# Neural network models of birdsong production, learning, and coding.

Ila R. Fiete[1] and H. Sebastian Seung[2]

[1]Kavli Institute for Theoretical Physics, University of California, Santa Barbara CA 93106
and Center for Neural Systems, Division of Biology, Pasadena CA 91125,
[2]HHMI and Department of Brain and Cognitive Sciences, MIT, Boston MA 02139

October 17, 2007

**Keywords and key phrases:** hebbian learning, associational learning, goal-directed learning, reinforcement learning, gradient learning, spiking neural network, spike-time dependent plasticity, STDP, birdsong, songbird, zebra finch, HVC, RA, LMAN, basal ganglia, empiric synapse, vocalization, sequence propagation, sequence learning, synfire, synaptic chain, sparse coding, sensorimotor, motor control

### Abstract

Birdsong involves motor sequence generation and goal-directed sensorimotor learning, and is controlled by a discrete set of premotor brain nuclei. These features make it an ideal system for theoretical explorations of the neural basis for motor learning and control. We review neural network models of various aspects of song production, in particular the formation of neural sequences to drive song, the learning of the motor map, and the reasons why song may be encoded the way it is by songbird premotor neurons. Our emphasis is on illustrating how theoretical work has contributed to our understanding of the song system and highlighting the resulting predictions for experiment.

## 1   Introduction

Birdsong has been compared to human speech, because both are examples of animal vocalizations that are at least partially learned (C. Darwin, 1971; P. Marler, 1970; A. Douple and P. Kuhl, 1999). Indeed, some of the neural pathways underlying these vocal behaviors may be similar in songbirds and humans (E. Jarvis, 2004). More generally, birdsong can be compared to any motor behavior that involves sequential control and learning, and is an ideal playground for experimental and theoretical explorations of the neural basis for motor learning and control.

The neurons involved in song production and learning are clustered in a small number of discrete nuclei, with well-defined boundaries, fairly well-characterized connectivity, and increasingly well-studied neurophysiology. The various species of oscine songbirds display a vast spectrum of complexity in syllable sequencing (syntax), improvisation, and vocal acquisition. For example, song can be as simple as repetitions of a set of syllables arranged in fixed order (e.g., zebra finches), or can be intricately improvised, with syllable ordering governed by probabilistic transitions between different syllables (e.g., domesticated bengalese finches). Mockingbirds can acquire novel songs and learn novel non-bird sounds throughout life. Yet the basic neural song circuit is similar in all oscines (E. Brenowitz 1997; H. Williams, 2004). This diversity of behavior subserved by an underlying similarity of structure makes birdsong an intriguing system for future comparative studies of sophisticated motor sequencing and control.

Models have so far focused on the zebra finch, a songbird with one of the simplest of oscine song repertoires, because most recordings of neural activity have been obtained from these birds. The male zebra finch learns one song motif with fixed syllable order, and sings repetitions of that motif throughout its life. Song acquisition takes place in a sensory phase, where the juvenile bird acquires a mental template of the song of a possibly unrelated adult tutor, and an overlapping sensorimotor phase, where the juvenile begins to vocalize and tries to match its vocalizations with its mental template (K. Immelmann, 1969; P. Price, 1979). In the laboratory the sensory and sensorimotor phases can be separated: Template acquisition is fast, and can be completed and the tutor bird removed from the juvenile's presence, before the juvenile has begun to practice its song. Sensorimotor learning is a slow process of self-generated trial-and-error, and can proceed successfully in complete auditory and social isolation after template acquisition. However, auditory feedback of the bird's own vocalizations is crucial during sensorimotor learning, because a bird deafened at this stage cannot learn to reproduce its mental template of the tutor song (M. Konishi, 1965; P. Marler, 1970). For a good review, see [1].

In this review, we focus on the motor and sensorimotor aspects of song production, in particular the formation of neural sequences to drive song, the learning of the motor map, and the reasons why song may be encoded the way it is by songbird premotor neurons. Our aim is to illustrate how theoretical work has contributed to our understanding of the song system and highlight the resulting predictions for experiment. Due to space constraints, we omit mention of notable studies on auditory neural coding and dynamics, and on the physics (acoustics and mechanics) of songbird vocal production.

## Basic functional anatomy of the song pathway

In the oscine vocal premotor circuit, high-level nucleus HVC projects to nucleus RA. RA drives the motor neurons that innervate the syringeal and respiratory organs for song production (Figure 1). Lesioning any of these nuclei immediately disables song production, while nuclei upstream of HVC are not necessary for singing in adult zebra finches: Therefore, HVC originates the top-level premotor drive for adult zebra finch song. Although the circuitry between the premotor nuclei is feedforward, connectivity within HVC and within RA is extensively recurrent. In addition to the premotor circuit, the song system contains the anterior forebrain pathway (AFP), which connects HVC to RA through an indirect path. Lesion studies reveal that the AFP is crucial for the sensorimotor phase of song learning and long-term song maintenance, but not for song production (S. Bottjer *et al.*, 1984; C. Scharff and F. Nottebohm, 1991; A. Doupe, 1993) [1]. The AFP is an avian homolog of the mammalian thalamocortical-basal-ganglia pathway.

Based on anatomical, neurophysiological, and lesion data, it is believed that HVC generates the spatiotemporal premotor drive for sequential motor activation in the form of sequential neural activity. The HVC activity is "abstract," in the sense that it encodes only temporal ordering, rather than song features, Figure 2 A–B. The synapses from HVC to RA, and from RA to the motor neurons, convert the sequence of HVC activity patterns into the sequence of motor commands that produces song (Figure 2).

According to this perspective, there are two types of learning. First, HVC acquires the capability of generating a long sequence of neural activity patterns, Figure 2 F–G. Second, the circuitry downstream from HVC acquires the capability of mapping HVC activity patterns to the correct sound-producing patterns of activity in RA. The HVC–RA synapses show extensive structural plasticity, with synaptic growth followed by massive retraction, precisely during the sensorimotor learning period (K. Herrmann and A. Arnold, 1991; H. Sakaguchi and N. Saito, 1996). Electrophysiological synaptic maturation accompanies such structural remodeling (R. Mooney, 1992; L. Stark and D. Perkel, 1999). Therefore, motor map learning from HVC to the vocal outputs is likely to involve plasticity in the HVC–RA synapses. Other possible loci of plasticity for sensorimotor learning,

not ruled out by experiments, might be the recurrent RA–RA synapses or the RA-motorneuron connections.

# Sequence generation

The high-level premotor area HVC displays clear sequential patterns of neural activation. In the awake singing zebra finch, each recorded RA-projecting HVC neuron (also called an $HVC_{RA}$ neuron) fires at most one brief burst of spikes (lasting $\approx$6 ms) per second-long song motif (Figure 2 A) (R. Hahnloser *et al.*, 2002). These neurons are otherwise inactive during singing. Different RA-projecting HVC neurons fire their single bursts at different fixed times in the motif. Thus as a population, the $HVC_{RA}$ neurons cover time during song.

In addition to RA-projecting neurons, HVC contains inhibitory interneurons which ramify within HVC and, X-projecting neurons which send their outputs to the AFP. Inhibitory interneuron activity is very different from $HVC_{RA}$ activity: these neurons fire tonically throughout song, with high rates and relatively little modulation in firing frequency across song, Figure 2 B.

In the adult zebra finch, deafening causes no immediate effect on song production, suggesting that auditory feedback is not important for sequence generation. Also, lesion studies indicate that input from the higher nucleus NIf to HVC is not necessary for singing in zebra finches. What network connectivity, and neural dynamics, underlie the sequence generation dynamics observed in the top-level premotor area, and how might that connectivity become established as hatchlings mature?

## The synaptic chain model

The "synaptic chain" model is an old idea for how networks of neurons might produce sequential patterns of activity. (The idea was already old when it was critiqued by Lashley in 1951 under the name "associative chaining" model.) In its simplest version, the neurons in a network are divided into groups, and the groups are ordered in a sequence. From each group of neurons, there are excitatory synaptic connections to the neurons in the next group of the sequence. The groups and their synaptic connections are like the links of a chain.

If the first group in the sequence is activated, it sends synaptic drive to the second group, which then becomes active. The second group in turn excites the third group to become active, and so on down the chain.[1]

More complex versions of the synaptic chain model also include inhibitory as well as excitatory synapses (S.-I. Amari, 1972), or different types of synaptic currents with multiple time constants (D. Kleinfeld and H. Sompolinsky, 1988). In other versions, precisely synchronous spiking of one group is required to activate the next group. These are known as synfire (or "synchronous firing") chain networks (M. Abeles, 1991).

### The role of intrinsic bursting

Two recent models (M. Li and H. Greenside, 2006; Jin *et al.*, 2007) have implemented the synaptic chain idea using excitatory conductance-based neurons to model sequence generation in the zebra finch HVC.

Recordings in the singing zebra finch show that when an RA-projecting HVC neuron is active during song, it fires a high-frequency burst of 4-6 spikes in a 6-10 ms interval (Figure 2 A). Previously we discussed the question of the mechanism of sequential activation in HVC. Another basic question

---

[1]An alternative mechanism for the propagation of sequential activity in songbird HVC is based on the removal of lateral inhibition between groups of neurons (P. Drew and L. Abbott, 2003); however, this mechanism does not allow for the self-propagation of sequential activity, instead requiring an external periodic drive (clocked input) to induce transitions between activity states.

concerns the mechanisms that give rise to the bursting behavior. One possibility is that the neurons burst because they are driven by synaptic input that lasts for the duration of the burst. Alternatively, the neurons might burst due to the dynamics of their intrinsic conductances.

Li and Greenside [8] have advanced a model that does not use intrinsic bursting. This is consistent with the fact that previous brain slice studies of HVC neurons have not revealed intrinsic bursting (P. Dutar *et al.*, 1998; M. Kubota and I. Taniguchi, 1998; M. Rosen and R. Mooney, 2006; J. Wild *et al.*, 2005). The network model is able to produce sequential neural activity with multi-spike bursts (Figure 3A). However, the realism of this model is compromised by the fact that the peak firing rate within the propagating multi-spike burst achieved by the model ($\approx 15$ ms interspike intervals) is approximately 10 times lower than seen in the actual HVC ($\approx 1.5$ ms interspike intervals).

In contrast, Jin et al. [6] have advanced a model in which HVC neurons are intrinsically bursting. The model generates multi-spike bursts at peak firing rates consistent with those seen in experiments (Figure 3B). For comparison, Jin et al. also simulated a synaptic chain model without intrinsic bursting. Unlike Greenside and Li, they were able to achieve higher peak firing rates more consistent with experimental data, but this required careful fine-tuning of parameters like synaptic strengths. In contrast, the model with intrinsic bursting was quite robust to changes in parameters like synaptic strengths.

Superficially, the Jin et al. model may seem inconsistent with previous experiments in vitro, in which somatic current injection did not trigger intrinsic bursting in RA-projecting HVC neurons However, one can imagine two scenarios in which this inconsistency is resolved. First, somatic current injection might be ineffective for revealing intrinsic bursting that is caused by a dendritic calcium spike. This depends on the strength of coupling between the soma and dendrite, and the voltage threshold for a dendritic spike. Second, the intrinsic properties of RA-projecting HVC neurons appear to be different *in vitro* and in the brain of a singing bird. For example, the peak firing rates observed *in vitro* are never as high as those observed during song. Possibly, HVC neurons *in vitro* lack some neuromodulator secreted during song, and this neuromodulator could be necessary for intrinsic bursting. Given the findings of Jin et al. regarding robustness, it seems important to look again for intrinsic bursting of HVC neurons.

### Directionality of excitatory connectivity

In the synaptic chain models mentioned above, there are excitatory synapses from each group to the next group in the chain. *Unidirectionality* is a fundamental feature. While the synaptic chain model is intuitively plausible, it is important to subject it to experimental testing and contrast the model's predictions with competing models. For example, one can imagine a *bidirectional* synaptic chain in which there are excitatory synapses between adjacent groups in the chain, but the synapses go in both directions, "forward" and "backward" (Figure 3 C, D).

In both the unidirectional and bidirectional models, it is assumed that song is initiated by activating the first group in the chain. This causes the entire chain to be activated sequentially in the forward direction (Figure 3 C and D , red 'electrode'). Both models can reproduce the sequential activity patterns observed in HVC during singing. However, the bidirectional model can also produce other kinds of activity patterns that the unidirectional model cannot (Figure 3 C, D). This could be used as the basis of experimental tests that distinguish between the two models.

In these tests, activity in the HVC network would be initiated by some means, and the subsequent response to stimulation would be observed. If the last group in the chain is stimulated (Figure 3 C and D, blue 'electrode'), there would be no subsequent activity in the unidirectional model. In the bidirectional model, the activity would propagate backward along the chain, and would result in a sequence of motor commands that is reversed from normal song. Another experiment would be to stimulate an intermediate group in the chain (Figure 3 C and D, green 'electrode'). According to the bidirectional model, subsequent activity would simultaneously propagate both forward and

backward along the chain.[2] In the unidirectional model, activity would only propagate forward.

While such experiments may sound straightforward, they are actually technically demanding. So far no topographic organization of RA-projecting HVC neurons has been detected. In other words, the location of a neuron in HVC does not seem to be related to the time at which it is activated during a song motif. If coactive neurons were colocated in HVC, it would be relatively straightforward to activate a hypothetical group in the synaptic chain models by stimulation through a microelectrode. Instead, the neurons of a hypothetical group appear to be scattered throughout HVC. Therefore, a group would be best identified through optical imaging of activity, and best stimulated through optical means also.

So far, no one has attempted to precisely stimulate a group of HVC neurons. However, cruder perturbation experiments have been performed. For example, HVC activity during song has been perturbed by stimulating RA, which leads to retrograde effects on HVC. Stimulation leads to skips in motif sequencing, but the subsequent vocalization is a recognizable fragment of the motif sung in the forward direction, beginning at a different position within the motif. This is more consistent with the unidirectional model than with the bidirectional model. But more conclusive tests will require precisely controlled stimulation of activity in HVC.

A more direct way of testing the synaptic chain models is to check their predictions concerning the synaptic connectivity of HVC. Since the real HVC is no doubt "messier" than in these idealized models, the predictions should be phrased probabilistically. Suppose that A and B are two RA-projecting HVC neurons that are activated in succession. In the unidirectional model, the probability of a connection from A to B is expected to be higher than that of a connection from B to A, or of a connection between two neurons chosen at random.[3] In the bidirectional model, the probability of a connection from A to B is equal to the probability of a connection from B to A.

Testing this kind of prediction could be done by comparing physiological and neuroanatomical measurements. Optical imaging could be used to find two neurons A and B that are activated in succession. Then optical stimulation could be used to see whether and how they are connected to each other. Alternatively, a completely neuroanatomical approach would be to reconstruct the entire connectivity of HVC using serial section electron microscopy, look for a chain within this connectivity, and determine its directionality.

**The role of inhibition**

The synaptic chain models above were based on synaptic excitation, without reference to the role of synaptic inhibition. Inhibition is often included in such models to globally "dampen" activity in a temporally non-specific way, and can enhance the stability of sequential activity in the excitatory neurons. To implement such a global inhibition effect, each inhibitory interneuron receives excitatory input from a large random set of projection neurons, and each projection neuron in return receives synapses from the interneurons. If the burst times of the projection neuron population were distributed across the entire song motif, then each interneuron would receive roughly continuous drive, and spike continuously.

Consistent with this picture, the temporal firing patterns of inhibitory interneurons in HVC are much less specific than those of the $HVC_{RA}$ neurons: an inhibitory interneuron is typically active throughout most of the motif. This suggests that the connectivity of inhibitory neurons is also much less specific than that of excitatory neurons.

---

[2]Note that the playback will not continue from the forward to the backward direction, or from the backward to the forward direction, because of the existence of a refractory period for activation. This point is well-known in the theory of excitable media.

[3]Alternatively, the predicted asymmetry of coupling between A and B might be reflected in synaptic strength, rather than in the presence or absence of connections.

## Sequence learning

Hand in glove with the question of what intrinsic neural properties and network architectures could underlie the stable propagation of activity in HVC, is the challenge of understanding how such architectures could emerge from initially less structured networks.

Intuitively, associative (Hebbian) learning rules like spike timing dependent plasticity (STDP) seem ideally suited to produce the kinds of asymmetric connectivity characteristic of synaptic chains: If neuron A fires before B, the connection from A to B is strengthened, while that from B to A is weakened. If a naive network is trained with this rule for synaptic plasticity on a set of input patterns presented sequentially, one would expect network weights to encode the temporal correlations in the training patterns, and later autonomously reproduce the same long, stable chains of neural activity when stimulated.

But contrary to intuition, past modeling studies show that associative learning rules are not sufficient to configure networks to autonomously generate long neural sequences. Networks trained in this way usually require finely tuned excitatory and inhibitory connectivity that is continuously re-normalized to balance network activity during learning, as well as additional pacemaker inputs to propagate activity from one time-step to the next. The situation is even worse if the network does not have access to training inputs that are themselves sequential (J. Hertz and A. Prugel-Bennett, 1996; N. Levy *et al.*, 2001). In such cases, the best response even with fine-tuning of parameters is the formation of extremely short chains ($< 5$ steps) using thousands of neurons.

The reason for this is that associative learning rules do not balance activity equitably across space (neurons) and time; instead, they tend to amplify inequities. For example, a neuron that by chance receives strong synapses from its inputs will fire often. Because of its frequent firing, it will tend to rapidly form strong outgoing synaptic connections as well. As a result, activity tends to accumulate around a few overly active neural hubs, producing epileptic or synchronous burst-like network states rather than long orderly chains. Empirically, these problems remain unsolved with the imposition of local constraints on the maximum strengths of individual synapses. For these reasons, the question of how plasticity rules that are local in both space and time can impose global constraints on activity balance over all space and all time, is non-trivial.

Two independent models provide a conceptual framework for how associative learning rules must be augmented to allow initially unstructured networks to self-organize and generate stable, long sequences of neural activation [3, 7]. Both models use spike-timing dependent plasticity (STDP) as the primary activity-dependent mechanism for synaptic change. But crucially, in addition to STDP, both models rely strongly on competitively enforced synaptic bounds that are not computed on the level of single synapses, but are triggered by neuron-wide constraints. Although the rules governing synaptic remodeling are not local to the individual synapses, they are local to single neurons. Remarkably, these local neural bounds nevertheless serve to uniformly and globally distribute activity throughout the network in both space and time.

The qualitative nature of these synaptic bounds, and how they are triggered and enforced, is where the two models diverge. The model of Jin and Jun [7] assumes that the formation of a finite number of strong connections from a neuron triggers axon remodeling and elimination of weak synapses. Fiete et al. [3] propose two separate schemes based on neural resource constraints: One assumes that the production rate of synapse-building resources is limited, therefore potentiation in one synapse triggers a uniform heterosynaptic weakening of all other synapses at that neuron. The other scheme assumes a cap on the total resources a neuron can devote to synaptic maintenance; therefore all synapses at a given neuron undergo uniform weakening whenever STDP drives their summed synaptic strength to exceed the cap.

Both models make predictions for the rules governing synaptic plasticity within HVC, and for the statistics of the resulting connectivity within HVC. The Fiete et al. model produces several sequences of different lengths, with a definite distribution of chain lengths. Determining if this distribution matches syllable length distributions in birdsong could indicate whether songbird syllables

attain their specific lengths largely due to bottom-up (low-level network properties) or top-down (behavioral, evolutionary or respiratory) constraints.

## Sensorimotor learning of the motor map

The main goal-directed aspect of learning in zebra finches takes place during the sensorimotor phase, when the juvenile learns to match its own string of vocalizations with its memorized mental copy of the tutor song, through auditory feedback and template comparison. Because the goal-directed aspect of song acquisition is so striking, and because natural goal-directed learning behaviors that are easy to quantify are rare, several models have focused on this aspect of song learning. Yet the microscopic synaptic plasticity rules that underlie such learning remain experimentally undiscovered and theoretically debated. Predictions from existing theoretical or computational models can serve to greatly narrow, in a principled way, the dauntingly large space of possible plasticity paradigms to investigate in experiment, and are now specific enough to be tested in detail in slice preparations of the song system[4].

Existing models of goal-directed sensorimotor song learning share the essential schema of the cartoon model (Figure 2) for the roles assigned to the different premotor nuclei: HVC produces a high-level sequence, and sensorimotor learning involves mapping that abstract sequence into the correct vocal motor sequence through plasticity in the HVC–RA synapses. The primary differences between models lie in the roles assigned to the AFP (Figure 4), in the particular rules for how activity triggers plasticity in HVC–RA synapses, and in their levels of biophysical realism and detail. The plausibility of different assumed roles for the AFP can already begin to be evaluated using existing electrophysiological and lesion results.

Early experiments showed that lesioning the terminal nucleus LMAN of the AFP, which projects back into the premotor pathway, abolishes the ability of juvenile songbirds to learn song (S. Bottjer *et al.*, 1984; C. Scharff and F. Nottebohm, 1991). Moreover, the immediate and lingering effect of the LMAN lesion, besides stoppage of song learning, is that it dramatically reduced the large trial-to-trial variability of juvenile song (C. Scharff and F. Nottebohm, 1991). These results suggested that the role of LMAN in song learning is possibly to drive vocal experimentation through perturbation of the premotor song pathway.

An early, and in several respects, prescient computational model [2] is based in part on these observations. The model applied the specific algorithm of weight perturbation (with momentum) from the reinforcement learning literature to suggest that LMAN directly generates small static weight changes in the HVC–RA connections (Figure 4A). According to the model, these LMAN-driven static weight changes last for the duration of one song rendition, after which they are selectively retained, based on correlation with a song evaluation (reinforcement or critic) signal. The reinforcement (critic) signal is a function of the match between the actual and desired (tutor) songs. Because the weight perturbation rule has been proven by others to perform gradient climbing on the scalar reinforcement signal, it will improve song performance, and do so independently of most parameter values. Their numerical simulation of a small network model of the birdsong network, consisting of rate-based units or neural assemblies with time coarse-grained into syllable-long chunks, demonstrated how this may happen. Experimental studies have supported several features of the model, including the high-level schema of actor-critic reinforcement learning (Figure 5), the role of perturbations in song-learning, and the role of the AFP in driving perturbations.

What are the biophysical assumptions and implications of this model? According to the model, both the song evaluation (reinforcement) signal and the perturbative drive are generated in the AFP, and delivered to the premotor pathway by LMAN. This means that LMAN must carry and convey two independent signals, for perturbation and reinforcement, to the HVC–RA synapses, Figure 4A.

---

[4]Experimental testing of the proposed paradigms is contingent on the identification of a reinforcement or critic signal, which is widely expected to exist but has not yet been found.

The model is based on batch-mode weight perturbation, and it assumes that LMAN can somehow directly perturb the weights of the HVC–RA synapses, in a way that the perturbations vary from song rendition to rendition but are static over the course of a rendition. Moreover, for the weight perturbation scheme to explore enough of the motor space for song learning, HVC–RA synapses that originate from different HVC neurons must receive different, and independent, perturbations from LMAN. There are a few experimental inconsistencies with these requirements in the birdsong system. First, single- and multi-unit activity in LMAN varies rapidly during a song motif, without an unusually large transient at the beginning. It is therefore difficult to imagine how LMAN neurons could provoke an instantaneous perturbation in the HVC–RA weights at the beginning of a song rendition, which is held constant for the rest of the rendition [5]. Second, each RA neuron receives an estimated 1000 inputs from up to 200 HVC neurons, but only approximately 20-50 inputs from LMAN. Thus, it is unlikely that synapses from different HVC neurons could receive independent perturbations to fulfil the requirements for weight perturbation.

Finally, Doya and Sejnowski found that learning with purely random inputs from LMAN was too slow, with poor convergence to the tutor song, even though the simulated network was far smaller than the songbird motor pathway, and even though a song motif consisted of very few time-steps (each time-step represented an entire syllable). Therefore, they assumed that LMAN activity itself evolves by reinforcement learning on the same global reinforcement signal, so that its output is composed of a term proportional to the true HVC–RA weight gradient, superimposed on a random component. With this assumption, the AFP must perform at least three functions: (1) generate and carry random perturbation signals whose dimension might be as large as the number of $HVC_{RA}$ neurons, (2) generate and carry a true gradient estimate after correlating its random activity with reinforcement, (3) function as a critic, generating and carrying the reinforcement signal. With these assumptions, the model indeed learns; however, the AFP is assumed to possess these capabilities de facto, without modeling how the necessary computations are performed, or how the resulting signals are carried by LMAN and delivered to the premotor pathway, leaving crucial questions unresolved. Whether the AFP actually possesses these capabilities is an experimentally open question.

A model proposed by Troyer and Doupe (2000) [9] suggests that plasticity in HVC–RA synapses and RA–RA synapses may be largely *associational* or *Hebbian* in nature. According to the model, plasticity is gated by a reinforcement signal, which LMAN delivers to the premotor network (Figure 4 B). The associational rule additionally includes a subtraction term for synaptic weakening, bringing it closer to reinforcement learning rules, but the specific form used is not proven in the literature to guarantee improved performance. As a result, there are no general assurances of convergence or stability, even in networks of simplified rate-based neurons. The model requires fine-tuning by hand of several homeostatic mechanisms to maintain stability during learning. The authors demonstrate through numerical simulation in rate-based networks that under particular parameter settings, with a small set of RA assemblies, HVC–RA connections and a few syllable-long time-steps, the network can learn to match a desired output.

A strength of this model is that, unlike others in the literature, it stresses the importance of lateral connections within RA, which, when subject to the same plasticity rules as HVC–RA synapses, form assemblies that are capable of pattern completion even with partially correct inputs, to generate "good" output syllables. The AFP's only role in the model is to generate and convey a spatially uniform reinforcement signal to RA (Figure 4 B). However, in the zebra finch, LMAN activity is spatially inhomogeneous (neural activity is not strongly correlated across the nucleus), and connections to RA are not highly divergent, which implies that different RA neurons are likely to receive fairly uncorrelated inputs from LMAN, rather than a uniform signal. As formulated, the model cannot explain why, in experiments, LMAN lesions reduce song variability. Experiments fail to reveal qualitative changes in LMAN activity when an adult bird's auditory feedback is altered

---

[5]It may be possible for weight perturbations to be driven by slow neuromodulatory action of LMAN onto the HVC–RA synapses; within the model, reinforcement would then have to be mediated by either ordinary synaptic transmission from LMAN to RA, or though a second neuromodulator released by LMAN neurons into RA.

(A. Leonardo, 2004), suggesting that LMAN may not primarily convey an error or match signal [6].

Subsequent informal proposals, not grounded in numerical simulation or modeling, have espoused more information-rich primary roles for LMAN in song learning. These include suggestions that LMAN may provide detailed instructive or supervisory error signals to RA. Unfortunately, these suggestions lack mechanistic proposals or models of how the AFP might correctly compute and generate these information-rich time-varying supervisory signals, leaving the bulk of the computational burden of song learning unsolved and at the door of the AFP. In reality it is possible that the AFP supplies such a supervisory signal to RA, but no experimental or computational studies exist yet to suggest that it does.

A recent model of song learning by Fiete and collaborators (2007) [4] returns to a reinforcement learning framework, and attempts to do two things: (1) To propose that, and illustrate how, spiking RA neurons might perform reinforcement learning by using rapidly time-varying LMAN synaptic inputs that perturb RA membrane voltages (rather than using static synaptic perturbations of the plastic HVC–RA weights) [2]) (Figure 4C); and (2) To show as a proof-of-principle that although it is widely thought to be slow, reinforcement learning can be fast enough to explain song acquisition in a full-scale model of the birdsong network with spiking, conductance-based neurons, delayed feedback, scalar reinforcement, and purely random independent perturbations.

The model assumes that LMAN perturbs the premotor pathway by injecting independent random currents into different RA neurons (Figure 4 C), through ordinary glutamatergic synaptic transmission (Figures 5 and 6 D–E). This is consistent with recent stimulation and pharmacological inactivation experiments in LMAN, which show that LMAN helps drive song variability through glutamatergic synaptic projections to RA. The model requires at most as many independent LMAN inputs as there are RA neurons. The LMAN inputs are assumed to be time-varying throughout song. Reinforcement is assumed to be computed and delivered from elsewhere in the brain, with learning following the high-level actor-critic schema. Under these conditions, the derivation of the rules for how activity should modify synaptic weights, and the rules themselves, are quite different from weight-perturbation like algorithms (Figure 6 D–F): Each HVC–RA synapse must somehow deduce an instruction for change based on time-varying fluctuations in the conductance of its postsynaptic RA neuron, rather than simply deciding whether to retain or discard direct static perturbations of its own weight through correlation with the reinforcement signal. The resulting rule is guaranteed to robustly (without parameter tuning) improve song performance, even in recurrent networks of conductance-based spiking neurons and with delayed reinforcement.

The model is able to produce learning in a biologically realistic network model of the song system, with spiking neurons and an impoverished, delayed reinforcement signal, Figure 6 A, B. Learning takes place on-line, and activity in the network evolves on the millisecond time-scale, driven by 5-6 ms bursts in the HVC inputs. Within 2000 iterations, the network output resembles the tutor song, which is a recording of an actual zebra finch song, Figure 6 C. Fiete et al. show that learning time even with random LMAN input should scale well in the full-sized premotor network, and may converge in far fewer iterations than sung by the typical zebra finch ($\approx$ 100000 renditions) during song acquisition.

The model's central prediction is a specific paradigm for the induction of bidirectional, heterosynaptic activity-dependent synaptic plasticity in the premotor vocal pathway (Figure 6 F) that should be possible to test in slice experiments.

The model above is consistent with a large body of experimental data in the song-system, and provides concrete predictions for synaptic plasticity. However, LMAN's role in the model is to provide purely random exploration. This assumption does not contradict the data, but is a limited view of what the AFP, a complex neural pathway, might be capable of. More sophisticated strategies for AFP-delivered perturbation may speed up learning. This avenue is wide open for discovery by

---

[6]If LMAN does generate an error signal, but bases it on learned predictions about auditory feedback, it is possible that prolonged rather than transient exposure to distorted auditory feedback is necessary to produce changes in LMAN activity.

theory and experiment.

More generally, models of song learning based on the AFP as a source of random or directed perturbations to the premotor system, could be used as comparative tools for studying the role of the basal ganglia thalamocortical loop in mammalian sensorimotor learning.

## Premotor codes and representations

Up to 30000 neurons are devoted to song production and learning in the zebra finch. Together, these neurons drive only 8-10 muscles that produce song. What are some of the functional requirements and costs that determine why so many neurons are required for song production?

The ultra-sparse or unary coding of song by $HVC_{RA}$ neurons (Figures 2 and 2 A) exacts a large energetic and space cost, in terms of the number of neurons required to encode a given set of patterns. To produce a sequence of length $L = 100$ steps with $m = 200$ neurons on per time, unary coding requires $N = Lm = 20000$ neurons. At the other extreme, a combinatorial code with a downstream readout capable of distinguishing between different combinatorial patterns, would require far fewer neurons, scaling as $\log(L)$ instead of $L$; intermediate possibilities would require intermediate numbers of neurons. What are some of the properties of a unary code that might offset its capacity costs and explain why it exists in HVC?

A possible explanation for ultrasparse HVC sequences could be mechanistic: recurrent networks cannot store many dense patterns due to increasing interference between patterns. Studies show that sparsely active attractor networks can store more patterns (M. Tsodyks and M. Feigelman, 1988). Nevertheless, in these studies, 'sparse' is far denser than unary, so the reason for ultrasparse coding must lie elsewhere.

Because oscine birdsong is learned, a natural place to investigate the role of unary HVC codes is in learning. There are actually two distinct roles that unary codes could play in song learning.

First, an analytical and numerical study by Fiete et al. (2004) shows that unary coding within HVC enables optimally fast learning of the feedforward motor map from abstract sequences in HVC to sequences of muscle activation. [5], If each HVC neuron were to fire twice instead of once, at random times in a song motif of fixed length, the fastest possible speed with which the input-output map can be learned is twice as slow (Figure 7 A–D). There is an intuitive explanation of the underlying mathematical reason for this effect. Suppose an HVC neuron, and its corresponding HVC–RA synapses, are active at twice, when the tutor song has opposing characteristics (e.g. pitch or amplitude) at those two times. The synapses will be pulled in opposing directions to best match the tutor song (Figure 7 C) at each time. Because the learning task is realizable, there is some value of synaptic strength that achieves the desired output, and in the end this value will be reached. But the conflicting demands on synapse strength at the two times in the motif make approach to the solution with a hill-climbing rule slower, even if the parameter controlling learning step size is optimized to produce the fastest possible learning without network instability.

Yet answers to "why" questions, however compelling, risk being just-so stories, unless they generate specific testable predictions. The work described above argued that unary codes may exist in HVC because of their utility for fast song learning. However, single-unit $HVC_{RA}$ activity and unary codes have only been observed in *adults*. Thus, the central prediction – which if false would invalidate the hypothesis – is that activity in $HVC_{RA}$ neurons, even if non-sparse initially (early in the sensory period), should be unary near the onset of sensorimotor learning in juvenile zebra finches, instead of emerging gradually after or towards the end of sensorimotor learning.

The analysis leads to an additional prediction. To accommodate a larger song repertoire, HVC could change in two *a priori* equally likely ways: (1) Keep HVC volume (number of neurons) fixed, but increase the number of bursts per neuron, or (2) Increase HVC volume, but maintain a unary code. Energetics might favor the first possibility. But according to the learning speed analysis, if zebra finches are under pressure to learn their songs quickly, then unary coding should be conserved,

even at the cost of a larger HVC. This suggests that coding in HVC across zebra finchs and related songbirds with somewhat larger repertoires (measured by number of unique syllables), should remain unary while HVC size scales with repertoire size. Indeed, existing experimental findings corroborate these predictions: HVC size scales with repertoire size after correcting for overall brain size, while RA, LMAN, or other song nuclei do not display similarly consistent relationships (T. DeVoogd *et al.*, 1993; D. Airey and T. DeVoogd, 2000; L. Garamszegi and M. Eens, 2004; J. Pfaff em et al., 2007).

Second, analysis and simulation by Fiete *et al.* (2007) [4] shows that if the HVC code is unary and if during a song rendition performance is continuously evaluated and a reinforcement signal continuously delivered (albeit delayed), then the time taken to learn song is independent of the length of the learned song. Because the HVC code is unary in zebra finches, this prediction can be used in experiments (where learning time is monitored for birds trained on tutor songs of different lengths) as an indirect probe of whether the elusive reinforcement signal in the song pathway is delivered online or in batch mode. If the reinforcement signal is evaluated and delivered only once, at the end of song, then the time taken to learn a tutor song should scale linearly with its length. Similarly, if the HVC code at each time consisted of a random pattern of activations, with half of all neurons active per time (dense code), learning time should scale with song length (even if reinforcement is delivered online).

**HVC encoding in other songbirds:** Do other songbirds also display unary HVC representations of song? At present, single-unit recordings of RA-projecting HVC neurons are only available from zebra finches. Learning with a pure time-code is non-generalizable: like one piece of magnetic tape, each pattern is used to produce song at only one specific time. An entirely different set of HVC neurons and HVC–RA weights must be used and trained to produce the same sound if it recurs elsewhere in the motif. Therefore, songbirds that rapidly acquire and imitate new songs (albeit after slowly acquiring a first song), may use a different encoding strategy (e.g. tuning codes based on acousitic features) than a time-code, Figure 7 E–G. In addition, songbirds with vastly larger repertoires, regardless of the speed of song acquisition, may display different HVC codes, because of capacity constraints.

## Conclusions

Advances in the acquisition of behavioral and anatomical data and neurophysiological recordings from awake, singing birds, are rapidly turning the birdsong circuit into an ideal system for unraveling the mechanisms underlying motor control and learning. Theoretical models have helped to lay a groundwork for understanding the dynamical and coding principles that lead to such functionality, by quantitatively validating or eliminating candidate explanations of network function. Theoretical models also provide a functionally motivated set of predictions about neural activity, connectivity, coding, and plasticity that narrow the choices of future experiments, and are presently ripe for experimental testing. Future experiments and modeling in songbird species with more flexible learning behaviors should greatly enhance our understanding of generalizable motor learning and control.

## Further Reading

[1] M.S. Brainard and A.J. Doupe. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature*, 404(6779):762–6, Apr 13 2000.

[2] K. Doya and T.J. Sejnowski. A novel reinforcement model of birdsong vocalization learning. In G. Tesauro, D.S. Touretzky, and T.K. Leen, editors, *Advances in Neural Info Proc Sys 7*, pages 101–108. MIT Press, 1995.

[3] I.R. Fiete, L. Burger, W. Senn, and R.H.R. Hahnloser. Enforcing global coding constraints by synaptic competition - a model for the formation of long ultrasparse sequences in the songbird. *Submitted*, 2007 (See also Society for Neuroscience Abstracts, 2005.).

[4] I.R. Fiete, M.S. Fee, and H.S. Seung. Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J. Neurophysiol.*, 98:2038–2057, 2007.

[5] IR Fiete, RH Hahnloser, MS Fee, and HS Seung. Temporal sparseness of the premotor drive is important for rapid learning in a neural network model of birdsong. *J Neurophysiol.*, 92(4):2274–82., 2004.

[6] D.Z. Jin, F. Ramazanoglu, and H.S. Seung. Intrinsic bursting enhances the robustness of a neural network model of sequence generation by avian brain area hvc. *J. Comput Neurosci, In press*, 2007.

[7] J.K. Jun and D.Z. Jin. Development of neural circuitry for precise temporal sequences through spontaneous activity, axon remodeling, and synaptic plasticity. *PLoS One*, 2(8): e723, 2007 (See also Society for Neuroscience Abstracts, 2005.).

[8] M. Li and H. Greenside. Stable propagation of a burst through a one-dimensional homogeneous excitatory chain model of songbird nucleus hvc. *Phys Rev E*, 74(1.1):011918, 2006.

[9] T.W. Troyer and A.J. Doupe. An associational model of birdsong sensorimotor learning i. efference copy and the learning of song syllables. *J Neurophysiol*, 84(3):1204–23, Sep 2000.
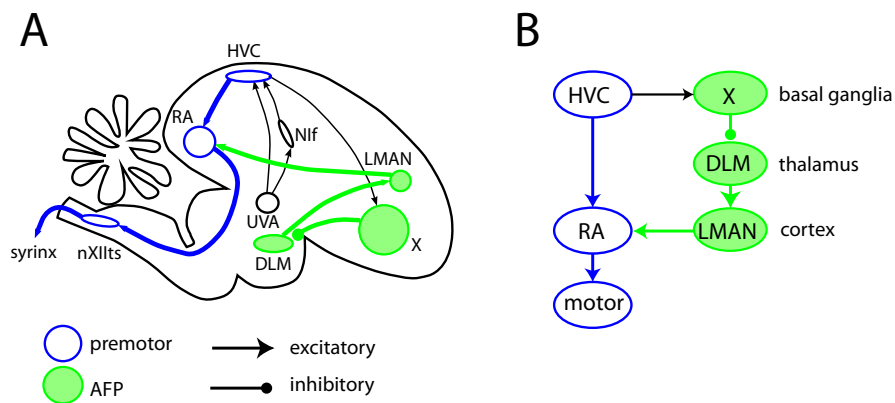
Figure 1: **The neural pathways underlying song learning and song production.** A, Song learning and production involves a discrete set of nuclei. B, the same areas, unfolded to better illustrate their connectivity. Lines terminating in arrows designate excitatory or putative excitatory connectivity. Lines terminating in circles designate inhibitory synapses. Blue nuclei are part of the premotor pathway. Green nuclei are part of the anterior forebrain pathway (AFP), and are important for sensorimotor song learning but not song production.
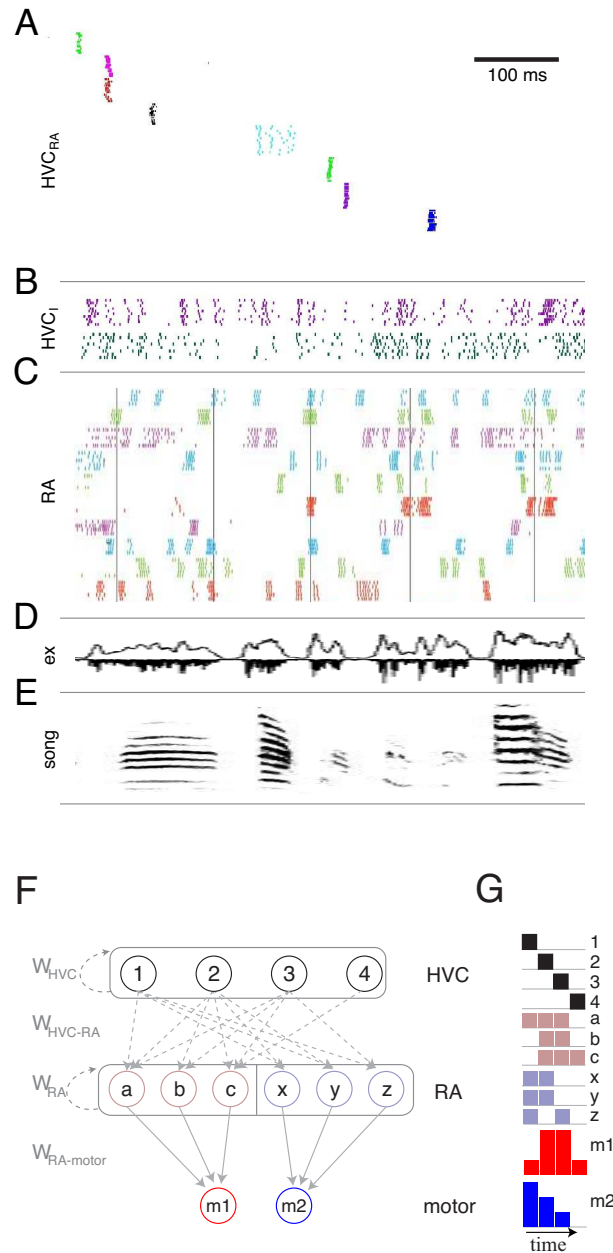
Figure 2: **Summary of data and schematic model. A–E: Activity in the premotor pathway.** Raster plot of single-unit activity in $HVC_{RA}$ neurons (A), in HVC interneurons (B), and in RA (C). Consecutive lines of the same color are multiple recordings of the same neuron. Ticks represent single spikes. D, EMG activity in the abdominal expiratory muscle. E, Spectrogram of resulting song. [A, B Adapted from Hahnloser *et al.*, Nature, 2002; C from Leonardo and Fee, J. Neurosci., 2005; D,E from Suthers *et al.*, Phil. Trans. Royal. Soc. B, 1999] **F,G: Schematic of loci of song production and plasticity** F, Flow of activity: HVC drives RA, which drives muscle contraction via motorneurons, in a feedforward path. In addition, connectivity within HVC and RA is recurrent. Dashed lines represent weights that are likely to be plastic; solid lines with arrows designate weights assumed to be fixed. HVC, with its recurrent connectivity, produces the high-level sequential drive. Motorneurons, designated by motorneuron pools m1 and m2, sum the activity from RA (G, middle 8 traces) to generate drive for the muscles (G, bottom two traces). Sequence learning involves plasticity in the recurrent HVC connectivity. In this simple picture, sensorimotor learning is dissociable from sequence learning, and involves plasticity in the HVC–RA and perhaps RA–RA connections.
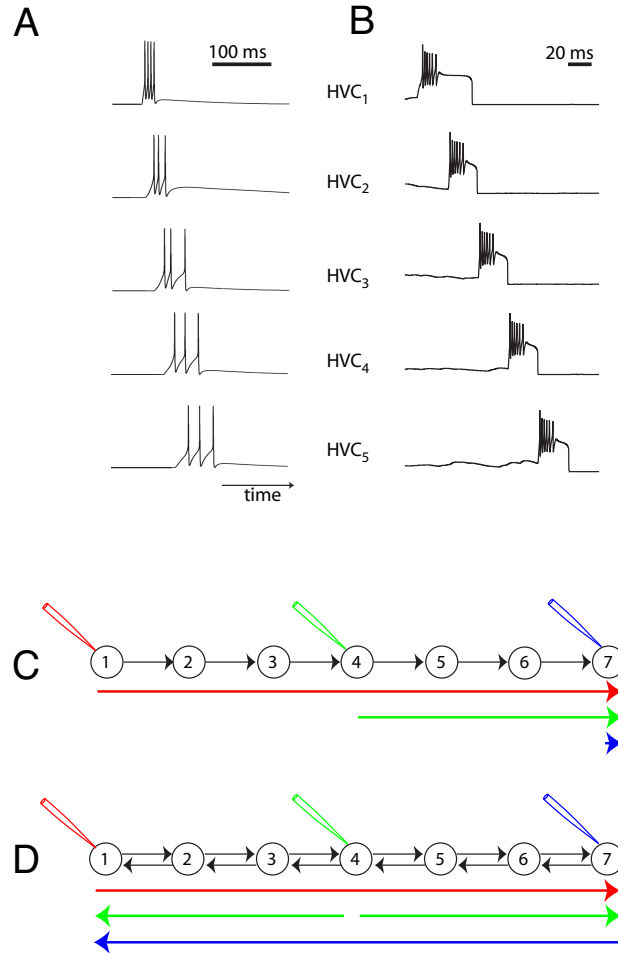
Figure 3: **A,B: Propagation of bursts in synaptic chain models of HVC.** A, Neurons in the model are not intrinsically bursting. Activity propagates due to synaptic chain connectivity, and the burst of spikes in each neuron is due to sustained synaptic drive from the network. Spike frequency during the burst is $\approx$ 10 times lower than in $HVC_{RA}$ neurons [adapted from Li and Greenside, 2006]. B, $HVC_{RA}$ neurons are assumed to possess an intrinsic mechanism for burst-generation. A synaptic chain network can robustly produce stable sequences of multi-spike neural activation [adapted from Jin *et al.*, 2007]. **C,D: Unidirectional synaptic chain versus bidirectional connectivity** C, In a synaptic chain, connections between different groups of neurons project in a predominantly 'forward' direction. Activity initialized at group 1 (red 'electrode') will drive sequential activity up to the end of the chain, at group 7. Stimulation of the last group of the chain (blue 'electrode') will produce no propagating activity or sequence playback. Activation of a group near the middle of the chain will generate playback of a fragment of the original activity sequence. The resulting vocalization should correspond to normal song, but beginning mid-song or mid-syllable. D, Bidirectional synaptic connectivity. Activity initialized at group 1 (red 'electrode') will produce a unidirectional sequence of activity propagating from beginning to end, as in the synaptic chain. Stimulation at the end of the chain (blue 'electrode') will lead to sequence playback in the reverse direction, from group 7 to 1. The resulting vocalization should be the song or a syllable sung in reverse. Stimulation of a central group in the chain will lead to activity propagation in both forward and backward directions. The resulting vocalization should be a superposition of the song (or a syllable) sung forward and in reverse, starting in the middle.
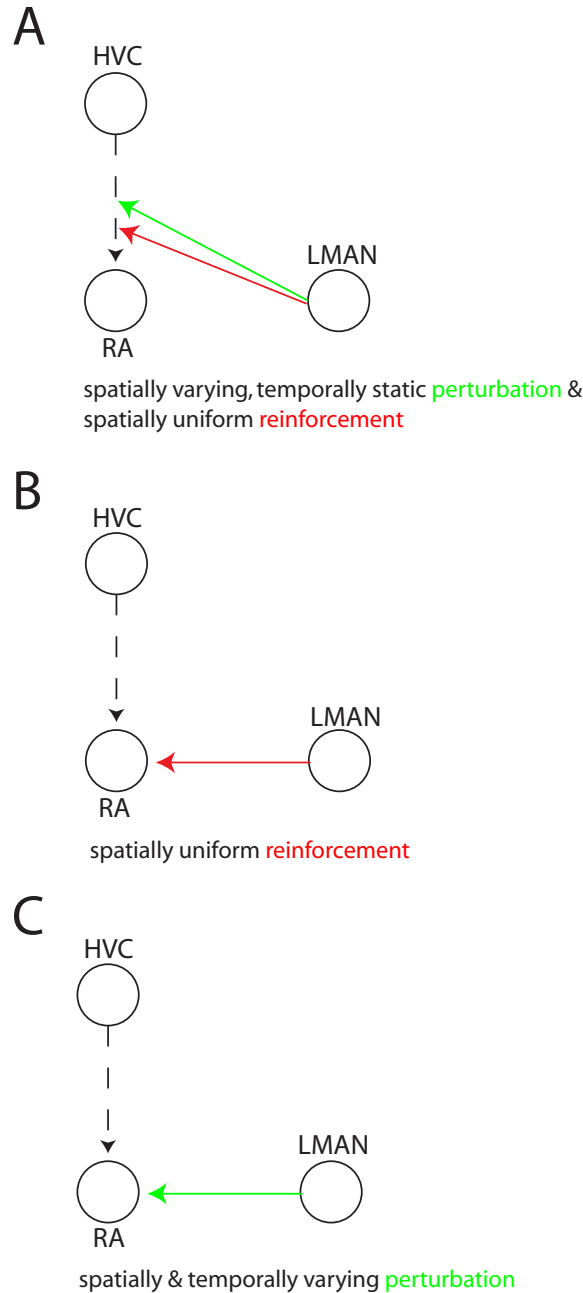
15

Figure 4: **Proposed roles for LMAN in sensorimotor learning.** The dashed line indicates a plastic synapse between HVC and RA neurons. A, The hypothesis of Doya and Sejnowski: LMAN must produce different, independent perturbations in each HVC–RA weight originating from different HVC neurons. These perturbations must be static over the duration of one song rendition, but vary across renditions. LMAN must also broadcast a separate, spatially uniform reinforcement signal to the premotor pathway. B, The hypothesis of Troyer and Doupe. LMAN's sole role is to provide a spatially uniform reinforcement signal to the premotor pathway. Learning is associative (Hebbian). C, The hypothesis of Fiete *et al.*: LMAN must inject different perturbations to different RA neurons. The perturbations are excitatory synaptic inputs to RA, and vary in time over one song rendition. (A spatially uniform reinforcement signal is assumed to arrive at RA from elsewhere in the network.)
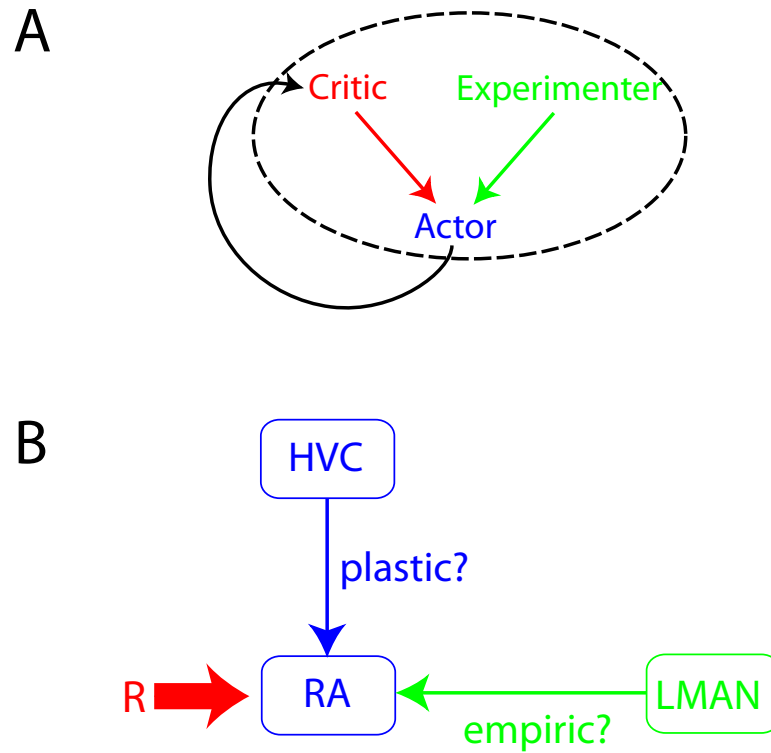
16

Figure 5: **The empiric synapse hypothesis.** A, The empiric synapse hypothesis for goal-directed learning consists of 3 parts: The actor network, which is responsible for performing the desired task and has plastic synapses, and two its inputs, one an experimenter, whose role is to drive perturbations in the actor through empiric synaptic inputs, and the second a critic, which provides a scalar assessment of the performance of the actor on the desired task, through a reinforcement signal. Actor neurons are hypothesized to correlate experimenter-driven perturbations with the reinforcement signal to derive a signal for synaptic change, in the direction of the gradient of the reinforcement signal. B, In the song system, the actor network is hypothesized to be the premotor pathway including HVC and RA, and plastic HVC–RA synapses. The experimenter is hypothesized to be the AFP. A critic-generated reinforcement signal (R) is widely assumed to exist in the song system, although its neural identity is unknown. Adapted from [4].
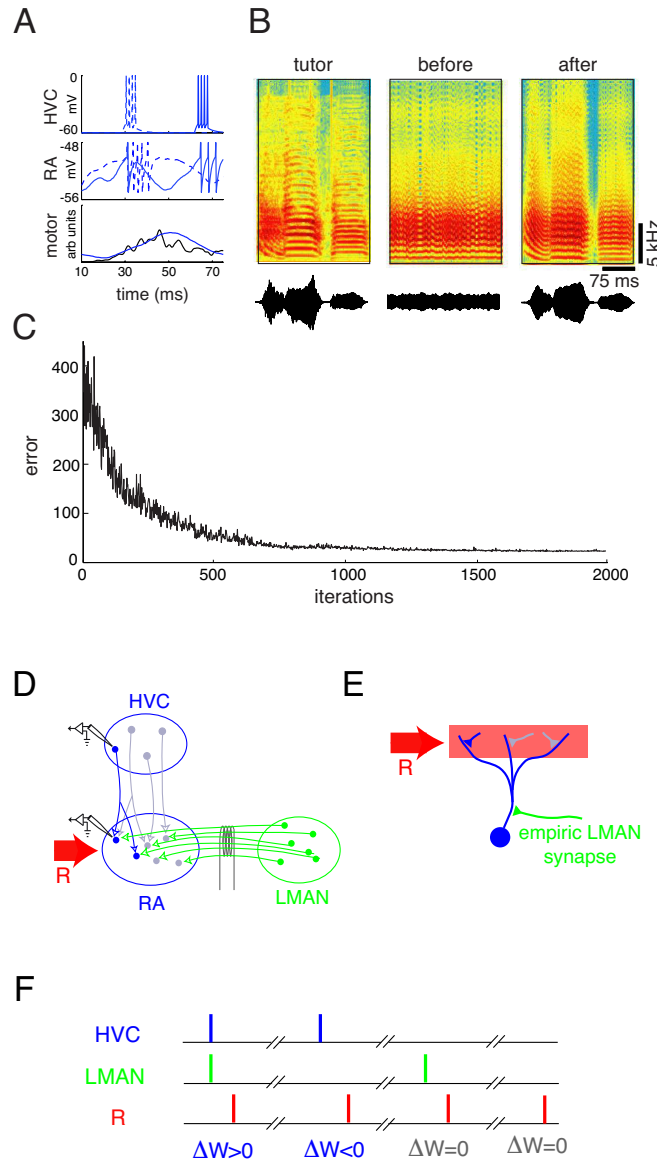
Figure 6: **Empiric synapses: results and predictions.** A, Activity of two model HVC neurons (top), two representative RA neurons (middle), and one of the motor output pools (bottom– black: tutor, blue: network response) during song learning. B, Spectrograms of the tutor song, and the output of the model song network before and after 1200 iterations of learning, together with the respective sound pressure waves. C, Mismatch between the network output and the tutor song decreases relatively rapidly. Learning has reached an approximate asymptote by 1000 iterations. **D–FPrediction for bidirectional synaptic plasticity from the empiric synapse model.** D, Possible experimental design for testing the empiric synapse rule in the birdsong pathway involves focused stimulation in HVC, intracellular recording in RA, and fiber bundle stimulation of the LMAN-RA axons, together with control of the reinforcement signal (R). E, Magnified schematic of a single RA neuron and its inputs: Each RA neuron receives a few empiric synapses from LMAN, and several from different HVC neurons; in addition, RA receives a reinforcement signal. Blue (gray): active (quiescent) premotor inputs; green: empiric input; red: reinforcement signal. F, Protocol for induction of bidirectional synaptic plasticity in active HVC–RA synapses: Simultaneous stimulation of an HVC neuron and the LMAN axon bundle, followed by administration of reinforcement, should lead to potentiation of all stimulated HVC synapses. Stimulation of an HVC neuron with no stimulation of the LMAN axon bundle, followed by administration of reinforcement, should lead to depression of the all stimulated HVC synapses. [Adapted from Fiete *et al.,* 2007 [4].]
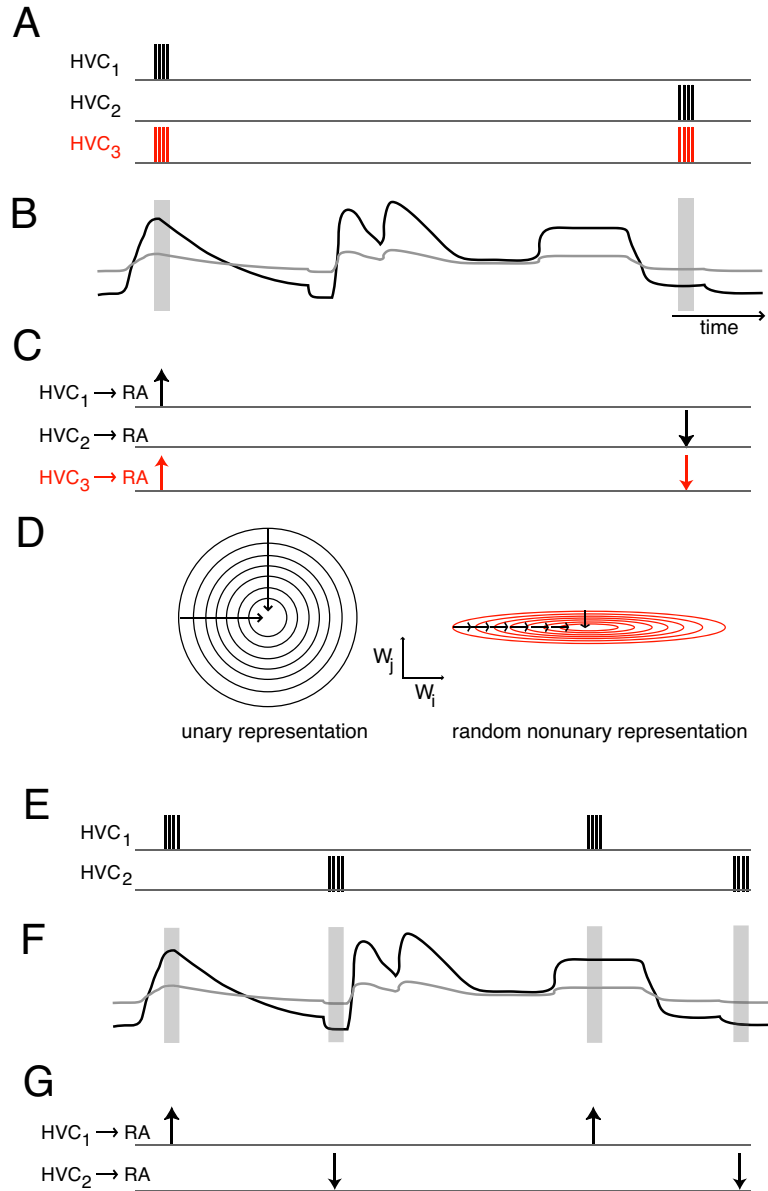
Figure 7: **Unary coding in HVC helps to speed the acquisition of the motor map.** A, The activity of three hypothetical HVC neurons. The first two are each active only once in the song motif, but the third is active twice. B, Hypothetical pitch of the tutor song (black), and the pupil network (gray). C, Direction in which synapses should change to reduce error between tutor and pupil pitch. Synapses from HVC neuron 3 should be strengthened (weakened) to improve the tutor-pupil match at the first (second) activity burst. Such conflicting demands on the synapses of neurons that are active at two or more random times in a motif cause a slow-down in learning speed. D, Contours of iso-error in the learning surface. Learning a feedforward map in a network with unary coding in the top layer is like learning on an isotropic cost surface, and can be fast. Denser coding in the input layer produces correlations and makes the learning surface anisotropic. To keep the error from diverging along the steep directions, the learning rate must be kept low. As a result, best-case learning is slower than in the isotropic case.**HVC activity tuned to acoustic features may be helpful for generalizable learning.** E, If HVC neurons fired multiple bursts, at selected points when the acoustic features of tutor song are similar, F, rather than at random times, there would be no interference in the learning update, F. If HVC neurons acquired such a tuning to features in the tutor song, it could be easy for the bird to quickly reproduce a specific heard sound, by activating the requisite sound-tuned HVC neuron.