birdsong syntax. Partially observable Markov model with adaptation, which is based on the branched chain networks model, can accurately describe the statistics of observed birdsongs in Bengalese finch. Future experiments should shed light on whether this is an accurate picture of the neural basis of birdsong syntax.

Acknowledgments

This research was supported by a Sloan Research Fellowship, the Huck Institute of Life Science at the Pennsylvania State University, and NSF Grant No. IOS-0827731.

References

Abeles, M. (1982). Local Cortical Circuits: An Electrophysiological Study, vol. 18. (Springer-Verlag).

Abeles, M. (1991). Corticonics. (Cambridge, UK: Cambridge University Press).

Amari, S. (1972). Learning patterns and pattern sequences by self-organizing nets of threshold elements. IEEE Trans. Computers.

Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2002). Parallel processing of serial movements in prefrontal cortex. Proc Natl Acad Sci U S A, 99, 13172–7.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? Q J Exp Psychol, 20, 249–64.

Bauer, E. E., Coleman, M. J., Roberts, T. F., Roy, A., Prather, J. F., & Mooney, R. (2008). A synaptic basis for auditory-vocal integration in the songbird. J Neurosci, 28, 1509–22.

Bottjer, S. W., Miesner, E. A., & Arnold, A. P. (1984). Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science, 224, 901–3.

Brainard, M. S. & Doupe, A. J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. Nature, 404, 762–6.

Brenowitz, E. A., Margoliash, D., & Nordeen, K. W. (1997). An introduction to birdsong and the avian song system. J. Neurobiol., 33, 495–500.

Callut, J. & Dupont, P. (2004). A markovian approach to the induction of regular string distributions. In Grammatical Inference: Algorithms and Applications, G. Paliouras & Y. Sakakibara, eds. (Heidelberg: Springer-Verlag Berlin), pp. 77–90.

Cardin, J. A., Raksin, J. N., & Schmidt, M. F. (2005). Sensorimotor nucleus nif is necessary for auditory processing but not vocal motor output in the avian song system. J Neurophysiol, 93, 2157–66.

Chang, W. & Jin, D. Z. (2009). Spike propagation in driven chain networks with dominant global inhibition. Phys Rev E Stat Nonlin Soft Matter Phys, 79, 051917.

Clower, W. T. & Alexander, G. E. (1998). Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. J Neurophysiol, 80, 1562–6.

Dave, A. S. & Margoliash, D. (2000). Song replay during sleep and computational rules for sensorimotor vocal learning. Science, 290, 812–6.

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., van der Linden, M., Cleeremans, A., & Maquet, P. (2003). Cerebral correlates of explicit sequence learning. Brain Res Cogn Brain Res, 16, 391–8.

Diesmann, M., Gewaltig, M. O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. Nature, 402, 529–33.

Dietrich, K. (1980). Model choice in the song development of young male bengalese finches. Z Tierpsychol, 52, 5776.

Doupe, A. J. & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci, 22, 567–631.

Doya, K. & Sejnowski, T. (1999). A computational model of avian song learning. In The new cognitive neurosciences, M. Gazzaniga, ed. (Cambridge, MA: The MIT Press).

Dutar, P., Vu, H. M., & Perkel, D. J. (1998). Multiple cell types distinguished by physiological, pharmacological, and anatomic properties in nucleus hvc of the adult zebra finch. J Neurophysiol, 80, 1828–38.

Estes, W. (1972). An associative basis for coding and organisation in memory. In Coding processes in human memory, A. Melton & E. Martin, eds. (Washington, DC: Winston).

Farries, M. A. (2004). The avian song system in comparative perspective. Ann N Y Acad Sci, 1016, 61–76.

Farries, M. A. & Perkel, D. J. (2002). A telencephalic nucleus essential for song learning contains neurons with physiological characteristics of both striatum and globus pallidus. J Neurosci, 22, 3776–87.

Fee, M. S., Kozhevnikov, A. A., & Hahnloser, R. H. (2004). Neural mechanisms of vocal sequence generation in the songbird. Ann N Y Acad Sci, 1016, 153–70.

Ferraro, F. R., Balota, D. A., & Connor, L. T. (1993). Implicit memory and the formation of new associations in nondemented parkinson's disease individuals and individuals with senile dementia of the alzheimer type: a serial reaction time (srt) investigation. Brain Cogn, 21, 163–80.

Fortune, E. S. & Margoliash, D. (1995). Parallel pathways and convergence onto hvc and adjacent neostriatum of adult zebra finches (taeniopygia guttata). J Comp Neurol, 360, 413–41.

Goller, F. & Cooper, B. G. (2004). Peripheral motor dynamics of song production in the zebra finch. Ann N Y Acad Sci, 1016, 130–52.

Grossberg, S. (1978). Decisions, patterns, and oscillations in nonlinear competitive systems with applications to volterra-lotka systems. J Theor Biol, 73, 101–30.

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

Hahnloser, R. H., Kozhevnikov, A. A., & Fee, M. S. (2006). Sleep-related neural activity in a premotor and a basal-ganglia pathway of the songbird. J Neurophysiol, 96, 794–812.

Hebb, D. (1949). The Organization of Behavior. (New York, NY: Wiley).

Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., & Putz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: a functional mri study. J Neurophysiol, 76, 617–21.

Honda, E. & Okanoya, K. (1999). Acoustical and syntactical comparisons between songs of the white-backed munia (lonchura striata) and its domesticated strain, the bengalese finch (lonchura striata var. domestica). Zoological Science, 16, 319–326.

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In Bird vocalization, R. Hinde, ed. (Cambridge, UK: Cambridge University Press), pp. 61–74.

James, W. (1890). The Principles of Psychology. (Cambridge, MA: Harvard University Press (reprint 1983)).

Janata, P. & Margoliash, D. (1999). Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. J Neurosci, 19, 5108–18.

Jin, D. & Kozhevnikov, A. (2011). A compact statistical model of the song syntax in bengalese finch. PLoS computational biology, 7, e1001108.

Jin, D. Z. (2009). Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. Phys Rev E Stat Nonlin Soft Matter Phys, 80, 051902.

Jin, D. Z., Ramazanoglu, F. M., & Seung, H. S. (2007). Intrinsic bursting enhances the robustness of a neural network model of sequence generation by avian brain area hvc. J Comput Neurosci, 23, 283–99.

Jurafsky, D. & Martin, J. H. (2000). Speech and Language Processing. (New Jersey: Prentice-Hall).

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. Psychol Rev, 110, 316–39.

Kleinfeld, D. (1986). Sequential state generation by model neural networks. Proc Natl Acad Sci U S A, 83, 9469–73.

Koch, C. (2004). Biophysics of computation: Information processing in single neurons. (Oxford University Press New York).

Kozhevnikov, A. & Fee, M. S. (2006). Singing-related activity of identified hvc neurons in the zebra finch. J Neurophysiol.

Kubke, M. F., Yazaki-Sugiyama, Y., Mooney, R., & Wild, J. M. (2005). Physiology of neuronal subtypes in the respiratory-vocal integration nucleus retroambigualis of the male zebra finch. J Neurophysiol.

Kubota, M. & Taniguchi, I. (1998). Electrophysiological characteristics of classes of neuron in the hvc of the zebra finch. J Neurophysiol, 80, 914–23.

Lashley, K. S. (1951). The problem of serial order in behavior. In Cerebral Mechanisms in Behavior (the Hixon Symposium),, L. A. Jeffress, ed. (New York, NY: Wiley), pp. 112–136.

Leonardo, A. & Fee, M. S. (2005). Ensemble coding of vocal control in birdsong. J Neurosci, 25, 652–61.

Li, M. & Greenside, H. (2006). Stable propagation of a burst through a one-dimensional homogeneous excitatory chain model of songbird nucleus hvc. Phys Rev E Stat Nonlin Soft Matter Phys, 74, 011918.

Long, M., Jin, D., & Fee, M. (2010). Support for a synaptic chain model of neuronal sequence generation. Nature, 468, 394–399.

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

Lu, X., Hikosaka, O., & Miyachi, S. (1998). Role of monkey cerebellar nuclei in skill for sequential movement. J Neurophysiol, 79, 2245–54.

Luo, M., Ding, L., & Perkel, D. J. (2001). An avian basal ganglia pathway essential for vocal learning forms a closed topographic loop. J Neurosci, 21, 6836–45.

Marder, E. (2000). Motor pattern generation. Curr Opin Neurobiol, 10, 691–8.

Marler, P. (1970). Birdsong and speech development: could there be parallels? Am Sci, 58, 669–73.

Miyachi, S., Hikosaka, O., Miyashita, K., Karadi, Z., & Rand, M. K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. Exp Brain Res, 115, 1–5.

Mooney, R. (2000). Different subthreshold mechanisms underlie song selectivity in identified hvc neurons of the zebra finch. J Neurosci, 20, 5420–36.

Mooney, R. & Prather, J. F. (2005). The hvc microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. J Neurosci, 25, 1952–64.

Mushiake, H., Inase, M., & Tanji, J. (1991). Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. J Neurophysiol, 66, 705–18.

Nakamura, K., Sakai, K., & Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. J Neurophysiol, 80, 2671–87.

Nissen, M. J. & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology, 19, 1–32.

Nottebohm, F., Kelley, D. B., & Paton, J. A. (1982). Connections of vocal control nuclei in the canary telencephalon. J Comp Neurol, 207, 344–57.

Nottebohm, F., Stokes, T. M., & Leonard, C. M. (1976). Central control of song in the canary, serinus canarius. J Comp Neurol, 165, 457–86.

Ohbayashi, M., Ohki, K., & Miyashita, Y. (2003). Conversion of working memory to motor sequence in the monkey premotor cortex. Science, 301, 233–6.

Okanoya, K. (2004). The bengalese finch: a window on the behavioral neurobiology of birdsong syntax. Ann N Y Acad Sci, 1016, 724–35.

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, 343–56.

Olveczky, B. P., Andalman, A. S., & Fee, M. S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol, 3, e153.

Page, M. P. & Norris, D. (1998). The primacy model: a new model of immediate serial recall. Psychol Rev, 105, 761–81.

Pearson, K. (2000). Motor systems. Curr Opin Neurobiol, 10, 649–54.

Pearson, K. G. (1993). Common principles of motor control in vertebrates and invertebrates. Annu Rev Neurosci, 16, 265–97.

Price, P. (1979). Developmental determinants of structure in zebra finch song. J Comp Physiol Psychol, 93, 260–277.

Rhodes, B. J., Bullock, D., Verwey, W. B., Averbeck, B. B., & Page, M. P. (2004). Learning and production of movement sequences: behavioral, neurophysiological, and modeling perspectives. Hum Mov Sci, 23, 699–746.

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., & van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. Hum Mov Sci, 26, 525–54.

Rosenbaum, D. A., Saltzman, E., & Kingman, A. (1984). Choosing between movement sequences. In Preparatory states and processes, S. Kornblum & J. Requin, eds. (Hillsdale, NJ: Erlbaum), pp. 119–134.

Rumelhart, D. E. & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. the contextual enhancement effect and some tests and extensions of the model. Psychol Rev, 89, 60–94.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., & Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. J Neurosci, 18, 1827–40.

Sakata, J. T. & Brainard, M. S. (2006). Real-time contributions of auditory feedback to avian vocal motor control. J Neurosci, 26, 9619–28.

Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. Psychol Bull, 99, 303–19.

Schmidt, M. F. (2003). Pattern of interhemispheric synchronization in hvc during singing correlates with key transitions in the song pattern. J Neurophysiol, 90, 3931–49.

Shallice, T. (1972). Dual functions of consciousness. Psychol Rev, 79, 383–93.

Sohrabji, F., Nordeen, E. J., & Nordeen, K. W. (1990). Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. Behav Neural Biol, 53, 51–63.

Solis, M. M. & Perkel, D. J. (2005). Rhythmic activity in a forebrain vocal control nucleus in vitro. J Neurosci, 25, 2811–22.

Sompolinsky, H. & Kanter, I. I. (1986). Temporal association in asymmetric neural networks. Physical Review Letters, 57, 2861–2864.

Spiro, J. E., Dalva, M. B., & Mooney, R. (1999). Long-range inhibition within the zebra finch song nucleus ra can coordinate the firing of multiple projection neurons. J Neurophysiol, 81, 3007–20.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In Information processing in motor control and learning, G. Stelmach, ed. (New York: Academic Press), pp. 117–152.

Sturdy, C. B., Wild, J. M., & Mooney, R. (2003). Respiratory and telencephalic modulation of vocal motor neurons in the zebra finch. J Neurosci, 23, 1072–86.

Suthers, R. A. & Zollinger, S. A. (2004). Producing song: the vocal apparatus. Ann N Y Acad Sci, 1016, 109–29.

Tanji, J. (2001). Sequential organization of multiple movements: involvement of cortical motor areas. Annu Rev Neurosci, 24, 631–51.

Thorpe, W. H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch. Fringilla coelebs. Ibis, 100, 535–570.

Vates, G. E., Broome, B. M., Mello, C. V., & Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. J Comp Neurol, 366, 613–42.

Verwey, W. B. (2001). Concatenating familiar movement sequences: the versatile cognitive processor. Acta Psychol (Amst), 106, 69–95.

Verwey, W. B. & Dronkert, Y. (1996). Practicing a structured continuous key-pressing task: Motor chunking or rhythm consolidation? J Mot Behav, 28, 71–79.

Vicari, S., Bellucci, S., & Carlesimo, G. A. (2003). Visual and spatial working memory dissociation: evidence from williams syndrome. Dev Med Child Neurol, 45, 269–73.

Vicario, D. S. & Nottebohm, F. (1988). Organization of the zebra finch song control system: I. representation of syringeal muscles in the hypoglossal nucleus. J Comp Neurol, 271, 346–54.

Vu, E. T., Schmidt, M. F., & Mazurek, M. E. (1998). Interhemispheric coordination of premotor neural activity during singing in adult zebra finches. J Neurosci, 18, 9088–98.

Wild, J. M. (1997). Neural pathways for the control of birdsong production. J Neurobiol, 33, 653–70.

Wild, J. M., Williams, M. N., Howie, G. J., & Mooney, R. (2005). Calcium-binding proteins define interneurons in hvc of the zebra finch (taeniopygia guttata). J Comp Neurol, 483, 76–90.

Williams, H. (2004). Birdsong and singing behavior. Ann N Y Acad Sci, 1016, 1–30.

Williams, H. & Mehta, N. (1999). Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. J Neurobiol, 39, 14–28.

Williams, H. & Vicario, D. S. (1993). Temporal patterning of song production: participation of nucleus uvaeformis of the thalamus. J Neurobiol, 24, 903–12.

Woolley, S. M. & Rubel, E. W. (1997). Bengalese finches lonchura striata domestica depend upon auditory feedback for the maintenance of adult song. J Neurosci, 17, 6380–90.

Yu, A. C. & Margoliash, D. (1996). Temporal hierarchical control of singing in birds. Science, 273, 1871–5.

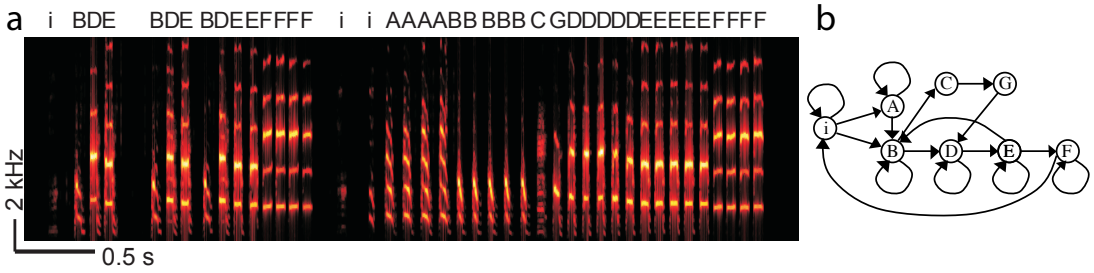

Figure 1: Spectrogram (a) and syllable transition diagram (b) of a Bengalese finch song. The syllables (i, A-G) are labeled on top of the spectrograms.

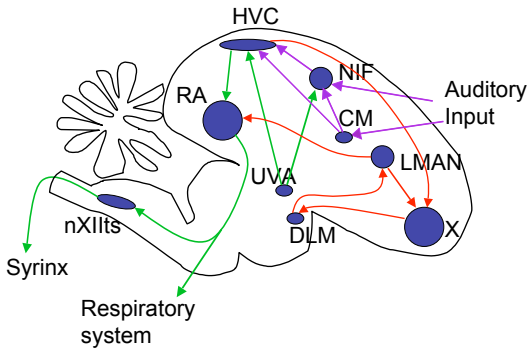

Figure 2: Key brain nuclei of the song control system. Major projections in the motor pathway are indicated with green lines. Red lines indicate the anterior forebrain pathway (AFP). Purple lines indicate the pathway of auditory input. Arrows indicate directions of projections, which are all excitatory.

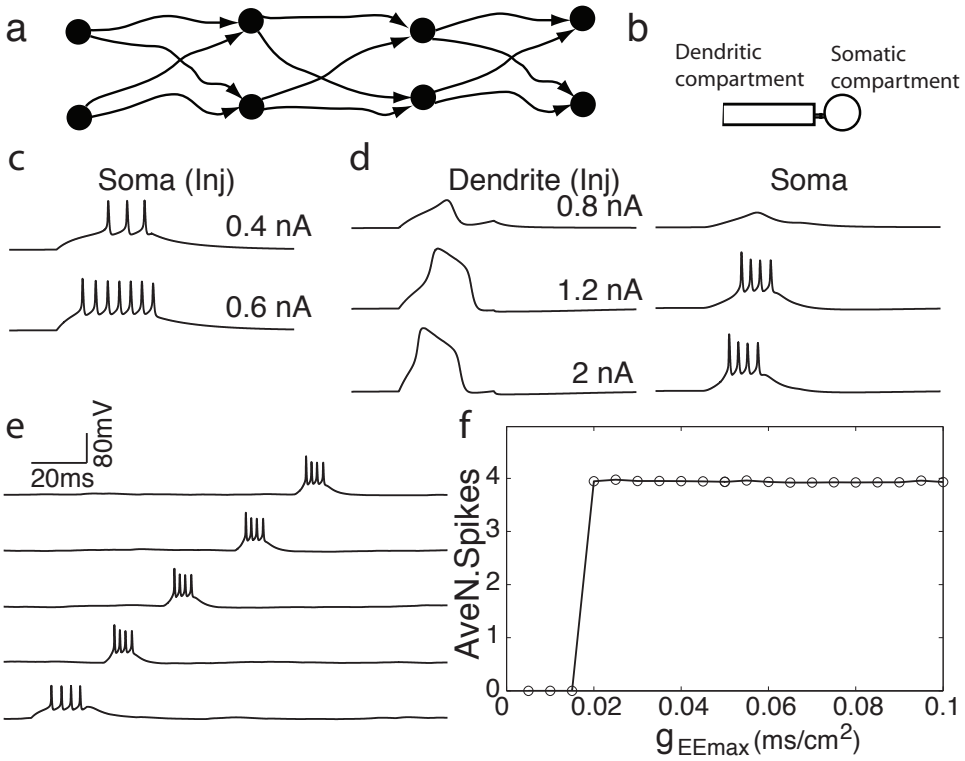

Figure 3: Robust propagation of bursts in the synfire chain network of HVC(RA) neurons with dendritic spike. a. The network connectivity. Neurons form successive groups, and each neuron connects to all neurons in the next group. b. Two-compartment model of HVC(RA) neuron. There are dendritic and somatic compartments. c. Somatic voltage responses with step current injections to the soma. The soma spikes regularly. d. Voltages of the dendrite (left) and the soma (right) with step current injection to the dendrite. Dendritic calcium spike induces a stereotypical burst in the soma. e. Stable propagation of bursts in a chain network of the two-compartment neurons. f. The burst propagation is stable for a wide range of the connection strength, as indicated by the number of spikes per spiked neuron.

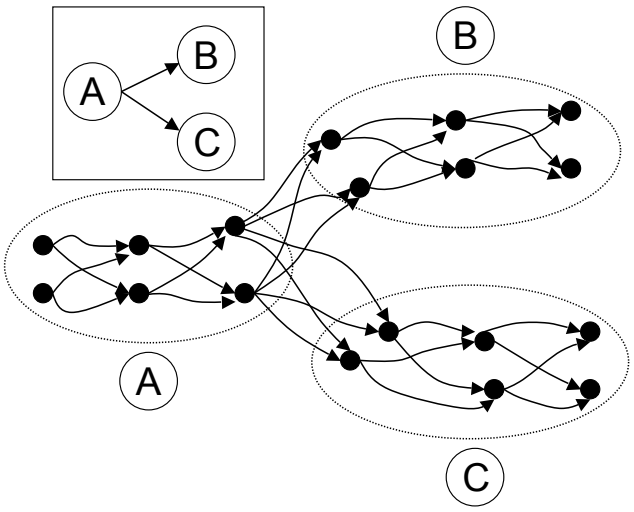

Figure 4: A network of HVC(RA) neurons for generating a probabilistic transition from syllable A to B or C. Each syllable is encoded by a synfire chain network. Chain A branches into B and C. HVC(RA) neurons inhibit each other through the interneurons (not shown). Spike activity propagates from chain A to either chain B or C but not both. The selection of B or C is probabilistic. Adapted from (Jin (2009)).

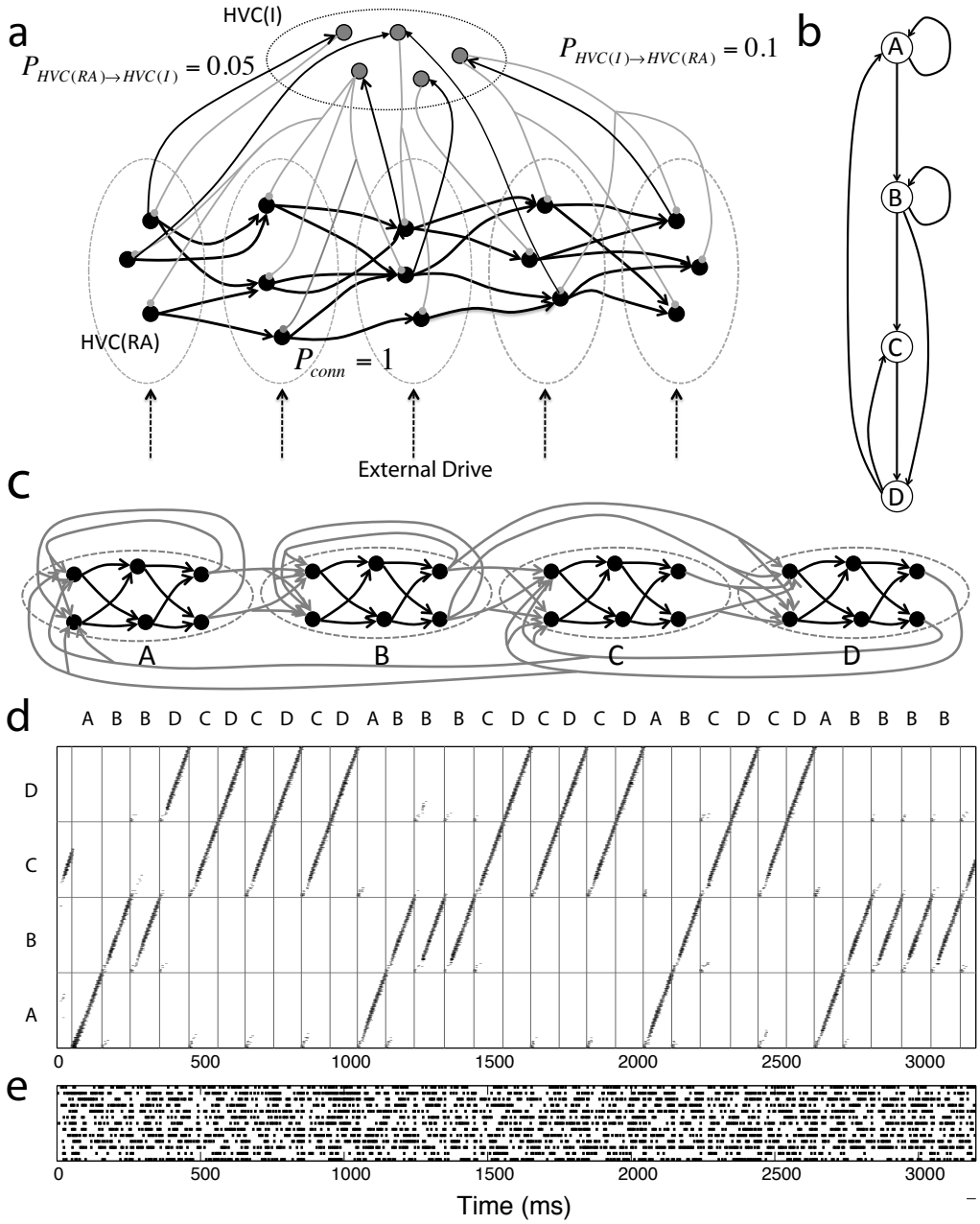

Figure 5: Simulation of branched chain networks. a. Network setup. HVC(RA) neurons form chain networks. They are also driven by external inputs. The HVC(RA) neurons and HVC(I) neurons are randomly connected. b. The simulated birdsong syntax. c. The branched chain network implementing the syntax. d. Spike raster of HVC(RA) neurons. e. Spike raster of HVC(I) neurons. Adapted from (Jin (2009)).

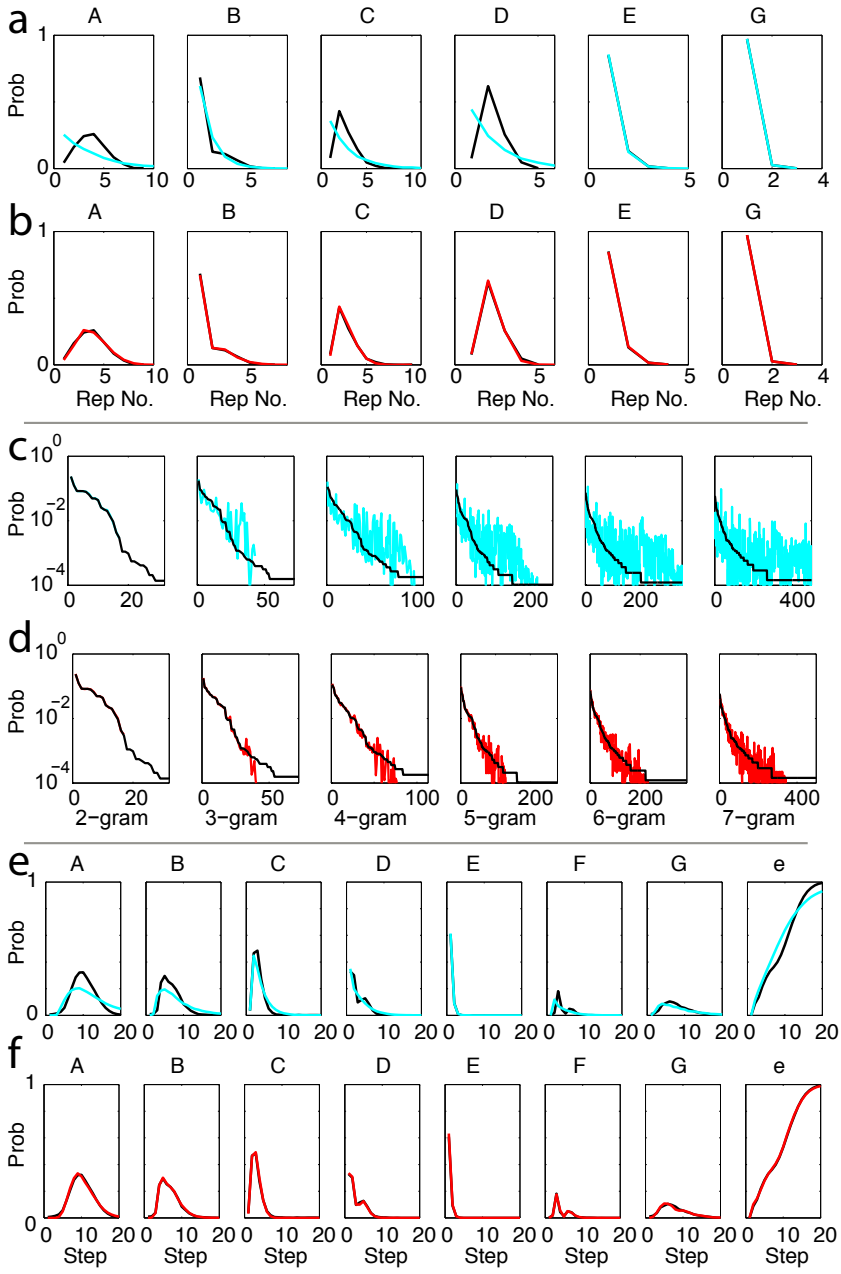

Figure 6: Comparisons of statistical properties of model-generated and observed syllable sequences of a Bengalese finch. Black, cyan, and red lines are probability distributions computed from the observed, the Markov model-generated, and the POMMA-generated sequences, respectively. a-b. Syllable repeat number distributions. The syllable IDs are displayed on top. c-d. N-gram distributions. e-f. Probabilities of finding the syllables and the end (denoted with e) at a given step from the start. The syllable IDs are displayed on top. Adapted from (Jin & Kozhevnikov (2011)).

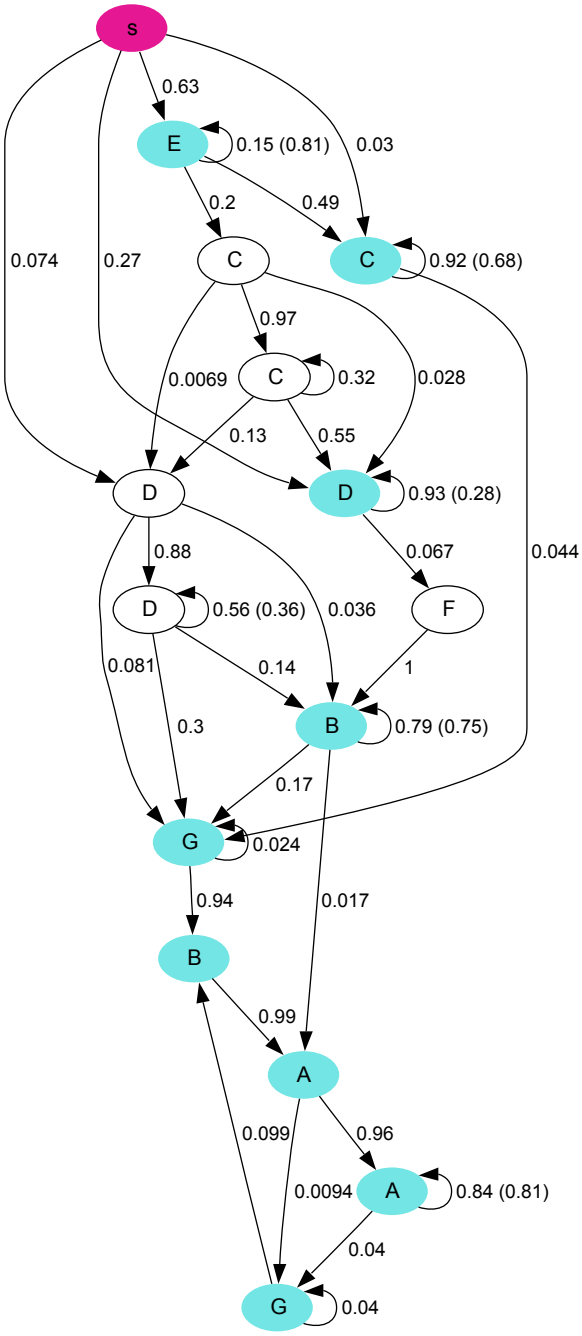

Figure 7: The POMMA derived from the observed syllable sequences of the Bengalese finch. The letters in each oval is the syllable ID associated with the state. Arrows indicate the directions of transitions, and the numbers near the arrows indicate the transition probabilities. The numbers in the parentheses are the adaptation parameters. The pink oval is the start state. Cyan indicates that the state can transition to the end state. Adapted from (Jin & Kozhevnikov (2011)).