The Pennsylvania State University

The Graduate School

Department of Physics

NEURAL MECHANISMS OF VOCAL SEQUENCE GENERATION IN SONGBIRDS

A Dissertation in

Physics

by

Yisi Zhang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2014

The dissertation of Yisi Zhang was reviewed and approved* by the following:

Alexay A. Kozhevnikov
Assistant Professor of Physics
Dissertation Advisor
Chair of Committee

Dezhe Jin
Associate Professor of Physics

Jun Zhu
Associate Professor of Physics

Christine Keating
Professor of Chemistry

Richard Robinett
Professor of Physics
Associate Department Head
Director of Undergraduate and Graduate Studies

*Signatures are on file in the Graduate School

ABSTRACT

Complex learned sequential behaviors, such as speaking, playing musical instruments and singing, occur in daily life. To understand the neural mechanisms governing the generation and learning of these complex motor sequences, the vocal system of songbirds has been used as an excellent model. Many studies have been focused on how the stereotyped song sequence is generated by the neural circuit of zebra finches. A more complex aspect of the birdsong, the syntactic organization of the vocal sequences that is common in both birdsong and human language, has not been well studied. The Bengalese finches sing variable songs that follow certain syntactic rules, and therefore can be used as a tool to study the syntax control of birdsong. This thesis progressively investigates where the song syntax is encoded in the Bengalese finch vocal system using various perturbation methods.

The first section of the thesis demonstrates that HVC, a premotor song control nucleus, plays an active role in encoding the syntax. Changing temperature in HVC strongly affects the length of the syllable repeats. In addition, the transition probabilities at several branching points are altered by temperature. In contrast, changing temperature in the downstream area RA does not affect song syntax. These findings suggest that the neural activity in HVC generates the syntax and is not instructed by inputs from higher order areas.

Auditory feedback plays an important role in shaping the syntax of Bengalese finch songs. The second section explored how the real-time distorted auditory feedback (DAF) affects the syntax. The syllable repeats are reduced by DAF, which is similar to the effect of cooling HVC. This suggests that the auditory feedback influence on syntax may be mediated via HVC.

Songbird singing is controlled by bilateral brain hemispheres. The third part of the thesis investigates the function of the two HVCs in timing and syntax control by unilaterally changing temperature in HVC. The left HVC is found to control the timing of the syllables, while both

HVCs control the timing of the gaps. Most syllable repeats are controlled by only one hemisphere. The transition probabilities at the branching points are controlled by both hemispheres.

The final part of the thesis develops a method that can wirelessly stimulate the neural circuit in behaving small animals. The device is tested in a zebra finch whose song is altered by electric pulses during singing. Potential experiments are proposed for studying the neural mechanisms of syntax generation using this device.

Taken together, all the four projects are dedicated to study the neural mechanisms of the syntactic sequence generation of birdsong.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The completion of this dissertation is not possible without the efforts of many people. First of all, I want to thank my advisor, Dr. Alexay Kozhevnikov. Alex has been extremely patient, supportive, and kind to guide me throughout my work at Penn State. His wisdom and optimistic attitude towards life have saved me many times from the frustration of work and life. I wish I can remember his advice to be confident, precise and loud in the remainder of my life.

I would like to thank Dr. Dezhe Jin for his inspiring advice for the direction of this research. I am grateful to my collaborators and colleagues in his theoretical group - Jason Wittenbach, Phillip Schafer, Seth Hulsey and Sumithra Surendralal, for their great advice and help. Particularly, I appreciate Sumithra's careful proofreading to make this dissertation a readable work. My colleagues in AK lab, especially Bruce langford, Mike Skocik and Guerau Cabrera, have made tremendous contributions to setting up the equipment used in my research. I would like to particularly thank Bruce, my 6-year comrade, for his generous assistance in every little step of my research and his creative technical advice.

I am indebted to many people in this department. Dr. Richard Robinett spared his time to fill in countless forms for me. Carol Deering, Melissa Diamanti, and many staff in the office have given me a lot of help to make my graduate life easier. I also want to thank my dissertation committee members for giving me valuable suggestions in my comprehensive exam and defense.

I thank my friends Wei Fan, Xuesong Wang, Enshi Xu, Xu Lu, Na Wang, Zhou Yu, Xiao Liu and Sikai Zhu for walking through this journey with me.

I want to thank my parents for raising me up and giving me lots of freedom in choosing my life. I am deeply grateful to my mother, for her continuous support and always being a model for me to be strong, responsible and kindhearted. Finally, dear father, you are and will always be in my memory, with all the happiness you gave me.

Chapter 1

Introduction

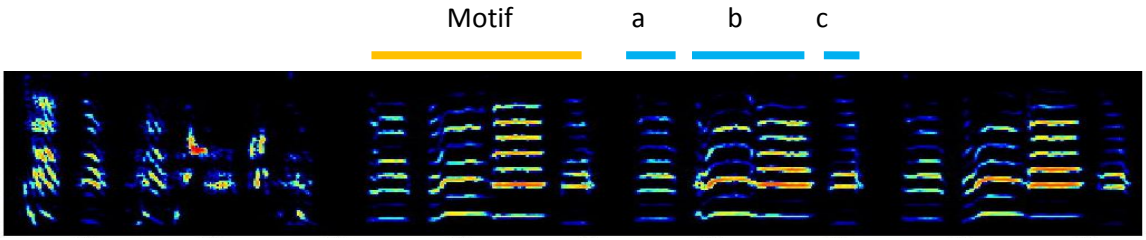
Sequential behavior is ubiquitous in humans and across various animal species. For example, a fly grooms its head, legs and body in a programmable manner; songbirds sing long bout of songs; and the humans can perform a variety of complex sequential movements, such as playing a musical instrument, speaking a language and so forth. Sequence learning is important for humans to acquire skills. How the brain generates complex sequences and how animals learn these sequences have been fundamental questions in neuroscience. Studies have shown that the sequence of actions are controlled hierarchically.[1,2] A typical example is the human language: a sentence consists of words that are grammatically linked, and a word is a combination of phonemes. Cognitive experiment on "artificial grammar learning" has shown that humans cannot learn a language if the strings of words do not follow a certain syntax.[3] The concept of syntax, as the rule that governs the sequence of discrete elements, can be generalized to other motor sequences as well.[4,5] How exactly the brain encodes the syntax of actions, however, has not been solved. The neural basis for generating simple motor sequences, such as locomotion, respiration and swallowing have been extensively investigated.[6–9] These types of rhythmic simple movements are controlled by central pattern generators (CPGs) that can generate oscillatory activity without rhythmic sensory or central input.[6,10,11] Long, complex movements, such as birdsongs and human speech, however, may not be simply explained by the CPG mechanism.[12]

The songbird has emerged as an excellent model system for studying the generation and learning of complex motor sequences.[13–15] Songbirds acquire their songs through learning. Like human infants learning to speak, the young birds learn the songs by imitating their fathers' songs.[15–17] Birdsong is comprised of a complex pattern of discrete sound bursts that are

structurally organized on various timescales.[18] The basic song elements are called syllables, which are several tens of milliseconds in duration and have identifiable temporal and spectral structures[18–23] (Fig. 1-1 b). The song complexity varies across species with syllable repertoires ranging in size from less than 5 syllable types such as swamp sparrows (*Melospiza georgiana*) to larger than 1000 syllable types such as brown thrashers (*Toxostoma rufum*).[24–27] The song of a zebra finch (*Taeniopygia guttata*), a songbird that has been most studied, is composed of repetitions of a few syllables that are strung in fixed sequences called motifs[18] (Fig. 1-1 b). The song motifs of a zebra finch have very little variation in duration and acoustic structure.[28] Their songs, due to the high stereotypy in sequence, are called linear songs. In contrast, canaries (*Serinus canaria*) have very variable songs[29–31]. A canary song is formed by blocks of syllable repeats and the blocks are linked by variable sequences[32]. The song of a Bengalese finch (*Lonchura striata domestica*, Fig. 1-1 a, c) lies between these two extremes - there is both a deterministic component and a stochastic component.[33,34] The stochastic component is characterized by a variable number of repeated syllables and multiple transition possibilities at the branching points.[33–36]

(a)

(b)



(c)

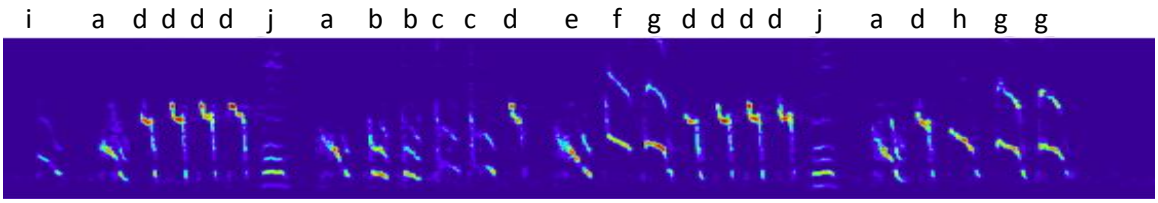


Figure 1-1. Zebra finch song and Bengalese finch song.

(a) Bengalese finch; (b) A segment of zebra finch song; (c) a segment of Bengalese finch song. Letters label the syllables.

Some birds, such as finches, start with highly variable songs called sub-songs.[19,37] Juvenile birds gradually improve their songs by listening to the adults and imitating their songs.[38] In about 3 months post-hatch, their songs are transformed to a sequence of highly stereotyped syllables in both acoustic structure and order.[39] Because these birds cannot learn new songs after

they are sexually matured, their syllable repertoires are fixed. They are called the age-limited learners or close-ended learners.[40–42] On the contrary, some birds, such as canaries, which can expand their repertoire by learning new songs even as adults, are called open-ended learners.[40] The Bengalese finch, a close-ended learner, is endowed with the capability to generate variable songs with its finite syllable repertoire, which allows us to study how the brain generates syntactical behavior that has a relatively simple mathematical representation.[33,34]

There have been a number of studies focusing on the neural mechanisms of song production in the zebra finch.[22,43–46] As is reviewed in the next section, song production in zebra finch is accomplished by the interplay between a few distinct brain areas with very stereotyped neural activities. Much less is known about the neural mechanisms behind how Bengalese finch produces the variable songs. The two finch species are closely related[47,48], indicating that similar neural mechanisms might also apply to the Bengalese finch[33]. Starting from what people have known about zebra finch, we can modify the model to make it produce more complex and flexible motor patterns. This thesis explores which brain areas of a Bengalese finch are involved in syllable sequencing using brain manipulation techniques.

1.1 Avian vocal system

Anatomy

Around ten songbird brain nuclei have been identified to be involved in singing. These nuclei are involved in two major pathways - a motor pathway and an anterior forebrain pathway (AFP) are found to be responsible for producing and learning the song[29,49] (Fig. 1-2). Each brain hemisphere has a complete set of these pathways.[50]

The motor pathway which consists of the premotor nucleus HVC (proper name), the robust nucleus of the arcopallium RA, and the tracheosyringeal part of the hypoglossal nucleus nXIIts, is organized in a hierarchical and feedforward way to control the vocal organs called syrinx.[29,49,51,52] The premotor nucleus HVC is the highest level of the vocal control pathway.[44,53] HVC projects to the downstream area RA, which in turn projects to the motor nucleus nXIIts that innervates the syringeal muscles. The AFP, starting from HVC and ending at RA, passing along the basal ganglia-like area X, which is homologous to the mammalian basal-ganglia the thalamic nucleus DLM and nucleus LMAN (the lateral magnocellular nucleus of the nidopallium), is found to be necessary for song learning but not for song production.[54–58] In addition to the feedforward pathways, there is also a feedback loop that connects the dorsal part of RA (dRA) back to HVC via the midbrain nucleus DM (dorsomedial intercollicular nucleus) and the pontomedullary respiratory areas[59], thalamic nucleus uvaeformis (Uva)[53,60,61] and nucleus interface of the nidopallium (NIf).[62–64] Unlike in mammals, birdbrains do not such structure that directly connects the two cerebral hemispheres homologous to the corpus callosum.[52] There is interconnection between the two sides in the brainstem including the respiratory area and DM, both of which projects bilaterally to the thalamic nuclei Uva[61,65,66]. Information from the two hemispheres can be exchanged in brainstem and the bilateral projections to Uva.

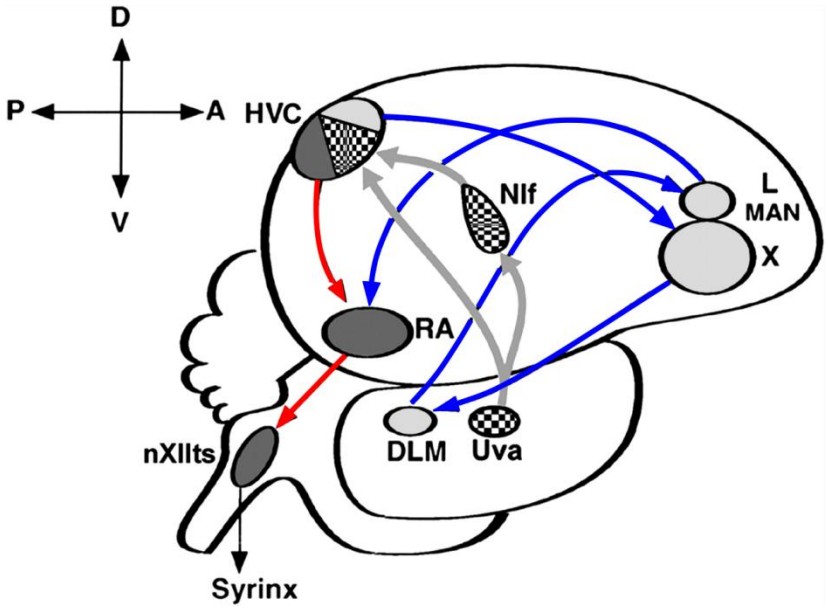


Figure 1-2. Vocal pathways in the brain of a songbird.

Red highlights the motor pathway; blue highlights the anterior forebrain pathway (courtesy of Eugene Akutagawa and Masakazu Konishi, 2005).

Function

The HVC has been found to encode the temporal structure of the song [45] and to be the driving force behind song timing.[46] Using antidromic activation, three kinds of neurons, including HVC-RA projection neurons (HVC(RA)), HVC-area X projection neurons (HVC(X)) and the inter-neurons (HVCi), can be identified.[45,67,68] The HVC(RA) neurons fire sparsely during singing - only once per motif in zebra finch - and are precisely time-locked to the song with a jitter for only about 1ms.[45] In zebra finch, a HVC(RA) neuron fires at a precise time of a motif.[45] The downstream RA neurons, whose firing pattern is also time locked to the song, however, fire at multiple time points. At each moment, a stereotyped subpopulation of RA neurons are turned on while others are off and in this way the acoustic structure is represented by the activities of RA ensembles.[45,69] Although multiple lines have pointed to feed-forward connections from HVC to

RA, reciprocal connections from RA back to HVC have also been observed.[70] How the RA activity can shape neural activity in HVC is not clear yet. The HVC(X) neurons display time-locked spiking activity during singing as well[68,71], but fire less sparsely so, compared to HVC(RA).[68] The HVC(X) neurons are believed to send a efference copy of the timing of the premotor sequence to the basal-ganglia-related circuit[68], where the song performance and the stored template may be compared and errors may be corrected.[17] The interneurons have tonic and non-sparse neural activity during singing that is time-locked to the song but the pattern is less stereotyped.[68] The HVCi neurons are believed to provide synaptic inhibition to the projection neurons.[67] Multiunit recording in zebra finch HVCs in both hemispheres shows short episodes of interhemispheric synchronization during singing.[66]

Forebrain lesion in juvenile birds leads to the development of abnormal songs.[54,55,72] Lesion of LMAN in adult zebra finch causes less variability of syllable structure, suggesting that the AFP provides exploratory variations to the motor pathway.[58,73,74] Controversial arguments have been made for the role of LMAN in generating the sequence variability. Hampton et al found that LMAN lesion in Bengalese finch adults reduced the syllable variability but did not affect the sequence variability.[75] However, Hamaguchi et al discovered that LMAN neural activity can propagate to HVC via the midbrain pathway, and augmented LMAN activity caused increase of sequence variability in zebra finch adults.[76]

Along the feedback loop from RA back to HVC, the respiratory brainstem network can drive HVC activity bilaterally via Uva.[77] It has been proposed that the respiratory feedback initiates the next syllable during the inter-syllable gaps and therefore it should determine the song sequence.[78]

As a major auditory input to HVC, NIf may hold the key of auditory feedback to HVC during singing.[62,63,79] Birds rely on auditory feedback to different degrees to produce songs.[80–87] Bengalese finch has real-time reliance on auditory feedback during singing.[86] Altering auditory

feedbacks affects song syntax, indicating a possible mechanism for syntax generation through the interplay between auditory feedback and HVC activity for syntax generation in Bengalese finch.[86,87] Lesion in NIf eliminates the song selective neural response in the HVC of zebra finch[62] and also reduces the song complexity of Bengalese finch.[88]

1.2 Theoretical models for birdsong production

The neural mechanisms underlying sequence generation have been studied for years. In one of the models, the sequence is thought to be generated with respect to a common clock. For example, the sequential activation of place cells in the hippocampus is phase-locked to the theta-oscillation.[89–91] In another model, the sequence is thought to be successively activated in a chain-like manner.[92–95] Although oscillations have been observed in HVC in vitro[96], intracellular recording in HVC(RA) neurons of singing zebra finches[97] and the HVC cooling experiment[46] has provided evidence for the chain model.

The neural sequence in HVC was reproduced by a synfire-chain model in which neural impulses propagate through groups of HVC(RA) neurons successively and drive the downstream nuclei RA for moment-to-moment vocalization.[95,98]

The synfire-chain model alone, however, can only generate fixed sequences like zebra finch songs. For Bengalese finch songs, new models are needed to describe the origin of the sequence variability, which has not been clearly identified by experiments. Ideally, if one can record HVC(RA) activity in Bengalese finch during singing, direct connection between the neural activity and sequence variability may be revealed. However, to date, only HVC(X) activity has been recorded in singing Bengalese finch.[71] Various computational models have been proposed, with the source of sequence variability being focused on 1) the competition between HVC(RA) neurons[99–101]; 2) active noise input from other areas[102]; 3) auditory feedback[103].

The Bengalese finch is a mutant strain of the white-rumped (*munia Lonchura striata*), a bird that is actually a linear song singer.[33] Some mutation must have occurred to one or more of the vocal control nuclei to cause sequence variability. Experiments have shown that the zebra finch can be fostered by Bengalese finch parents and learn songs from them.[47,48,71] Moreover, one of the features of Bengalese finch song, the variable number of repeated syllables, can be acquired by zebra finch as well.[71] Therefore, the neural network that is thought to be well-understood in zebra finch should be able to extend to the Bengalese finch. It has been shown computationally that, with a few modifications, a chain network of neurons can generate probabilistic sequences.[101] The synfire-chain model is modified in such a way that, at the branching points, multiple chains compete with each other via mutual inhibition mediated by the interneurons and finally only one chain can propagate. The transition probability is determined by the synaptic strength between the branching chains. The model also allows self-transitions so that the syllable repeats are possible. Each synfire-chain represents a state in a Markov chain. To generate non-Markovian songs, same syllable can be represented by different states of a Markov chain, so there is a many-to-one mapping from neural chains to song syllables. Sometimes, the lengths of the observed syllable repeats follow a bell-shaped distribution, which cannot be captured by a Markov model.[34,71] To see this, suppose we have a two-state Markov chain, for which P(A|A) = p and P(B|A) = 1-p. The probability to have an n repeat sequence given the first is A is $P(A^{n-1}B|A) = p^{n-1}(1-p)$, which follows a geometric distribution but not bell-shaped. The repetition distribution can be realized by imposing a synaptic adaptation on the probability of the recurrent transitions.[34] Auditory feedback could also bias the transition probability via NIf.[101]

The chain model for zebra finch has been well supported by experiments. In addition to that, we want to test in Bengalese finch if the branching chain model is correct. In this work, I use various perturbation methods to tackle this question.

1.3 Methods to study the neural basis for song production

Two methodologies have been generally used to study the brain. One is direct observing the neural activity using electrophysiological methods and imaging. Thanks to the developments in finer probe fabrication, microscopy, stain methods and so on, single neural activity can be monitored in behaving animals[104,105] and population activity can be visualized using fluorescent dye under two-photon microscope[106–108] or using fMRI.[109,110] Before the technology was available to observe the neurons directly, an indirect way was commonly used to relate the brain area with its function by manipulating the brain. Many early studies have been done by brain lesion.[111] However, the irreversible damage of brain tissue may cause side effects that make the conclusion ambiguous. Later on, as the neuron's electrophysiological basis was revealed, electrical stimulation was widely used to transiently alter the neural activity in a controllable way.[112,113] Biochemical studies on the interaction of the neurotransmitters allow to locally change a population of the neurons' behavior using pharmacological methods.[114] Recently, genetic tools have given researchers the freedom to vary the expression of certain neuron-related molecules so as to investigate the neural mechanisms in a more microscopic level[115], and also later on, derived the powerful optogenetic tools that allow selectively activation or inhibition a subpopulation of neurons by light in behaving animals.[116] Temperature manipulation has also been used to locally change the dynamics of neural activity.[117–121,46,122,123] Compared to brain lesion, temperature manipulation is non-damaging and reversible; compared to electric stimulation, temperature can target a larger population of neurons and give them a persistent effect; compared to pharmacological methods, temperature is much easier to monitor than the concentration of chemicals.

In the songbird field, all the above methods have been used for various purposes. The branching chain model predicts that HVC neural activity itself can generate the sequence

variability.[101] To test this, one can manipulate the temperature of HVC without affecting other areas and see if the syntax could be changed. If the syntax is not affected by the manipulation, there would be two possible explanations: the perturbation cannot effectively alter syntax; or some input from some higher order nucleus is enforcing the syntax. On the other hand, if the syntax is changed due to this manipulation, it has strong implication that HVC plays an active role in syntax generation.

By locally cooling HVC, a series of chemical reactions, such as the ion channel open/close rate, are slowed down, resulting in a slowed propagation velocity through the chain network and consequently, a slowed song production. This result was observed in the zebra finch whose songs are linearly stretched by cooling.[46,122] In zebra finch, the songs are deterministic, indicating a chain connected in a linear fashion with no splits and thus no change in syntax by cooling would be expected. However, in Bengalese finch, the transition probabilities between two chains could be changed due to the change of relative synaptic strengths by temperature. For example, the short-term synaptic plasticity[124], shaped by the pre-synaptic calcium concentration, vesicle depletion, neurotransmitter reuptake, receptor saturation, etc, could be affected.

To achieve a finer investigation on the moment-to-moment singing control by the two HVC hemispheres, one can use electric stimulation that is delivered to either side of the brain at a time during singing[125], especially at the transitions where repeats or branching points occur. The link between perturbed HVC activity and possible syntax change in this way would also be an evidence for HVC encoding syntax.

In addition, direct neural activity recording during singing can provide important information about whether the syntax is encoded within the neural activity in HVC.[45,71] If, for example, different spiking pattern is observed on the same neuron at the same syllable type that presents at different locations of a song, the many-to-one mapping can be proved.

1.4 Summary of work

This work has been focused on using the perturbation methods to locate the brain areas that are involved in generating the syntax of the song.

Chapter 2

How syntax is encoded in songbird brain is disputable. The focus of debate has been whether HVC, the high vocal center that is responsible for the timing control of the song, can encode the song sequence by its own dynamics[101] or has to be instructed by a higher order area[88]. In Chapter 2, I attempted to answer this question by carrying out temperature manipulation in the HVC of a Bengalese finch and analyzed the change of song syntax. I found that in addition to stretching the song timing, temperature also affected the syntax in terms of the distribution of repeat length and the transition probabilities at the branching points. The repeat lengths were reduced by cooling down HVC and increased by warming it up. Some of the transition probabilities at the branching points were altered by cooling. These findings are consistent with the idea that the syntax is encoded within HVC. To rule out the possibility that the downstream area RA may affect the syntax through the reciprocal connections back to HVC, I carried out cooling in RA. The syntax was not found to be altered with RA cooling.

Chapter 3

Auditory feedback has been suggested to affect the syntax of Bengalese finch song.[86,87] We hypothesize that auditory feedback is important to maintain the bell-shaped distribution of the repeats. In this chapter, I explored the role that auditory feedback may play in the repeats using real-time distorted auditory feedback (DAF). The repeat length was reduced with DAF presented.

This effect resembles the cooling effect on repeats and suggests that the effect of cooling on repeats are auditory feedback mediated.

Chapter 4

There is high inter-hemispheric coordination between the two HVCs, even though they are not directly connected. Episodes of synchronized activity by multiunit recording were observed in the zebra finch.[65] The left and right HVCs were found to quickly switch their dominance in controlling singing.[46,125] In Bengalese finch, lateralization of the syrinx movement was observed[126], indicating an asymmetric control by the two HVCs. The branching chain model also indicates that both timing and syntax control should be dominant in only one hemisphere.[101] In Chapter 4, I further study the role of the two HVC hemispheres using unilateral cooling. I found that the timing of the syllables was only affected by left HVC cooling while the syntax was affected by both sides.

Chapter 5

Electric stimulation has been an important tool for studying the moment-to-moment singing control by HVC.[53,125] In Chapter 5, I describe the design of a simple wireless device that can deliver electric pulses in behaving small animals. This device was tested in a zebra finch. Songs were successfully interrupted by electric stimuli. A paradigm about how to stimulate either hemisphere at a random time of a song is described.

In conclusion, all the projects were dedicated to study the mechanisms for syntax generation in the songbird.

Chapter 2

Neural substrate of birdsong syntax identified by cooling the brain

Syntactical rules underlie temporal order in many behavioral sequences, such as human language and birdsong. The neural substrates of syntax are poorly understood. Here we use reversible temperature change to study the role of the neural activity within the premotor brain area HVC (proper name) of a songbird in syntactic organization of the song. Changing HVC temperature alters the song syntax: it affects the distributions of the lengths of most syllable repeats and changes the syllable transition probabilities of some syllable branching points. The finding strongly implicates HVC as a key player in the song syntax generation. Our results demonstrate the power of local reversible manipulations of neural circuitry for localizing the seat of syntax in the brain.

2.1 Introduction

Syntax is a critical aspect of human language that enables generation of virtually unlimited sequences of words with a finite set of rules.[127] Other behavioral sequences in humans and animals, such as music[128], birdsong[129], songs of rock hyraxes[130], and grooming in rodents[131] also follow syntactical rules. Understanding the neural mechanisms of syntax across animal species and humans might shed light on the evolutionary origin of human language, and help to treat speech and language disorders such as stuttering or agrammatism. Where and how syntactic rules are encoded in the brain remains a critical unsolved problem in systems neuroscience. In human brain, the brain loci controlling speech syntax have been searched using functional magnetic resonance imaging[132], positron emission tomography[133], and brain lesions in aphasic patients[134,135]. In animal systems, lesioning has been the main method for locating the sites

controlling syntax.[88,131] However, lesions can disrupt the brain circuitry beyond the site of lesion, for example, by damaging the fibers passing through the lesion site. Furthermore, brain circuitry may have compensatory reactions to lesions, obscuring the true role of the lesioned site in the intact brain. Imaging techniques, while noninvasive, are correlational and lack direct causal proof that a brain site is enforcing syntactical rules. In this work, we use rapid and reversible manipulations of local brain temperature in an intact brain to establish a causal link between the songbird forerbain premotor area HVC (used as a proper name) and birdsong syntax.

Songbird has emerged as a model system for investigating the neural mechanisms of production and learning of vocal sequences.[14,136–139] Birdsong consists of sequences of stereotypical acoustic elements – song syllables (Fig. 2-1a, b). Syllable sequences of many species are variable and follow syntactic rules.[32–34] The syntax of the Bengalese finch song can be characterized by a syllable transition diagram (Fig. 2-1c). The syllable transition has branching points: a syllable can be followed by one of several different syllables with different probabilities.[140] Some syllables can also transition to themselves and repeat a variable number of times.[141] Longer repeats are usually more variable.

A forebrain premotor area HVC has been shown to be critical for singing[29] and to play a prominent role in defining the fine scaled timing of acoustic features in the song.[28,43,45,68,69,97] Brief bursts of spikes of HVC projection neurons time-lock to the acoustic features down to 1ms, making the HVC neural activity one of the most temporally precise patterns known to be generated by neural circuits.[45] Manipulations of neural activity in HVC were shown to alter the song tempo.[46] These observations are consistent with the idea that the propagation of spikes on synaptic chains in HVC is encoding the song timing down to the millisecond timescale.[95,97] Where and how birdsong syntax is encoded is not understood, and the role of HVC in syntax is debated. Auditory feedback can have an effect on the song syntax[142,86,143]; also, neural correlates

of the syllable sequence have been observed.[71,144] However, no direct causal link between activity in any brain area and syntax has been found.

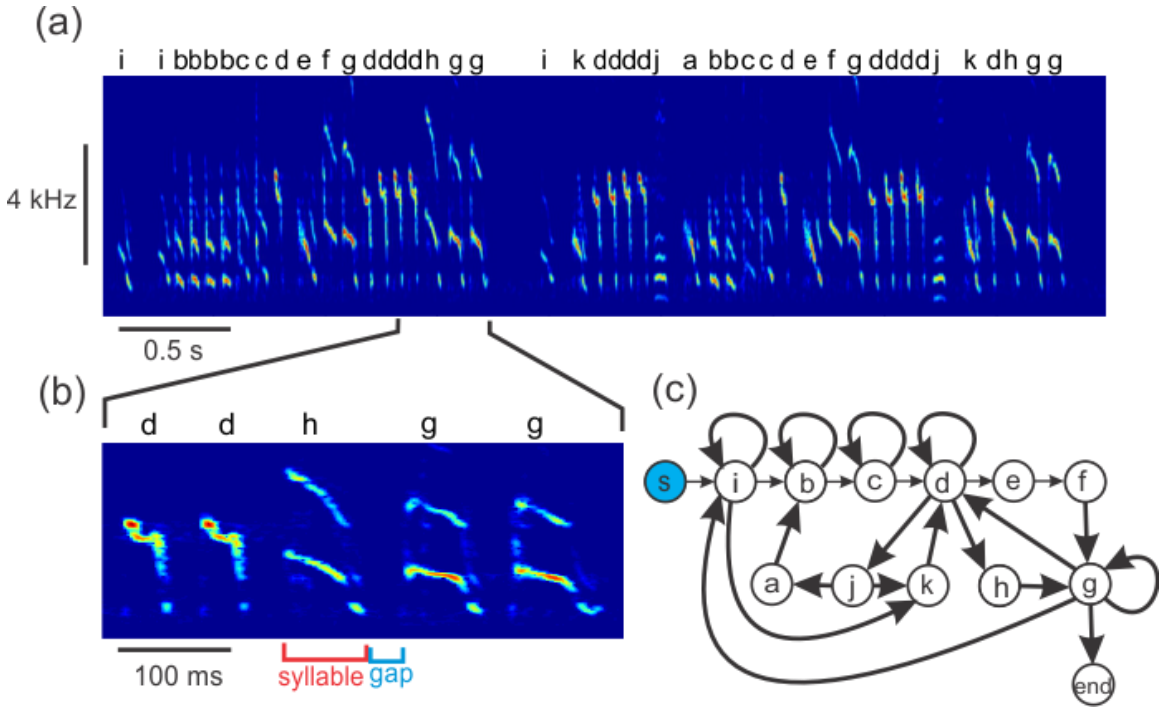


Figure 2-1. Bengalese finch song structure on several time-scales

(a) The song bout consists of a sequence of song syllables (labeled above the plot with letters a through k). (b) Song syllables are separated by silences (syllable gaps). (c) The song syntax can be visualized by a transition diagram where arrows show the allowed transitions between the syllables.

Several distinct candidate mechanisms of syntax generation in birdsong have been proposed. In the hierarchical model, song syllables are encoded in HVC and are activated by input from the higher brain area (for example, nucleus interfacialis (NIF)).[88] In this model, HVC encodes the song syllables but plays no active role in the syntax. An alternative possibility is that HVC itself is the primary site of syntax generation.[101] In this picture, synaptic chains of HVC

projection neurons encode the song syllables; the chains are connected into branching patterns. Competition between the chains at the branching points through a "winner-take-all" mechanism is generating variable sequences. The song syntax is encoded in the connection patterns of the synaptic chains.[101] It has also been proposed that the neural feedback through the brainstem is predominantly responsible for the syntax.[78]

Our experiments directly probe the role of HVC in song syntax through real-time reversible manipulations of HVC temperature in the Bengalese finch. As will be shown below, cooling HVC reduces the length of syllable repeats and changes the transition probabilities of branching points. HVC is thus critically involved in generation of the song syntax. These results cast serious doubt on the hierarchical model. Our findings are consistent with the branching chain network model, modified to allow auditory feedback to influence syllable transitions.

2.2 Methods

2.2.1 Manipulations of HVC temperature

We engineered a device for reversible local temperature manipulation of the brain, similar to a recently described device.[123] Briefly, the device consists of a Peltier thermoelectric element, gold-plated contact pads, and a heatsink. The contact pads of the device are placed on the surface of the brain, and the device is attached to the skull by dental acrylic (Fig. 2-2 a, c). Application of a direct current to the device causes changes of the brain temperature in the vicinity of the cooling pad. The temperature of HVC is monitored by a miniature thermocouple inserted into the HVC at a depth of 0.5mm (Fig. 2-2 b). All procedures are carried out in accordance with the IACUC protocol approved by the local Institutional Animal Care and Use Committee.

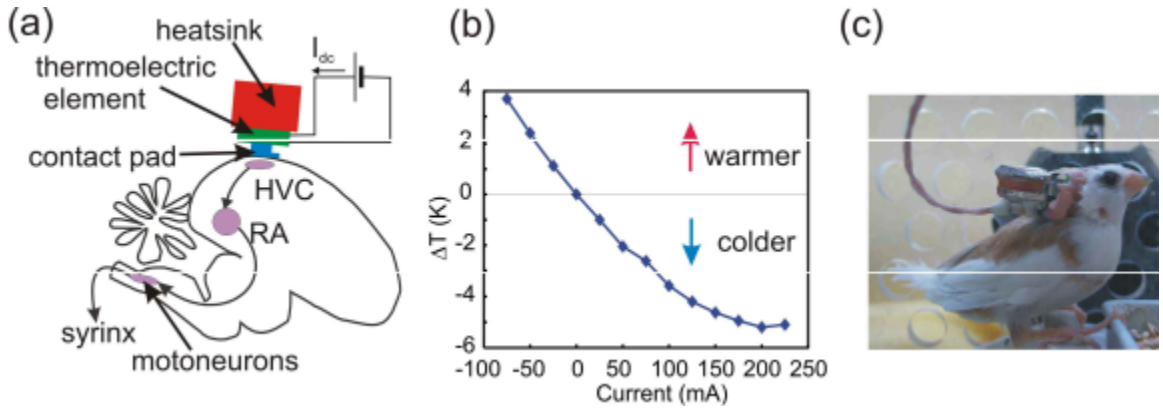


Figure 2-2. Cooling device

(a) Schematic of the experiment. The contact pad is placed on the surface of the brain and its temperature is altered by electric current flowing through the thermoelectric element. (b) The dependence of temperature within HVC (at a depth of 0.5mm) vs. the direct current applied to the thermoelectric element. (c) The device on a freely behaving Bengalese finch.

2.2.2 Changing temperature in RA

The cooling probes to be inserted into RA were made by gold-plated silver wires that were 300µm in diameter and 4.5mm in length, with 4mm separation (Fig. 2-3 a). The probes were coated by polyimide tubing with 510µm inner diameter and 560µm outer diameter. Small amount of Torr Seal was used to glue the tubing to the silver probe. A bare tip around 0.5mm long was left to contact the brain tissue for cooling. Temperature drop at the tip is about one half of the temperature drop at the cold side of the Peltier module, yielding the feasibility to use the same Peltier module with copper mesh heat sink to cause sufficient temperature change in RA. A very fine thermocouple (IT-24P, Physitemp Instruments, 230µm in diameter) was inserted in the same tubing along with one of the probes into the brain. The distance between the tip of the thermocouple and the tip of the probe is about 0.25mm.

The cooling device was implanted in an anaesthetized bird with probes bilaterally inserted into RA, at a beak angle of 60°, about 1mm posterior and 2mm lateral to the midsagittal bifurcation and 1.7mm in depth. We demonstrated that the probes reached RA and effectively changed temperature in RA using electrophysiology. After the device was fixed on the skull with dental acrylic, the head angle was rotated backward by 15°. RA area was probed by a carbon fiber electrode (Carbostar-1, Kation Scientific) inserted about 200μm anterior to the spots where the right cooling probe was inserted (Fig. 2-3 b). The spontaneous neural activity was recorded with different currents applied to the Peltier module (TE-18-0.45-1.3, TE Technology, Inc). To eliminate the AC pickup by the power supply to the cooling device that severely reduced the signal to noise ratio, a 9-volt battery with an adjustable resistor was used as the current supply. The neural signal was amplified by 100 times and filtered with a cut-off frequency at 2kHz (FLA-01, Cygnus Technology) and then fed into an A/D converter (USB 6211, National Instrument) and recorded at a sample rate of 40000 Hz. After the neural signal was taken and the effect was confirmed, the bird was put back into the cage and allowed to wake up.

To estimate the temperature change in RA, we first took the temperature measured by the thermocouple that was implanted along with one of the probes. To avoid damaging RA, the probes were inserted right across the edge of RA (where RA activity was first observed). We assumed an exponential decrease of temperature change with distance. If we model RA as a sphere with 0.5mm radius, the temperature change at the center of RA would be about $e^{-0.5/\lambda}$ of the amount measured by the thermocouple. According to Long et al (2008)[46], the distance constant λ is about 1.2mm. Therefore, we estimated the temperature change in RA to be 0.66 times the measurement, yielding a range between -5K to 3.5K in our experiment.

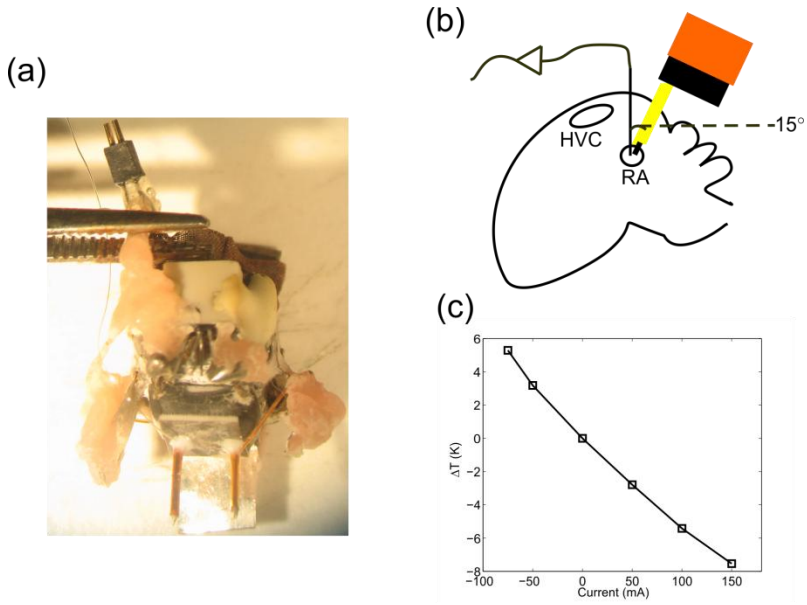


Figure 2-3. Device for RA cooling

(a) Photo of the probes on the cooling device. (b) Surgery illustration. (c) RA temperature measured by the thermocouple vs dc current applied to the Peltier device.

2.2.3 Histology

After the experiment was done on a bird, the brain was taken for histology to confirm that the RA probes were inserted at the right spots. The brain tissue was submerged in the fixation solution (4% buffered paraformaldehyde) for two days. Brain slices of 100 μm thickness were obtained using Oxford vibratome sectioning system. The brain areas were confirmed with the trace of the probes under the optical microscope.

2.2.4 Data acquisition and analysis

Undirected songs were recorded (16-bit resolution, sample rate 40 kHz) and analyzed off-line. Recording was triggered when the acoustic intensity crossed the threshold. Recordings

shorter than 2 seconds were discarded immediately. The songs were then screened offline based on the rhythm of the acoustic profile. To obtain the song rhythm, the raw signal was first converted to the logarithm scale. The signal was then decimated with a cutoff at 30 Hz, after which, fast Fourier transform (FFT) was performed on this signal. A song was recognized if its mean spectrum amplitude between 8Hz and 10Hz was above a threshold. Segmentation of the song into song syllables and inter-syllable gaps was done based on the envelope amplitude crossing a threshold. The separations that were less than 5ms were merged. Syllables that were shorter than 5ms were discarded. A finer search for the actual syllable boundaries was made by finding the maxima around the detected onsets and minima around the offsets of the first order derivatives of the acoustic power. Segmentation was then manually checked and corrected.

About 10 songs were classified manually based on the syllable structures. The spectrogram of a song was converted to a binary version to reduce the effect of amplitude variation. The first 65ms (as a mean of all syllable durations) of the syllables was used for classification. Those that were shorter than 65ms were padded with zeros. The spectrogram segments of syllables were vectorized into a single dimensional, and were used for training a support vector machine (SVM). The trained SVM was then used to predict the classification of other datasets from the same bird. After the songs were sorted automatically, an offline inspection was made to correct any wrong classifications (see Appendix A for details).

To assess the effects of temperature changes, we compare the durations of the song syllables, the durations of the inter-syllable gaps and the sequences of song syllables in the songs under the control condition and when HVC temperature was altered. To eliminate the possible long-term changes in the song, data comparison is done between the control data and the test condition data collected on the same day. Time intervals when brain temperature is altered are interleaved with the time intervals when the brain temperature is not changed to ensure that temperature manipulation does not cause irreversible song changes.

Fractional stretch of the duration of song elements

At each temperature, the mean duration of each song element (each type of song syllable and each type of inter-syllable gap) is computed. For each song element, a least squares linear fit of the mean duration vs. temperature change is performed; data at each temperature are weighted by the inverse square of the standard deviation. The slope of the least squares fit is divided by the mean duration of the song element in the control condition to give a fractional stretch of the song element, expressed in per cent per degree Kelvin (%/K).

Analysis of temperature effects on the length of syllable repeats

We computed the mean lengths of the repeats for all the repeating syllables at different temperatures. For each syllable repeat, we computed the least squares fit of the mean repeat length vs. temperature. The slope of the fit was divided by the mean length of the repeat in the control condition to give the fractional change of the repeat length vs. temperature, expressed in per cent per degree Kelvin (%/K). To assess whether the syllable repeat length is significantly affected by temperature, we tested whether the fractional change of the repeat length is significantly greater than zero. We used a one-tailed t-test and significance level $p= 0.05$ as a criterion for the slope of linear regression to be significantly different from zero.

To assess the correlation between the variability of the repeat length and the cooling effect on repeat length, we calculated the standard deviations of the repeat lengths from all the control data. We computed a linear fit of the rate of change with temperature vs. the standard deviation of the repeat length (Fig. 2-5 d). We used a t-test to test that the computed slope is significantly different from zero and, thus, that there is a statistically significant correlation between the standard deviation of the repeat length and the fractional change of the syllable repeat length.

To test if the change of repeat length is due to the conservation of the repeat duration, we simulated the repeat length distribution in this scenario. The simulated repeat length distribution

was generated from the distribution of the total repeat duration in control condition as $\{N(\Delta T)\} = \{[\frac{D^0_{repeat}}{1+\Delta T\alpha}]\}$, where $\{D^0_{repeat}\}$ is the distribution of the repeat duration of the control group and α is the fractional change of timing by temperature, which is about -3.5%/K by taking the average of syllable stretch and gap stretch. The simulated distribution was compared to the experimental result using a one-tailed t-test.

Analysis of temperature effects on the branching points

To construct the transition probability matrix, we count the number of transitions from syllable i to j (i≠j) and divide it by the total number of transitions from syllable i, so that the elements in the transition probability matrix of each row sum up to 1. To make the transitions into an irreducible Markov chain, a start/end state was introduced so that the sequence chains were cyclic. The introductory notes were merged into the start/end state to eliminate their effect on the sequence of the song. To separate the effect on repeats from the effect on branching points, we did not count self-repeating transitions and treated the repeats as a single 'repeated syllable'. Thus, the diagonal elements of the transition matrix are set to zero. To assess the difference between the sets of transition probabilities at each branching point of the control songs and the songs cooled by 4K, we used two-sided Fisher's exact test[145] for the branching points that have rate transitions and Pearson's χ^2 test for the rest. We ignore the branches that occur very rarely: we only consider the branches that occur at least ten times in both conditions. The branching points are defined as those that have more than one type of transition after we omitted the rare ones. For RA cooling data, we used Pearson's χ^2 test to see if there is dependence of the transition frequencies on the temperature.

For those branching points that were significantly affected by temperature, we confirmed the effect by carrying out a day-to-day comparison. We constructed the "$2 \times c$" table for each day with the rows as the temperature conditions ("control" and "cooled") and the columns as the

branches. Cochran–Mantel–Haenszel (CMH) statistic[146] was used to test if there was consistent change of transition probabilities by temperature across days. The CMH statistic, as an extended χ^2 statistic, testing for repeated categorical data, in the case of a "q 2×2" table, is defined as $\chi^2_{MH} = \frac{\{\sum_h (n_{h11} - n_{h1\cdot} n_{h\cdot 1}/n_{h\cdot\cdot})\}^2}{\sum_h \frac{n_{h1\cdot} n_{h2\cdot} n_{h\cdot 1} n_{h\cdot 2}}{n_{h\cdot\cdot}^2 (n_{h\cdot\cdot} - 1)}}$, with 1 degree of freedom. We used the odds ratios between the two groups to visualize the consistency of the day-to-day variation of the probability distribution. For example, if the branching point has two-way branches, the odds ratio between cooled and control groups would be $OR = \frac{p_1^{cool}/p_2^{cool}}{p_1^{control}/p_2^{control}}$. We used the natural logarithmic scale of OR to make the plots and the corresponding standard error for $\ln(OR)$ is $Se_{\ln(OR)} = \sqrt{\frac{1}{p_1^{cool}} + \frac{1}{p_2^{cool}} + \frac{1}{p_1^{control}} + \frac{1}{p_2^{control}}}$. If there are more than two branches in a branching point, the subgroup that has the least change by temperature is merged.

Analysis of sequence variability

We approximate the song sequence as a Markov chain in which the selection of a syllable is only dependent on the immediately preceding one. For the long repeats, reduction of repeat length by cooling will reduce the very biased transition probability of the self-transitions and consequently increase the transition entropy. To separate the entropy changes due to change of transition probabilities of branching points, but not due to syllable repeats, we merged the sequences of repeated syllables into one repeated syllable by setting the diagonal elements in the syllable transition matrices to zero. We used transition entropy, defined as $H_i = -\sum_j P_{ij} \log_2 P_{ij}$, to quantify the variability of the transitions at the i[th] branching point. To compare the change of transition entropies between control and cooled songs, we bootstrapped the distribution of the transition entropies by carrying out the following steps:

1) Count the number of transitions $\{N_{ij}^{control}\}$ and $\{N_{ij}^{cool}\}$.

2) Resample the counts to have the same N_{total} as the actual data but are with replacement.

3) Calculate the transition entropies for each condition using the resampled transition probabilities.

4) Compute $dH = H^{cool} - H^{control}$.

5) Repeat step 2-4 for 1000 iterations. If the 5% quantile of the bootstrapped distribution is greater than zero, the increase is considered to be significant.

For the transition entropy at different temperatures, we find the slope of transition entropy vs. temperature for each branching point. The trend is considered to be significant if the p-value of the t-test on the slope significantly smaller than 0 is less than 0.05.

The principal transition probability in a branching point is the probability of the most frequent transition. The rest of the transitions are defined as minor transitions.

Analysis of the RA electrophysiology

Three trials of RA activity were taken for each temperature condition. Each trace was 20s long. Spikes in a trace were automatically detected by threshold crossing and were plotted in one graph aligned with their onsets. The spikes that had the same waveform were selected manually in a Matlab GUI and the spike times were stored. The real time spike rates were estimated by calculating the inter-spike-intervals and taking the inverse. The mean and standard deviation were obtained from the real time spike rates. The slope of spike rate vs. temperature change was estimated by linear regression.

2.3 Results

2.3.1 HVC temperature effects on timing

We made a device for reversible local brain temperature manipulation and used it to change HVC temperature in Bengalese finches and study the effects of temperature changes on the song (see Methods). Cooling HVC increases the duration of the song syllables and the inter-syllable gaps of the Bengalese finch song (Fig. 2-4 a,b). Overall, these effects of temperature on the song are similar to the effects previously observed in zebra finches.[46,122] Cooling increases the durations of song syllables by 2.8+/-0.9 %/K (mean +/- s.d., N=34 different syllables in 5 birds). Durations of the inter-syllable gaps are increased by 4.2 +/- 2.5 %/K (mean +/- s.d., N=50 inter-syllable gaps in 5 birds). The fractional duration change of the intersyllable gaps is significantly larger than that of song syllables (p=3x10-4, Kolmogorov-Smirnov test).

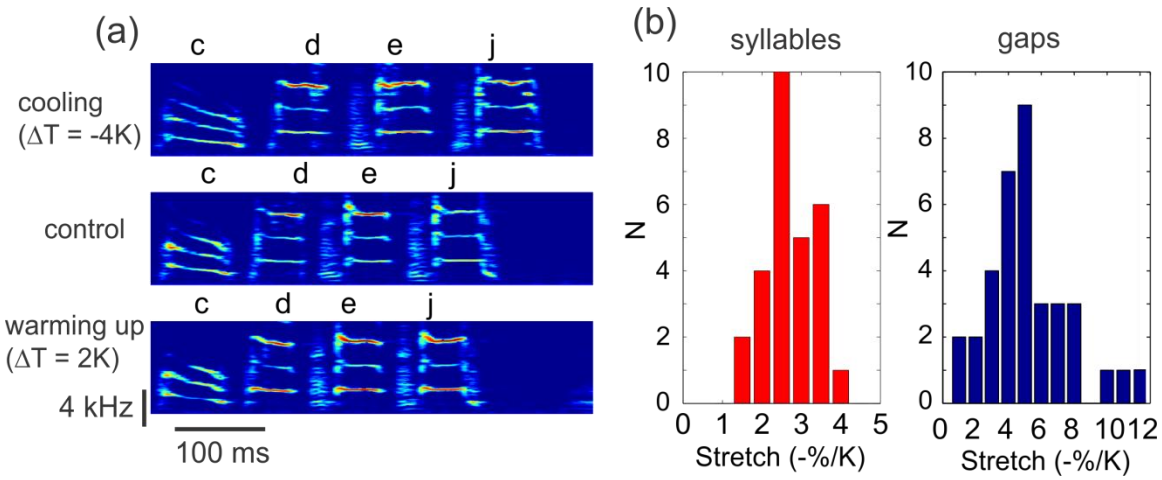


Figure 2-4. Cooling effect on song timing

(a) A song segment consisting of 4 syllables at different HVC temperatures, aligned to the onset of the first syllable. Decreasing HVC temperature increases the syllable durations whereas increasing the temperature reduces the durations.(b) Histograms for the fractional

duration change of the song syllables (left) and inter-syllable gaps (right) for all the syllables and gaps studied.

2.3.2 HVC temperature effects on syntax

In addition to affecting the duration of the song elements, changing HVC temperature changes the syntax of the song. An example of the effect of temperature on the song syntax is shown in Fig. 2-5 a. The song has a repeat of syllable A, and cooling HVC decreases the number of times syllable A is repeated, or the repeat length (Fig. 2-5 a, b). Increasing HVC temperature has the opposite effect – the repeat length of the syllable is increased.

Fig. 2c shows the mean repeat length vs. temperature for all the repeated syllables studied (N=13 repeated syllables in 5 birds). We use the slopes of the repeat length with respect to temperature (Fig. 2-5 c) for each syllable repeat to quantify the effect on the repeat length change (see Methods). Overall, the repeat length of the syllable changes by 4.4 +/- 4.2 %/K (mean +/- s.d. for all the repeated syllables). A statistically significant decrease of repeat length with cooling was observed for 6 syllable repeats out of 13 repeats in 5 birds ($p<0.05$, t-test on whether the slope of regression is greater than zero, see Methods). The degree to which the repeat length is affected by temperature is correlated with the repeat variability and with the mean repeat length. The length of the shortest repeats is not altered when the temperature is changed, whereas majority of the longer repeats display a pronounced length change (Fig. 2-5 c).

For Bengalese finches, longer syllable repeats tend to be more variable in length, so the degree to which the length of the syllable repeat is affected by temperature is also correlated with the variability of the syllable repeat. Fig. 2-5 d shows the dependence of the rate of the change of the repeat length on the standard deviation of the repeat length. Repeats having a larger length variance are more strongly affected by temperature change, the effect is highly statistically

significant: the slope of the fit is significantly different from zero (p=5×10-6, n=13, t-test on the whether the slope is greater than 0).

Repeat length could be reduced if the time of the whole repeat (time between the onset of the first syllable of the repeat to the offset of the last syllable of the repeat) had to be constant while individual syllable is stretched. However, this possibility cannot account for the change of repeat length by temperature. To address that, we first show that the total time of the repeat is not conserved (Fig. 2-6 a). 5 of the 6 syllable repeats that are affected by temperature also display a significant change in total duration with temperature. On the contrary, 4 short syllable repeats that are not altered by temperature show a significant decrease of total time with temperature due to the shortening of the syllable and inter-syllable gaps. These observations strongly suggest that the repeat length is not constrained by the timing of the whole repeat. We next show that the distribution of the repeat length, if constrained by the total duration, would yield a smaller effect than what we observed. To simulate the time-conservation model, we generated the distribution of repeat length at different temperature from the control data using $\{N(\Delta T)\} = \{[\frac{D^0_{repeat}}{1+\Delta T\alpha}]\}$ (see Methods). As shown in Figure. 2-6 b, by cooling HVC by 2K, the distribution of repeat length is significantly shifted to the left of the simulated distribution ($p = 1.3\times10^{-11}$, one-tailed t-test). Therefore, our observation suggests that the change of repeat length by temperature is not due to the necessity to compensate the total duration of the repeat but indeed a change of the number of the repeats.

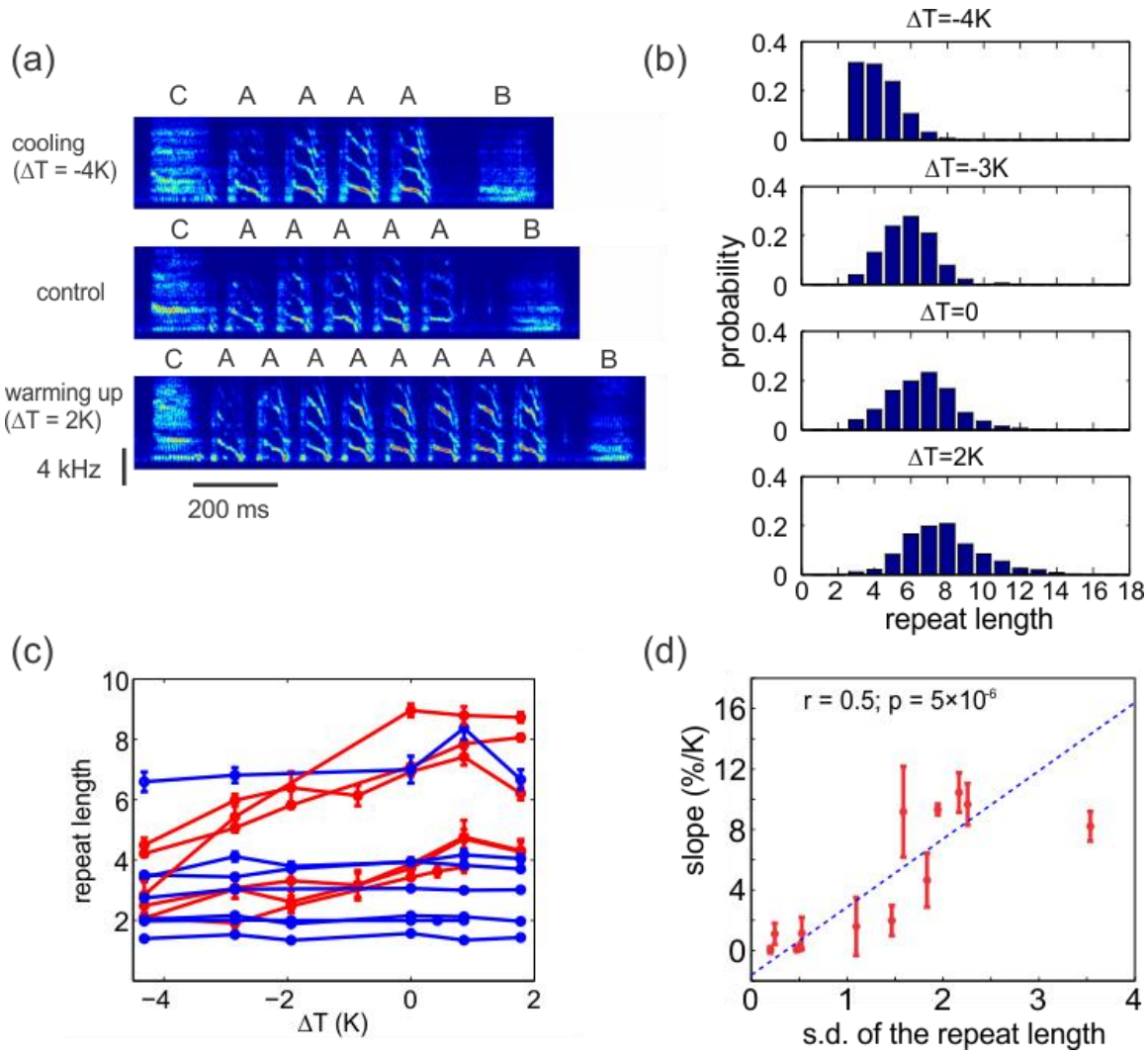


Figure 2-5. Effects of HVC temperature on repeats

(a) Representative song segments at different HVC temperatures, aligned to the onset of the first syllable A. Decreasing HVC temperature increases the length of the repeat of syllable A. (b) Histograms of the repeat lengths of syllable A at different HVC temperatures. The temperature change is indicated at the top of each plot. Cooling systematically shifts the distribution, decreasing the mean number of repeats.(c) Dependence of the repeat length on temperature for all repeated syllables. Error bars are s.e.m. Red curves indicate the syllable repeats significantly ($p<0.05$, linear regression t-test for determining if the slope is different from zero) affected by temperature; blue curves indicate the syllable repeats which are not significantly

affected by temperature. Shorter syllable repeats tend to be unaffected by temperature whereas most of the longer repeats are significantly affected by temperature. (d) Dependence of the change of the repeat length with temperature on the standard deviation of the syllable repeat length. The slopes are calculated from the 13 curves in (c) divided by the mean repeat length of control. The plot shows that more variable syllable repeats (having larger s.d.) are more strongly affected by temperature. The slope of the fit $\beta = 4.5$, it is significantly greater than zero ($p=5\times10^{-6}$, n=13, t-test on the slope of linear regression).

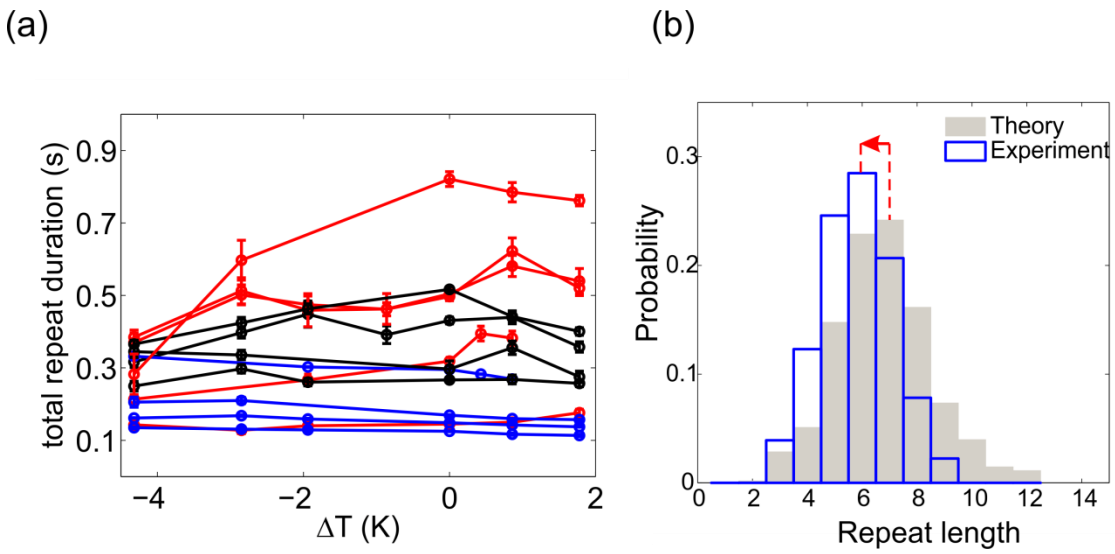


Figure 2-6. Change of repeat duration

(a) Change of repeat duration with temperature. Red curves have positive slopes (n=5) and blue curves have negative slopes (n=4). (b) Compared to the time-conservation model, cooling by 2K yields greater reduction of repeat length ($p=1.3\times10$, one-tailed t-test).

We found HVC temperature can also affect the transition probabilities of branching points. To assess the effects of temperature on that, we compare the probabilities of the first-order

syllable transitions in the control condition and when HVC is cooled by 4K. Fig. 2-7 a shows an example of a branching point which is affected by HVC cooling. Syllable ‘k’ in the song can be followed by either syllable ‘b’ or syllable ‘d’. The transition 'kb' is dominant under control condition, but, when HVC is cooled, the transitions 'kb' and 'kd' both have probabilities closer to 0.5 (Fig. 2-7 a), the change of probability is statistically significant ($p = 1.8\times10\text{-}5$, Fisher's exact test).

Out of 23 branching points in 6 birds, 11 are significantly altered by cooling HVC by 4K ($p<0.05$, two-sided Fisher's exact test). To rule out the inteference of the day-to-day variation of the transition probabilities, Cochran–Mantel–Haenszel (CMH) test was used to examine if the change of transition probability distribution is consistent across days. All those 7 branching points chosen show consistent distribution change on different days. The natural logarithm of the odds ratios (ln(OR)) between the control and cooled songs for choosing one branch over others from 3 different days are used to visualize the consistency of probability change (Fig. 2-7 c). They all have the ln(OR) of each branching point either above or below 0 in all the 3 days.

We further explored if the change of the transition probability distributions is related to the change of sequence variability. For each branching point, we calculated the transition entropy (see Methods for definition) to see how random or deterministic the transition is. The transition entropy can quantify the transition variability in such a way that the more deterministic a transition is, the lower the transition entropy will be and vise versa. We observed an elevation of transition entropy in 9 out of 23 branching points (tested by constructing the 95% confidence interval of the entropy difference using bootstrap method), a decrease in only one branching point and all the rest are not affected. The 9 branching points that show increase of transition entropy belong to the branching points that are changed in transition probability distribution by temperature. Therefore, most branching points that are affected by temperature are due to the increase of randomness of the transitions, i.e., the decrease of transition probabilities of the

frequent ones and the increase of transition probabilities of the rare ones (Fig. 2-7 c). The direction in which the transition entropy changes is also consistent across various temperatures (Fig. 2-7 d). In 7 out of 18 branching points, the transition entropy decreases as temperature increases (Observed in 5 birds. We did not get enough data for the 6th bird). In conclusion, about half of the branching points are affected by temperature, among which, most are driven in such a way that higher temperature causes more deterministic transitions while lower temperature causes more random transitions.

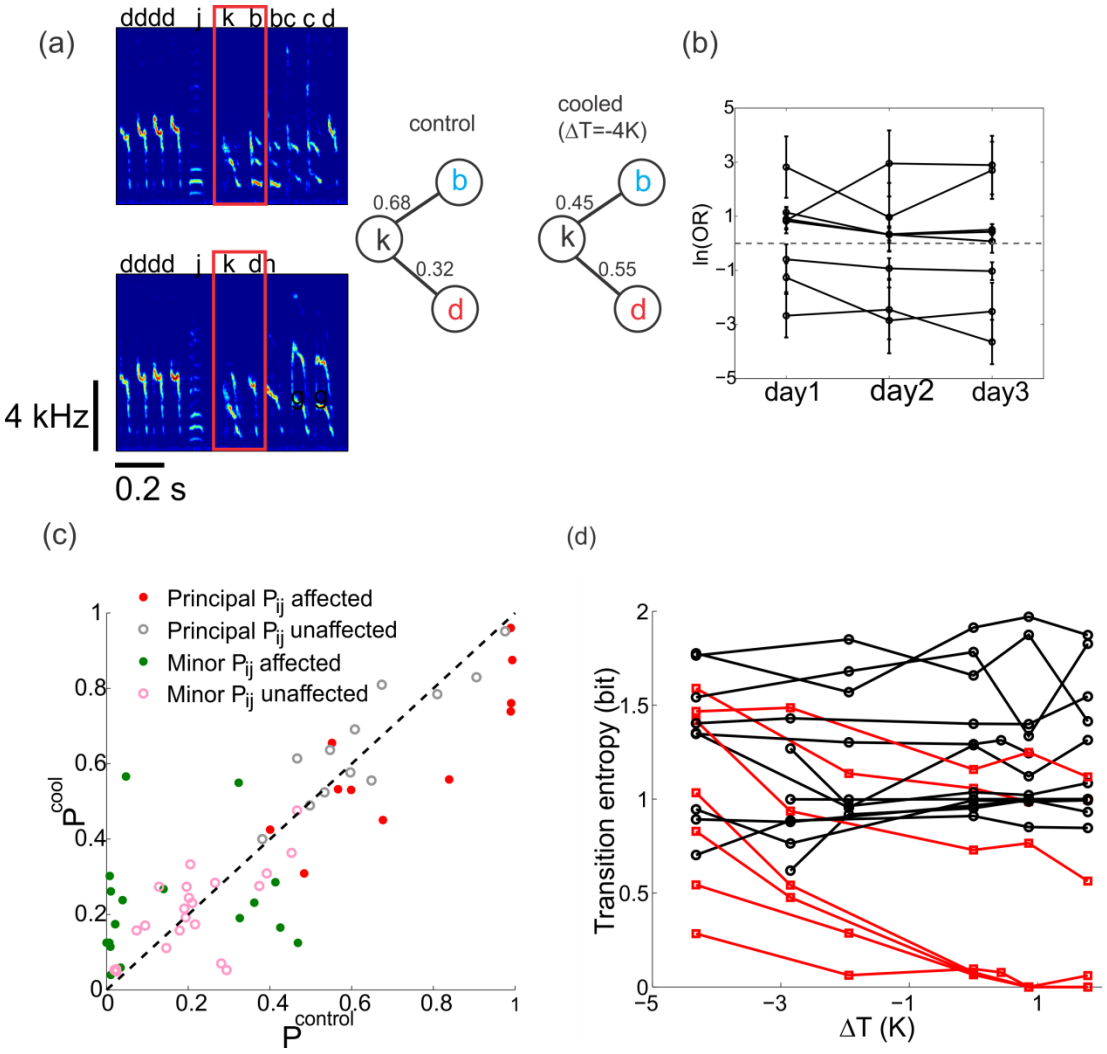


Figure 2-7. Effects of HVC temperature on branching points

(a) Examples of a branching point affected by temperature. Top: spectrograms of song segments showing that syllable j can be followed by either syllable a or k; bottom: the transition probabilities to a and k from j under control and 4K cooling. Right: transition probabilities at different temperature. (b) The day-to-day variation of the transition odds ratio (OR). 8 of the 11 branching points that are significantly altered by temperature are shown in this graph (the 3 that are not shown have 0 count in some of the days). Error bars are the standard error of ln(OR). All the 8 branching points have consistent odds ratios going to the same direction ($\ln(OR)>0$ or $\ln(OR)<0$) in all the three days that are examined. (c) Transition probabilities in cooled vs control songs. Solid dots are the transition probabilities belonging to the branching points that are significantly affected, of which most principal probabilities (red dots) are below the 45° line, while most minor probabilities (green dots) are above the 45° line. (d) Transition entropies of 18 branching points from 5 birds at different temperature. The red curves (n=7) have significant decreasing trend ($p<0.05$, one-tailed t-test on slopes) with temperature.

2.3.3 Changing temperature in RA

The dynamics of RA during singing has been found to be driven by its upstream nucleus HVC.[46] Whether it would reshape the HVC activity via the reciprocal connection[70] or the feedback loop to affect the syntax of the song is not clear. We further argue that the HVC temperature effect on the syntax is not a result of neural activity change in RA. In order to address that, we bilaterally change the temperature of RA directly.

To demonstrate that the RA activity is actually altered by temperature, the spike rate of the spontaneous spiking activity was measured with different RA temperatures in an anaesthetized bird. The spikes become more sparse as temperature decreases (Fig. 2-8 a, b). The

change of spike rate as a function of temperature is 0.45 ± 0.02 Hz/K. The spike rate is increased from the lowest to the highest temperature by 1.6 times. The implantation was also confirmed by histology after the experiment was done (Fig. 2-8 c).

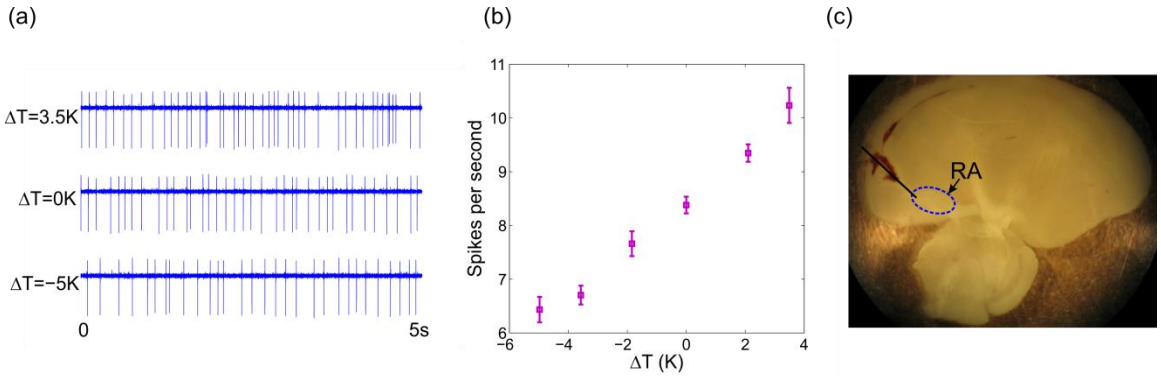


Figure 2-8. Cooling RA

(a) RA spiking in an anaesthetized bird at three different temperatures. (b) Change of spikes per second with temperature. Error bars are the standard deviation of spike rate. (c) Histology for RA cooling implantation. The black bar indicates where the cooling probe was inserted.

Neither the song timing nor the syntax, however, is found to be affected by temperature change in RA during singing (Fig. 2-9 a). The syllables are stretched by 0.35+\-0.20%/K (n=13) and the gaps are stretched by 0.23+\-0.27%/K (n=16). Both the syllable and gap stretch are significantly greater than 0 (stretch of syllables: $p=2\times10^{-5}$, one-tailed t-test; stretch of gaps: p=0.01, one-tailed Wilcoxon signed rank test), but significantly smaller than cooling in HVC (stretch of syllables: $p=3\times10^{-7}$, one-tailed t-test; stretch of gaps: $p=8\times10^{-9}$, Mann-Whitney U-test). The center of RA to the center of HVC has a distance for about 2.5mm. If the decay of temperature change follows $e^{-d/\lambda}$, where $\lambda \approx 1.2$mm[46], HVC should change about 12% as the temperature change in RA, which is consistent with our observation of about 1/10 stretching

effect as cooling HVC. Therefore, the change of song timing by cooling RA accounts for the slight temperature change in HVC and RA temperature does not affect song timing.

Furthermore, the distribution of the repeat length stays the same at different temperatures. By examining a long repeat that has great variability in repeat length, the variation between groups of different RA temperatures is equal to the variation within the group of the same temperature (p=0.9, Kruskal Wallis-test). The slopes of mean repeat length vs temperature on each day do not show a consistent trend either (Fig. 2-9 b). Therefore, changing temperature in RA does not have an effect on the repeat length as seen in the HVC temperature change experiments.

RA temperature does not affect the transition probability distribution at the branching points as well. For all the 5 branching points we investigated, the transition frequencies are not found to be dependent on temperature (p>0.05, Pearson's chi-squared test for independence). In addition, the transition entropies of the branching points are not changed by temperature either (Fig.2-9 c).

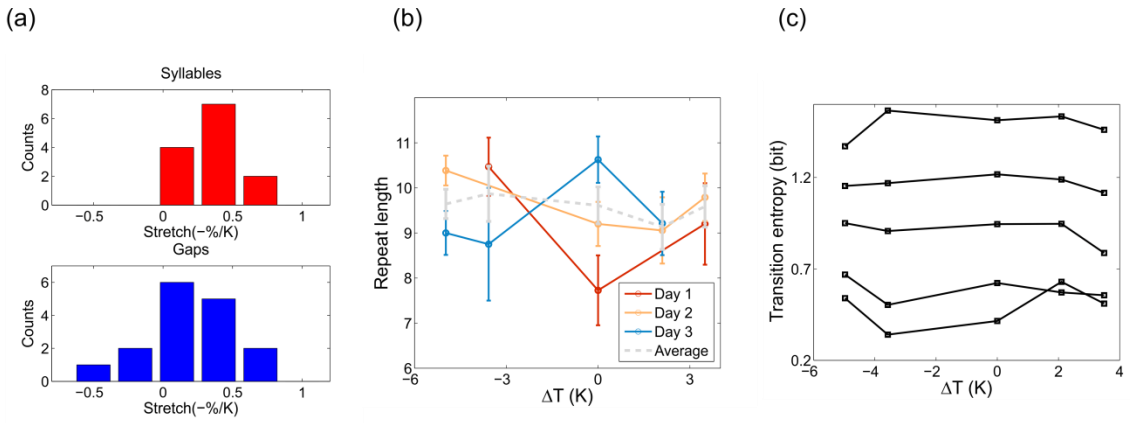


Figure 2-9. RA cooling effects

(a) Distribution of fractional syllable stretch and gap stretch. (b) Mean repeat length with different RA temperature on 3 days. None of them has a slope that is significantly different from 0 (p>0.05, t-test on the slopes). (c) Transition entropies of 5 branching points at different

temperatures. the five curves all have slopes that are not different from 0 ($p>0.05$, t-test on the slopes).

2.4 Discussion

Our experiments demonstrate the first real-time reversible brain circuitry manipulation altering the syntax of the action sequence. Cooling effects on the song syntax show that HVC is critically involved in syntax generation. This is the first direct demonstration of the significance of neural activity within HVC in the syntax formation. Brain area HVC has attracted a lot of attention as a site that plays a key role in the song generation, particularly in the song tempo control. The relation between the "fine-scale" song timing that determines the duration of song syllables and the higher level temporal control that defines the sequencing of syllables in the song remains unclear. The involvement of neural activity within HVC in the syllable sequencing of the song was not established until this study. Our findings cast serious doubts on the models of purely hierarchical song organization, where HVC is determining the song tempo, but external command input plays the dominant role in the song syntax. Involvement of HVC in song syntax shows that there is overlap in the neural circuits and significant interdependence of the neural mechanisms underlying the temporal patterning on a fine scale (timing within the song syllable) and the temporal organization of the song on a larger timescale (organization of syllables into a sequence).

The strongest temperature effect is on the change of repeat length. Syllable repetition is a universal phenomenon in Bengalese finch songs. Studies on the effect of social context on repeats show an increase of repeat length in the directed songs.[141] Therefore, a well learned Bengalese finch song should contain a considerable number of repeated syllables.

The distribution of a repeated syllable's length remains stable across days (Fig. 2-10). Our experiment links the encoding for repeat length to HVC, where the timing of the song is encoded. The computation in HVC has been well described by a chain model.[95,101] By allowing a probabilistic transition among multiple chains, variable sequences like Bengalese finch songs can be generated.[101] If the transition between syllables is simply Markovian, the distribution of the repeat length should follow a geometric decay, which is not what we see in the Bengalese finch songs.[71] As a result, the neural activity activated by the recurrent propagation should have been modified each repetition to generate such a bell-shaped distribution. Changing temperature in HVC can shift the distribution of the repeat length, indicating that temperature may have altered the modification to each repetition cycle. Fujimoto et al observed neural activity in some of the HVC(X) neurons that either increases or decreases in spike rates through a repeat. These neurons seem to be acting as counters for the repeats. Temperature may have changed the rate of increase or decrease in those HVC(X) neurons. Two mechanisms might be able to produce the counter-like behavior: the HVC neurons, such as those HVC(X) neurons, might be able to be self-excited to produce a decaying or cumulative activity; or the neurons could be excited by external inputs, such as the auditory feedback input and the feedback input from Uva, with which, even though the neurons' own activity may be always decaying, they are kicked to produce either a slower decay or accumulation. Changing temperature in HVC cannot distinguish those two possibilities, because it can either change the neuronal properties for self-excitation or the synaptic properties to receive the external inputs. In the next chapter, we will discuss a situation when the auditory feedback is distorted, how the repeat can be affected, which sheds some light on explaining the role of external input in producing repeats.

Our results on the change of repeat length does not support the idea that there is "planning" of the temporal duration of the syllable repeat: the syllable is repeated during the "planned" temporal interval. If the "planning" hypothesis is true, some area should encode for a

longer time scale that holds for 1-2s.The observation that repeat duration is dependent on HVC temperature, and possibly the subsequent state of dynamics or chemical properties of HVC neurons, casts doubt on the existence of such areas or neural mechanisms. It is more likely that how long the repeat is sung, is a dynamic decision that is affected by the change of neural activity each cycle. Thus our result supports a Markov chain model with adaptation over a hierarchical model.

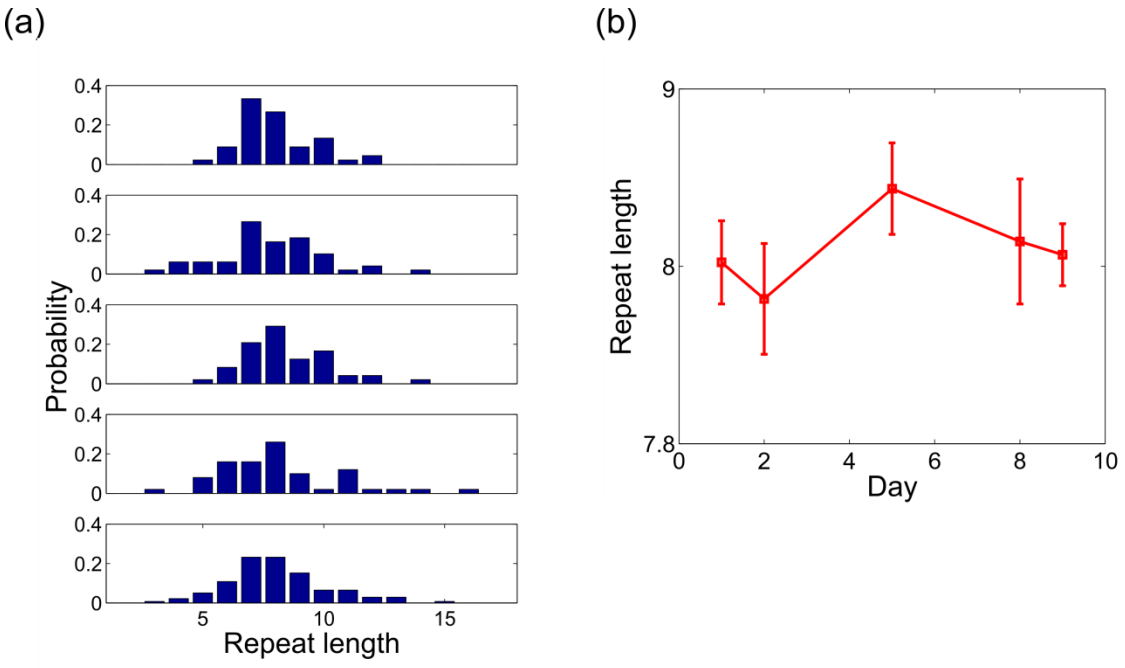


Figure 2-10. Repeat length across days

(a) Distribution of repeat length on 5 different days. (b) The mean of the repeat length in a period over 9 days. The same data are used as (a). Error bars are the s.e.m.. There is no difference among the repeat lengths (p=0.67, one-way ANOVA test).

We saw a change of repeat length in only the longer repeats, but not the short ones. One may argue that the short repeat has not much room to reduce in length further, however, warming

HVC did not drive them to increase either. The short repeats do not have as much variability in length naturally. These repeats are more like a deterministic sequence, only that the syllables in this sequence have the same acoustic profile. The short repeats may be explained by a many-to-one mapping that same syllables can be encoded by different neurons. In a neural network, these neurons may not be connected in a recurrent fashion but rather a feedforward fashion as those in the chain model. A desirable experiment to carry out would be to record HVC neuron activity during singing and see how the neurons act at the short repeats and long repeats. If the hypothesis is correct, it would also be interesting to find out how the long repeats are formed through learning - is it a degeneracy of states or a primitive state before differentiating to multiple states?

When HVC temperature is changed, inter-syllable gaps stretch more than song syllables for Bengalese finches, and the difference between the stretch of the syllables and gaps is larger than the difference reported for zebra finch.[122] This difference might be related to the differences between the syntax of the zebra finch song and the Bengalese finch song and their underlying neural mechanisms. The gaps in a Bengalese finch song are more variable in duration compared to the syllables. In our model, the inter-competition among the candidate branching chains occurs during the gaps. This competition process could cause greater variability in the duration of gaps and may be less robust against the temperature change, resulting in a greater stretch in the gaps compared to song syllables.

The finding that reversible manipulations of HVC circuitry alter the syntax of the song may be consistent with several alternative syntax mechanisms, including the neural feedback-mediated syntax[78] and co-articulation phenomenon[144]. Given the complexity of the neural circuitry and the level of current understanding of biological circuits, further experiments will be needed to decisively establish the neural mechanism that is consistent with all experimental evidence. Additional questions about birdsong syntax remain, such as the interaction between the motor circuits in the two hemispheres. It has been shown in zebra finch that the coordination of

the two hemispheres plays a role in tempo generation[46,125]; however, the role of two hemispheres in syntax control remains unknown.

To characterize the syntax changes in our experiments, we analyzed the changes in the syllable transition probabilities. Although the changes in transition probabilities unambiguously indicate the syntax change, transition probabilities alone do not fully capture the statistical properties of the syllable sequences in the Bengalese finch.[34,71,147] It will be interesting to see whether more subtle changes in the syntax can be detected with statistically correct representations of the song syntax, such as the partially observable Markov model with adaptation[34] and the long-range order of the song sequence.[32]

2.5 Conclusions

In summary, our experimental findings demonstrate the close inter-dependence of the syntax and timing neural mechanisms and suggest that brain area HVC is the key site for the birdsong syntax generation. This result restricts the plausible birdsong syntax models to the models where the tempo and the syntax mechanisms are inter-dependent, and HVC is strongly involved in both. This finding will motivate further theoretical and experimental studies elucidating the neural mechanisms of temporal order.

Chapter 3

The effect of distorted auditory feedback on Bengalese finch song syntax

Auditory feedback has been suggested to affect the birdsong syntax.[86] The HVC neurons have been found to respond to the altered auditory feedback when a Bengalese finch is singing[87], suggesting that the auditory feedback effect on syntax may be mediated via HVC. We use distorted auditory feedback (DAF) to test how the repeats and branching points are affected by auditory feedback. This result, in combination with the cooling effects on syntax, can help us understand how HVC may have functioned to control syntax. We find that DAF reduces repeat length in a similar way as cooling. The DAF effects on branching points are also investigated.

3.1 Introduction

Auditory experience is necessary for young birds to learn songs.[81] Deafened young birds cannot develop normal songs.[38,80,148] Even when the songs are fully learned and stay stable, some birds still rely on auditory feedback to maintain the quality of their singing performance. The dependence on auditory feedback varies across species. Some close-ended learners, such as white-crowned sparrows and chaffinches produce normal songs even 2 years after removal of audition.[16] Deaf zebra finch adults have gradual song deterioration within 6-8 weeks[82,149], while deaf Bengalese finch adults develop abnormal songs in only 5 days.[83–85] In Bengalese finch, auditory deprivation causes degradation of sequence stereotypy within one week post deafening.[84] The deterministic sequences are broken up and novel transitions emerge.

Birds such as canaries and Bengalese finches that sing variable songs depend on real-time auditory feedback.[86] The transition probabilities at the branching points can be affected by altered auditory feedback.[86] In zebra finch, nuclei within premotor area and anterior forebrain

pathway (AFP) have not been found to be responsive to the bird's own song during singing[68,150–153], even though HVC is selective to the bird's own song when listening.[154–156] In contrast, multiunit recording in the HVC of a singing Bengalese finch shows response to altered auditory feedback.[157] This finding is consistent with the on-line reliance on auditory feedback for singing in Bengalese finch.

As demonstrated in Chapter 2, HVC is actively involved in syntax generation. In this chapter, it is tested if HVC determines the sequence by taking the auditory feedback into account. To check this, the auditory feedback is continuously disrupted by presenting white noise pulses[68,151,158] during the repeats. A reduction of repeat length was observed with the distorted auditory feedback (DAF), which is similar to the cooling effect on the repeats.

3.2 Methods

3.2.1 Distorted auditory feedback

Birds were housed in a sound insulated chamber, where song recording took place. A targeted syllable was chosen, and the first 15-20ms of that syllable were used as the template in the algorithm for on-line syllable detection. The real-time auditory feedback system is described in Skocik et al (2013).[159] Once the syllable was detected, the system generated a white noise pulse, 30-60ms long, and 70-80dB SPL, with a systematic delay for about 30μs. This intensity approximates or is a little louder than the bird's natural voice.

The probability of DAF was set to affect 30%-50% of the songs. The random generation of white noise pulses was all-or-none: either generating the noise following the detection of all the targeted syllables in a song or generating nothing. In this way, if the targeted syllable was a repeated one, the bird would hear the noise throughout the repeat (Fig. 3-1).

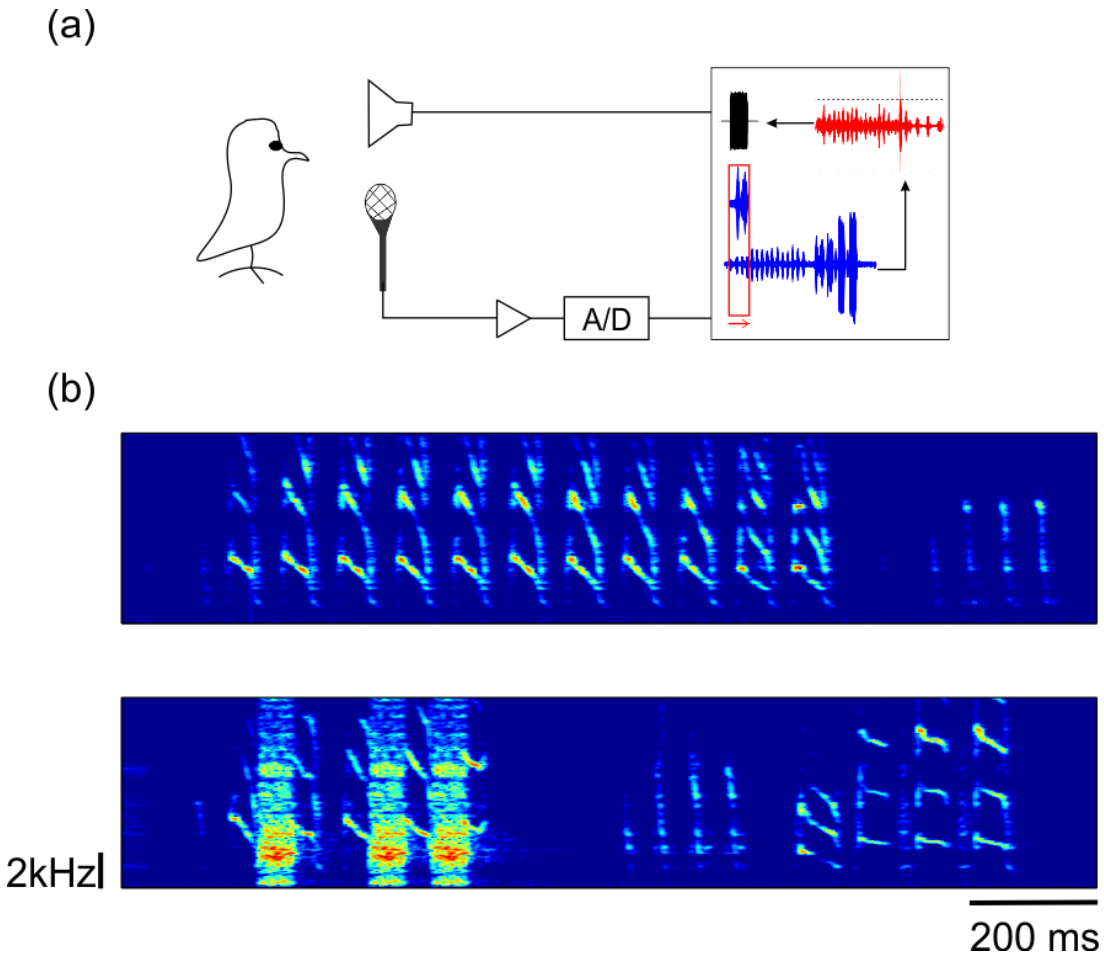


Figure 3-1. Distorted auditory feedback

(a) Diagram for the real-time auditory feedback system. The acoustic signal is picked by the microphone, amplified and digitized by an A/D convertor and recorded by the customized RTXI software at a sample rate of 30303Hz. The signal stored in the buffer is matched with a pre-stored template of the targeted syllable using cross correlation. Once the cross correlation crosses a threshold, a brief white noise is generated. (b) Top: a segment of intact song with a long repeat. Bottom: Disturbed song from the same bird.

3.2.2 Data analysis

DAF on repeats

The number of repeats with distorted playback was counted as follows. First, for the intact bird songs, an average syllable duration was estimated by dividing the total interval of the repeats by the total number of repeats. Then this average was used to divide the total interval of the repeats with playback to get the estimated number of repeats in the disturbed songs. The effect of DAF on repeats was assessed by comparing the repeat length from the intact songs with the DAF songs. The distributions of the repeat lengths were compared by Mann-Whitney U test. Unless otherwise indicated, the significance level was set at α=0.05 criterion.

Comparison to cooling results

The distribution of the repeats with DAF was compared to the distribution with cooling. Kolmogorov-Smirnov test (KS-test) was used to compare the DAF distribution to the cooling distribution of various temperatures.

DAF on branching points

The distribution of the transition probabilities at the branching points between the DAF songs and intact songs on the same days was compared using Pearson's χ^2-test. To rule out the possibility of probability change in the interleaved intact songs due to DAF, these songs were also compared to the songs from the days before DAF was introduced. We did not take the song interruption into account. It was concluded that an interruption occurred if the inter-syllable-gap following the white noise was longer than 0.3s or if it was the end of the song.

3.3 Results

3.3.1 DAF effect on repeats

With DAF presented, two birds show significantly reduced length of the long repeats (Fig. 3-2 $p_{bird7} = 2.5 \times 10^{-11}$, $p_{bird11} = 9.7 \times 10^{-7}$, Mann-Whitney U-test). While presenting DAF during a short repeat does not change repeat length significantly (Fig. 3-2 p=0.99, two-tailed Mann-Whitney U-test). It is consistent with what we saw in the cooling experiment: longer repeats are more affected by temperature, while short repeats are not changed by temperature. This result suggests that the generation of short repeats has less reliance on the auditory feedback than the generation of long repeats. If the change of auditory feedback is the only factor that causes the reduction of repeats, we can infer that cooling HVC may have disturbed the integration of the auditory feedback signal during the stage of repeat production.

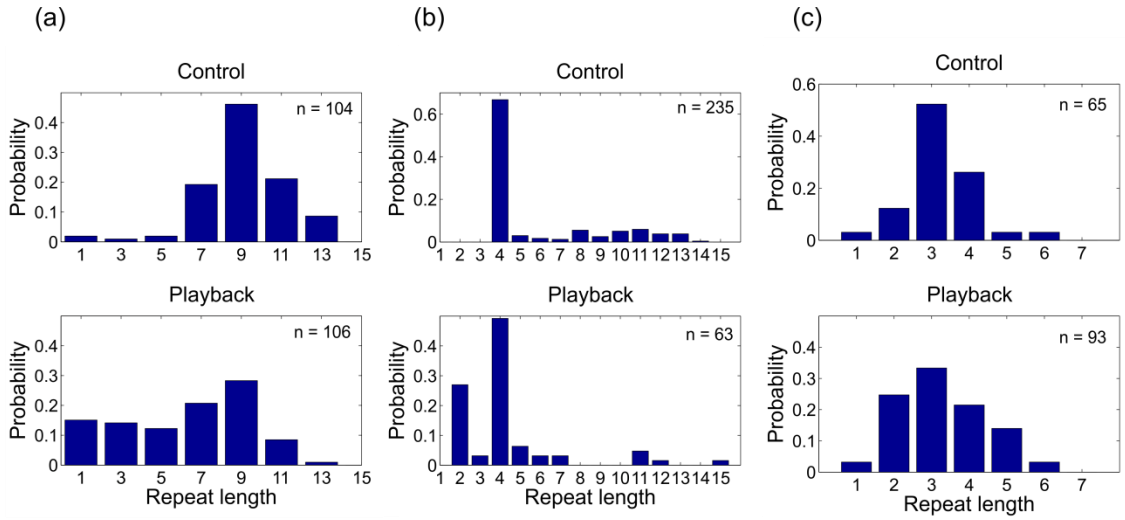


Figure 3-2. DAF effect on repeat length.

(a) Bird 7, syllable #7. Repeat length distribution without (top) and with (bottom) DAF. DAF significantly reduces repeat length ($p=2.5\times10^{-11}$, Mann-Whitney U-test). (b) Bird 11,

syllable # 6. Repeat length is reduced by DAF. Without DAF, all repeat lengths are above 3 while with DAF, repeat length less than 4 emerges. (c) Bird7, syllable #2. DAF effect on a short repeat with the mean around 3. DAF does not change the mean repeat length (p=0.99, two-tailed Mann-Whitney U-test).

If the reduction of repeat length by cooling and by DAF are attributed to different mechanisms, the repeat length distribution in the two cases may have different distribution shape. Both cooling and DAF have been done in Bird 7. Syllable #7 is also reduced in length by cooling HVC. As shown in Fig. 3-3, both DAF and cooling (by 2K) shift the distribution to the left. To find at what temperature the cooling effect resembles the DAF effect, p-values of the Kolmogorov-Smirnov test (KS-test) between the two distributions at different temperatures was used. The probability for the two distributions to be the same has a maximum at $\Delta T = -2K$ (Fig. 3-3 c, d, p=0.07). Therefore, altering auditory feedback with our paradigm seems to affect the repeat length similarly to a mild temperature drop in HVC.

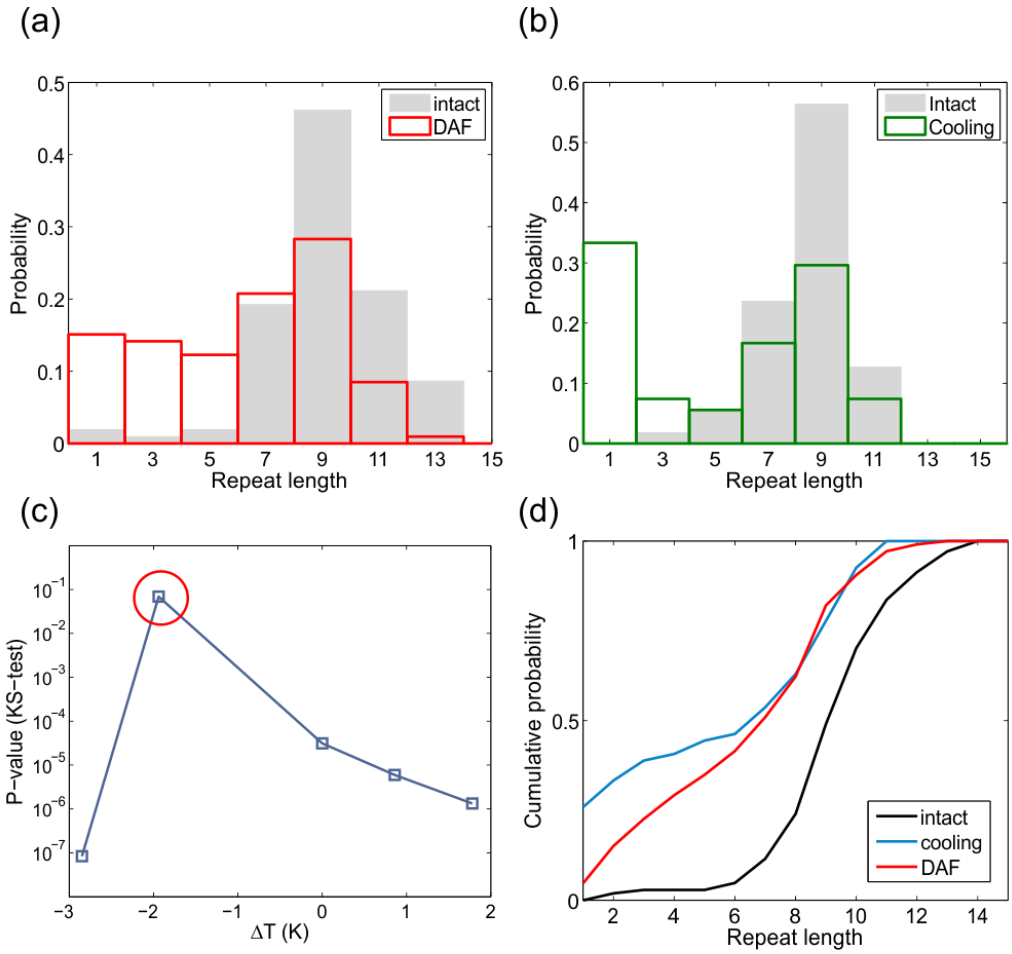


Figure 3-3. Comparison between the effect of DAF and cooling.

(a) Distribution of the repeat length of syllable #7 from Bird 7 with and without DAF presented. $n_{intact} = 104$, $n_{DAF} = 106$. (b) Distribution of the same syllable's repeat length with and without cooling by -2K. $n_{intact} = 55$, $n_{-2K} = 54$. (c) The p-values of the KS-test between repeat length with DAF and with different temperature changes. At dT = -2K, the similarity between cooling and DAF is largest, p=0.07. (d) Cumulative probability of the repeat length distribution from intact, -2K cooling and DAF songs.

3.3.2 DAF effect on branching points

DAF also affects the transition probability at the branching points. We tested 3 branching points in 2 birds. All of them are significantly affected by DAF ($p<0.05$, Pearson's χ^2-test, Table 3-1, Fig. 3-3). DAF seems to increase the probabilities of the transitions that occurred rarely or never occurred in the intact songs (Table 3-1). If we look carefully at those transitions, as shown in Figure 3-4, there are two novel transitions 'cf' and 'ci'. Syllable 'i' is an introductory note and syllable 'f' is always the first syllable followed by a series of repeated syllables. In intact songs, 'i' and 'f' follow a deterministic sequence 'cdee'. With DAF presented, sometimes the 'dee' after the 'c' is skipped and the song resumes from the beginning of the next chunk. The transition 'ei' and 'bi' in Bird 11 are also transitions to introductory notes. As we see in Table 3-1, with DAF, all those kind of transition probabilities are enhanced. We want to be careful here because song restart might be different from the probabilistic transitions at the branching points. If we disregard the song restart and only consider the branches of the normal transitions, 'bc' in Bird 11 becomes a deterministic transition with both control and DAF, while the other two branching points show no difference between control and DAF ($p>0.05$, Pearson's χ^2-test). Therefore, we think DAF changed transition probabilities mostly by increasing the frequency of song restart.

Again, we compare the DAF effect to the cooling effect on the branching points. As shown in the last block of the table, cooling HVC ($\Delta T = -4K$) not only increases the number of transitions to 'f' and 'i', but also affects the transition probabilities of 'cc' and 'cd' ($p = 2.7 \times 10^{-7}$, Pearson's χ^2-test). It suggests that cooling not only caused song restart but more importantly, it affects the normal transitions. We did not see such kind of effect in the DAF experiment. But since our data are very limited, more experiments need to be done to draw a convincing conclusion.

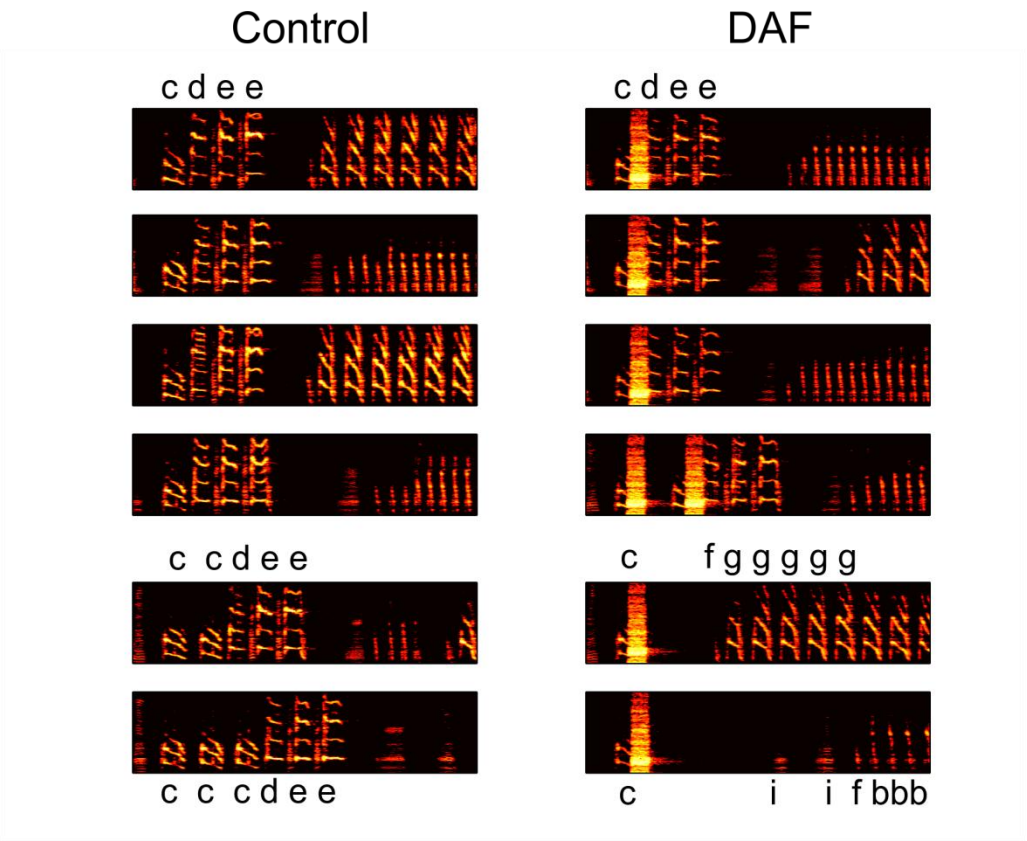


Figure 3-4. Effect of DAF targeting at syllable 'c'.

Table 3-1. Transition probabilities at the branching points without and with DAF.

B11	P_{ef}	P_{ek}	P_{ei}		N_{total}
Control	0.77	0.13	0.10		192
DAF	0.61	0.11	0.28		44
B11	P_{bc}		P_{bi}		N_{total}
Control	0.95		0.05		65
DAF	0.71		0.29		97
B7	P_{cc}	P_{cd}	P_{cf}	P_{ci}	N_{total}
Control	0.26	0.74	0	0	183
DAF	0.18	0.66	0.08	0.08	203
B7	P_{cc}	P_{cd}	P_{cf}	P_{ci}	N_{total}
Control	0.05	0.94	0.01	0	264
Cooling	0.19	0.69	0.07	0.05	169

3.4 Discussion

The goal of this chapter is to investigate the distorted auditory feedback's effect on the syntax of Bengalese finch song so as to address the question of whether the syntax change we see in the cooling experiment is due to the internal change of HVC dynamics or the change in receiving the external drive.

DAF has shown to shorten the repeats which is similar to the effect of cooling HVC. It suggests that cooling HVC has affected the strength of auditory feedback input during singing. Sakata and Brainard observed a decrease of HVC interneuron activity with altered auditory feedback[157] and they also observed syntax change[86]. All these observations suggest that auditory input to HVC is an important factor in shaping the syntax.

DAF affects branching points mostly by causing the song to stop or restart. Cooling also more frequently truncates the song or inserts introductory notes in between syllables. But cooling also influences the transition probabilities of the normal transitions. We do not observe the probability change in the normal transitions by DAF. But given such a small dataset, it is hard to make a comparison to cooling. Further experiments are needed to answer this question.

3.5 Conclusions

In summary, our results suggest that 1) DAF reduces repeat length similar to cooling; 2) DAF affects the branching points, but the effects can be viewed as song stopping or restarting. The similar effects on the syllable repeats by DAF and cooling HVC suggest that auditory feedback may have shaped the syntax via HVC. Further studies will be needed to establish whether the effects of DAF on branching points are similar to effects caused by HVC cooling.

Chapter 4

Using unilateral cooling to study the two HVC hemispheres in singing control

Activities of HVCs in the two hemispheres are needed to be tightly coordinated to produce complex songs. We use unilateral cooling to differentiate the role of the two HVCs in controlling song timing and syntax. We find that the timing control of syllables is dominated by the left hemisphere, while the timing control of inter-syllable gaps is alternating between the two hemispheres. Most syllable repeats are controlled by one hemisphere. The transition probabilities at the branching points seem to be influenced by both sides.

4.1 Introduction

Songbirds have a complete set of motor pathway in each hemisphere that controls the ipsilateral vocal organ (syrinx), respectively. Each side of the syrinx can be controlled independently, but the two sides jointly produce the complex sounds.[160–162] Therefore, there exists a question about how the two hemispheres are coordinated while conducting such complex motor activities. Direct connection between the premotor areas of the two hemispheres is absent, though synchronization via the midbrain during singing has been observed.[163,65] It is conceivable that motor activities of certain complexity should not be simultaneously controlled by both hemispheres, since it will require a great deal of inter-hemispheric communication which costs time. As a result, the complexity of the motor pattern may break the symmetry of the bilateral control and cause lateralization.[164]

In humans, the production and perception of language is lateralized in the left hemisphere.[165] It has also been shown that non-human vertebrates also exhibit lateralization for vocal control, such as songbirds. Canary songs are primarily produced by the left HVC, according

to the unilateral HVC lesion studies.[166,167] The two hemispheres of a zebra finch rapidly switch their dominance in singing control.[46,125] The Bengalese finches are found to be left-side dominant in both singing and perception of the songs.[168,169] Peripheral studies in Bengalese finch motor neurons and syrinx show specialized control of frequency by the two sides of the syrinx.[126] These observations suggest that left and right HVCs are functionally differentiated.

Most studies mentioned above were done by lesioning HVC in one of the hemispheres. However, this method is irreversible and cannot compare the effect of unilateral lesion in the same bird. Temperature manipulation provides a convenient way to compare the effect on the control of the two sides. In the previous chapters, we have argued that HVC plays a key role in encoding the syntax of the song. In this chapter, we further study the role of HVC in each hemisphere using unilateral cooling.

4.2 Methods

4.2.1 Unilateral cooling

Two Peltier modules (TE-18-0.45-1.3, TE Technology) were used to independently change temperature of one side of the brain. The two modules were connected in series with the polarity pointing to the same direction, i.e., one's positive side was connected to the other's negative side. 4 jumpers were used to determine which side or both sides would be used (Fig. 4-1 a). Both Peltier modules shared a common heat sink. Two gold-plated silver pads, each of which had a size of 2mm×2mm, with a 2mm separation, were attached to the cold side of the Peltier modules by thermally conductive adhesive (3M Thermally conductive Adhesive TC-2810).

To measure the temperature change in HVC by unilateral cooling and compare it to the bilateral cooling, two thermocouples were inserted in HVC of each side at a depth of 0.7 mm

underneath the cooling pads in an anaesthetized bird. Voltage across each thermocouple was read on a voltmeter (HP 3457A multimeter). The readings were recorded 2 minutes after every time the current was changed so that the temperature change could be stabilized. Brain temperature change in case of unilateral cooling was similar to the change in case of bilateral cooling (Fig. 4-1 b). In case of unilateral cooling, there was a slight temperature change in the contralateral HVC due to the heat transferred from the shared heat sink. This change was negligible compared to the other side. Therefore, when current was passes through one Peltier module, it effectively changed the temperature of one side of brain without significantly affecting the other side.

During the experiment, two thermocouples were attached to the top surface of the left and right cooling pads respectively to monitor the temperature change. The signal was amplified 1000 times in order to be recorded via the analog to digital converter (NI USB-6211) along with the songs.

Four different conditions, cooling left HVC, cooling right HVC, bilateral cooling, and control, were interleaved. Data was collected for about 2 hours in each condition.

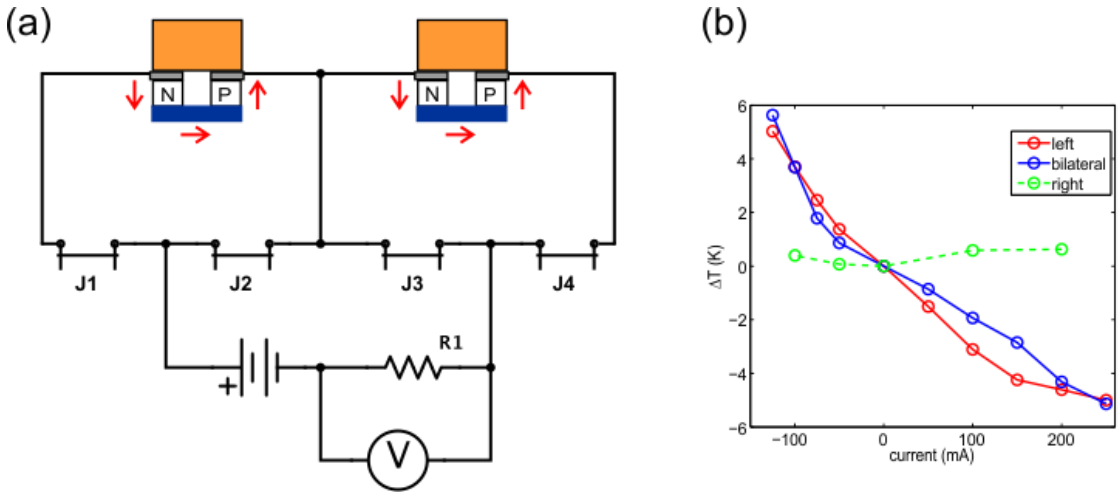


Figure 4-1. Unilateral cooling device

(a) Schematic of unilateral cooling device. The three operation modes: J1, J4 connected and J2, J3 disconnected; J1, J3 connected and J2, J4 disconnected; and J2,J4 connected and J1, J3 disconnected will conduct the current through both, left and right, respectively. (b) Temperature change with current in the left HVC measured at depth of 0.7mm below brain surface, under the operation with both (blue), left (red) and right (green) Peltier module.

4.2.2 Data analysis

Effects of unilateral cooling on song timing

We measured the syllable duration and the inter-syllable gap duration. The quantification of the cooling effect on song stretch, as the slope of the fractional change of duration vs. temperature, was the same as described in Chapter 2. Below, we denote the stretch (in %/K) of the i[th] syllable as ss_i^l and ss_i^r, and the stretch of the i[th] gap as sg_i^l and sg_i^r. The superscript "l" (or "r") denotes that left (or right) HVC is cooled. Wilcoxon signed rank tests were performed to compare the stretch effects between the left cooling and right cooling groups.

The correlation between the two unilateral cooling effects on syllable and gap stretch was calculated by Pearson correlation coefficient:

$r_s = \sum_{i=1}^{n}(ss_i^l - \mu s^l)(ss_i^r - \mu s^r) / \sqrt{\sum_{i=1}^{n}(ss_i^l - \mu s^l)^2 \sum_{i=1}^{n}(ss_i^r - \mu s^r)^2}$ for the syllable stretch, where μs^x is the mean of ss_i^x and

$r_g = \sum_{i=1}^{n}(sg_i^l - \mu g^l)(sg_i^r - \mu g^r) / \sqrt{\sum_{i=1}^{n}(sg_i^l - \mu g^l)^2 \sum_{i=1}^{n}(sg_i^r - \mu g^r)^2}$ for the gap stretch. The significance of whether the correlation coefficient was above 0 was tested using the one-tailed t-test with n-2 degrees of freedom with the t-statistic calculated by $t = r / \sqrt{\frac{(1-r)^2}{n-2}}$. The correlation was also confirmed by a shuffle test. One of the groups was shuffled by 1000 times,

each of which the correlation coefficient was calculated. The p-value was estimated by $p = N_{R<r}/1000$.

Syllable features

Five syllable features, syllable duration, mean fundamental frequency, mean spectral entropy, mean temporal entropy, and spectro-temporal entropy were extracted. The features were calculated for songs from the control group only.

The fundamental frequency (FF) was calculated from the middle part of the syllable that spans 1/2 of the total syllable duration. The FF was obtained by finding the first peak of the autocorrelation of the signal segment. For real signal, the autocorrelation is $R_{xx}(j) = \begin{cases} \sum_{i=0}^{N-j-1} x_{i+j}x_i, j \geq 0 \\ \sum_{i=0}^{N+j-1} x_i x_{i-j}, j < 0 \end{cases}$, which is symmetric around zero. For the right half of the autocorrelation, suppose the first peak is at the m[th] time lag, then the corresponding fundamental frequency is Fs/m, where Fs is the sample rate.

Let the spectrogram of a syllable be $S_{mn} = \tilde{x}(\omega_m, T_n)$. The spectral entropy is defined as $H_s = -\sum_m P_m log P_m$, where $P_m = \frac{\sum_n |S_{mn}|^2}{\sum_{m,n} |S_{mn}|^2}$. To calculate the temporal entropy, the absolute value of the raw signal was first smoothed. Then we calculated the normalized power from the smoothed signal as $P(t) = |x_{smooth}(t)|^2 / \sum_t |x_{smooth}(t)|^2$. The temporal entropy was calculated from $H_t = -\sum_t P(t)\, log P(t)$. The spectro-temporal entropy was calculated using $H_{st} = -\sum_{m,n} P_{mn} log P_{mn}$ with $P_{mn} = \frac{|S_{mn}|^2}{\sum_{m,n} |S_{mn}|^2}$. The cluster center of each syllable type, defined as the median of the points in the feature space, was used to represent the acoustic feature of that syllable type.

Relationship between syllable features and syllable stretch by unilateral cooling

We first normalized the measure of the syllable stretch to eliminate the systematic difference of cooling power in different birds. To do that, we divided ss_i^x by the median of

$\{ss_i^{bilateral}\}$ from the same bird. We made fit of the normalized stretch $\widetilde{ss}_i{}^x$ vs. each of the features of syllable i. To test the dependence of the stretch on that feature, the slope of the fit was compared to 0 using t-test.

In order to verify that this slope could really represents the relationship between $\widetilde{ss}_i{}^x$ and syllable feature and was not a result of bird-to-bird variation, we tested the null hypothesis that $\beta_1 = \beta_2 = \cdots = \beta_5 = \beta$ by ANOVA, where β_i is the slope of $\widetilde{ss}_i^x$ vs. feature of the ith bird. For the slopes from the left cooling data, the corresponding p-values for the five features are 0.935, 0.155, 0.163, 0.9995 and 0.1918. While for the slopes from the right cooling data, the p-values are 0.017, 0.086, 0.101, 0.031 and 0.164. Therefore at least for the left cooling, it was fair to pool the data from 5 birds together.

Stretch dependence on syntactic role

Depending on the syntactic role of syllables and gaps, they were categorized into 1) repeated syllables, 2) syllables in deterministic sequences, 3) branching points, 4) gaps between repeated syllables, 5) gaps between deterministic syllabls and 6) gaps at branching points. A syllable was considered a deterministic syllable if the transition probability following it was greater than 95%. If a syllable could be followed by more than one syllable types, it was considered a branching point. If the transition probability associated with an inter-syllable gap was greater than 95%, the gap was considered a deterministic gap; otherwise, it was considered a branching gap. Transitions with probability lower than 5% were not taken into account. Introductory notes were not counted.

The dependence of the stretch on the sequence categories was assessed by comparing the distribution of $\tilde{S}^x$ with the conditional distribution on its syntactic category using Kolgmonov-Smirnov test. If $\tilde{S}^x|category$ and $\tilde{S}^x$ had the same distribution (p>0.05, KS-test), the stretch was not considered to be dependent on the category of the syllable.

Analysis of hemispheric dominance on syllable repeats

The slope of the repeat length vs. temperature was used to quantify the temperature effects on the repeats. The significance of the effect was tested by t-test with N-2 degree of freedom.

After we looked at the slopes, we found 7 out of 8 repeats were primarily affected by changing temperature in only one of the hemispheres. To assess to what extent changing temperature in the dominant hemisphere can capture the bilateral effect, we first separated the repeat lengths with temperature change on the dominant side from the repeat lengths with temperature change on the non-dominant side. For the one repeat that was controlled by both hemispheres, we took the average repeat length of left and right, and assigned the mean to both dominant and non-dominant groups. Each data point $l_{i\Delta T}$ within one group was the mean repeat length of syllable i with ΔT temperature change divided by the mean repeat length of the same repeat in control condition. We mapped the $l_{i\Delta T}$ in each group to the $l_{i\Delta T}$ from the bilateral cooling data and calculated correlation coefficient between them. Then we compared the correlation coefficient to 1 using one-tailed t-test with the t-statistic as $t = (r-1)/\sqrt{\frac{(1-r)^2}{n-2}}$.

For the analysis of the relationship between the hemispheric dominance in controlling the repeats and other aspects, we quantified the dominance as $D_i^{repeat} = \frac{R_i^l - R_i^r}{R_i^l + R_i^r}$, where R_i^x is the slope of repeat length vs. temperature of syllable i.

We investigated the cooling effect on timing at different locations with respect to the repeats. Those locations were the gap before the repeat, gaps within the repeat, the gap right after the repeat and the repeated syllables. The difference of the stretch effect by cooling in the left and right at those locations was quantified by $D_i^{timing} = \frac{s_i^l - s_i^r}{s_i^{bi}}$. We also calculated the correlation coefficient between the repeat dominance and the difference of the stretch effect.

Comparison of cooling effect on branching points

The method to calculate the transition probabilities was the same as described in Chapter 2 for bilateral cooling. Changes in transition probabilities were analyzed using Pearson's χ^2-test. The randomness of the song sequence was assessed by the transition entropies at each branching point (see definition in Chapter 2).

To compare the transition entropies at all the branching points under different conditions - left cooled, right cooled and intact songs, Friedman's test was used. After getting the p-value, post-hoc test was performed based on a paired Wilcoxon signed rank test on the "control vs. left", "control vs. right" and "left vs. right" pairs. To estimate the dependence of transition entropy on temperature, the slopes of the transition entropy vs. temperature was calculated. The slopes was tested by t-test on whether it was greater than 0.

4.3 Results

4.3.1 Unilateral cooling effect on song timing

The cooling effect on syllables stretch is much more salient with the left HVC cooling than the right (Fig. 4-2 a, 37 syllables in 5 birds,$p=3.5\times10^{-7}$, Wilcoxon signed rank test). Stretch by the left cooling has an effect of -2.0+\-0.9%/K (mean +\- s.d., n=37 p=2e-15, t-test on mean smaller than 0), whilst stretch by the right cooling has an effect of -0.3+\-1.0%/K (mean +\- s.d., n=37, p=0.05, t-test on mean smaller than 0).

There is, however, no significant difference between the left and right cooling effect on the gaps (n=63, p=0.8830, Wilcoxon signed rank test). Therefore, both HVCs are responsible for the timing control of the gaps. The stretch of gaps by the left and right cooling are anti-correlated (r = -0.301, p=0.008, t-test with df=N-2), indicating that the control by the two HVCs on the gap timing is more alternating than simultaneous.

-

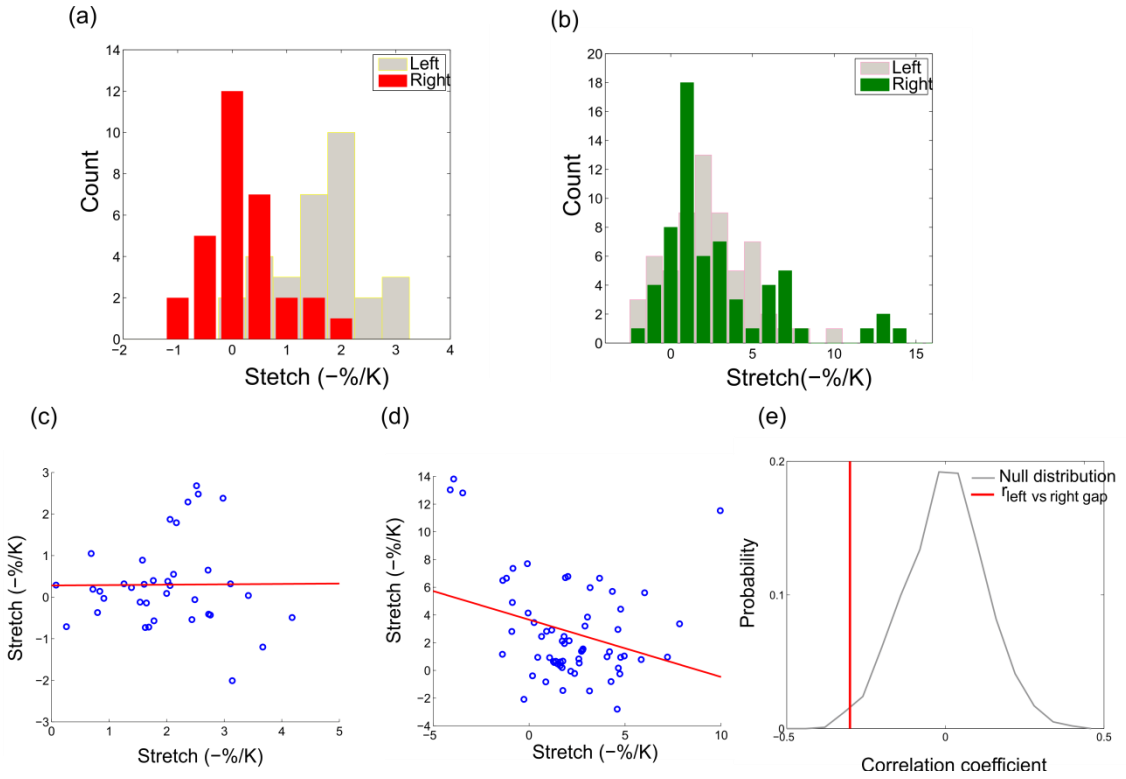


Figure 4-2. Unilateral cooling effect on timing

(a) Unilateral cooling effect on the stretch of syllables (n=37 in 5 birds). (b) Unilateral cooling effect on the stretch of inter-syllable gaps (n=63). (c) Stretch by left and right HVC cooling on the same syllables. X-axis: cooling left HVC; y-axis: cooling right HVC. r=0.0081, p=0.96,n=37, t-test on r=0. (d) Stretch by left and right cooling on the same inter-syllable gaps. X-axis: cooling left HVC; y-axis: cooling right HVC. r=-0.3005, p=0.017, n=63, t-test on r=0. (e) Comparison to the null distribution generated by shuffling the order of the gap stretch of one group for 1000 times, p=0.008.

Peripheral study on the Bengalese finch vocal organ show differential singing control by the left and right side of syrinx.[126] We want to see if the selective control on the syllable

structures has occurred in the higher order stage. Five features, syllable duration, mean fundamental frequency, mean spectral entropy, mean temporal entropy and the total spectro-temporal entropy have been chosen to test the relationship between the unilateral cooling effect and the acoustic structure of the syllables. It shows that, overall, cooling left HVC stretches more for the syllables that have relatively high fundamental frequency, relatively low spectral entropy and spectro-temporal entropy (Fig. 4-3). The duration of the syllables and the temporal entropy (the randomness of amplitude over time) are not correlated with how strongly the syllables are stretched. The right cooling effect on the syllables, most of which are quite minimal, does not display a significant correlation with any of the features.

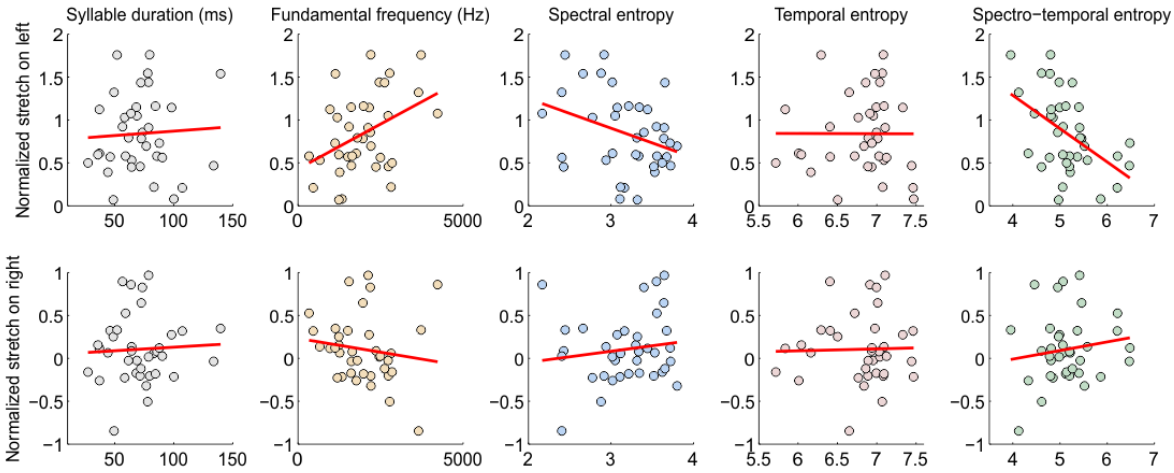


Figure 4-3. The relationship between syllable stretch and syllable features

The stretch by left HVC cooling (top chart) is dependent on the mean fundamental frequency of ($\beta^{FF,l} = (2 \pm 1) \times 10^{-4}, p = 0.023$, one-tailed t-test), spectral entropy ($\beta^{SE,l} = -0.34 \pm 0.17$, $p = 0.046$, one-tailed t-test) and spectro-temporal entropy ($\beta^{STE,l} = -0.39 \pm 0.11, p = 0.01$, one-tailed t-test) of the syllable but not the duration ($p = 0.48$) or the temporal entropy ($p = 0.49$). The stretch by right HVC cooling (bottom chart) is not dependent on any of the features ($p > 0.05$).

Next, we check if the stretch effect is specific to the syntactic role of the syllables. The syllables can be categorized into those within a deterministic sequence, repeated syllables and the syllables at the branching points. We compare the distribution of the stretch by unilateral cooling conditioned to the syntactic category of the syllables to the marginal distribution of the stretch effect. The unilateral cooling effect on the syllable stretch, however, is not dependent on the syntactic role of the syllables ($p>0.05$, Kolmogorov-Smirnov test). The transitions can also be categorized into deterministic transitions, transition between repeated syllables and probabilistic transitions at the branching points. The unilateral cooling effect on the gap stretch is not found to be dependent on the category of the transition as well ($p>0.05$, Kolmogorov-Smirnov test).

4.3.2 Unilateral cooling effect on repeats

From Chapter 1, we know that bilaterally cooling HVC reduces the repeat length. We want to know if that effect was due to the temperature change in one of the hemispheres or both hemispheres. For the 8 relatively long repeats that we investigated, 3 are significantly affected only by right cooling, 3 only by left cooling and 2 by both ($p<0.05$, t-test on the slope of repeat length vs dT greater than 0, Fig. 4-4 a). One of the two repeats (syllable #1) that are affected by both is actually much more affected by the right cooling than the left, so we also count it as right dominated. Therefore, for the control of most syllable repeats, only one of the hemispheres is dominant. The bilateral cooling effect is well captured by the dominant side. As shown in Figure 4-4b, the correlation between the repeat length with bilateral cooling and with the dominant hemisphere cooling is close to one ($r=0.88$, $n=40$, $p=0.066$, t-test on correlation coefficient smaller than 1); while the non-dominant side's correlation with the bilateral is significantly smaller than 1 ($r=0.59$, $n=40$, $p=0.0018$, t-test on correlation coefficient smaller than 1).

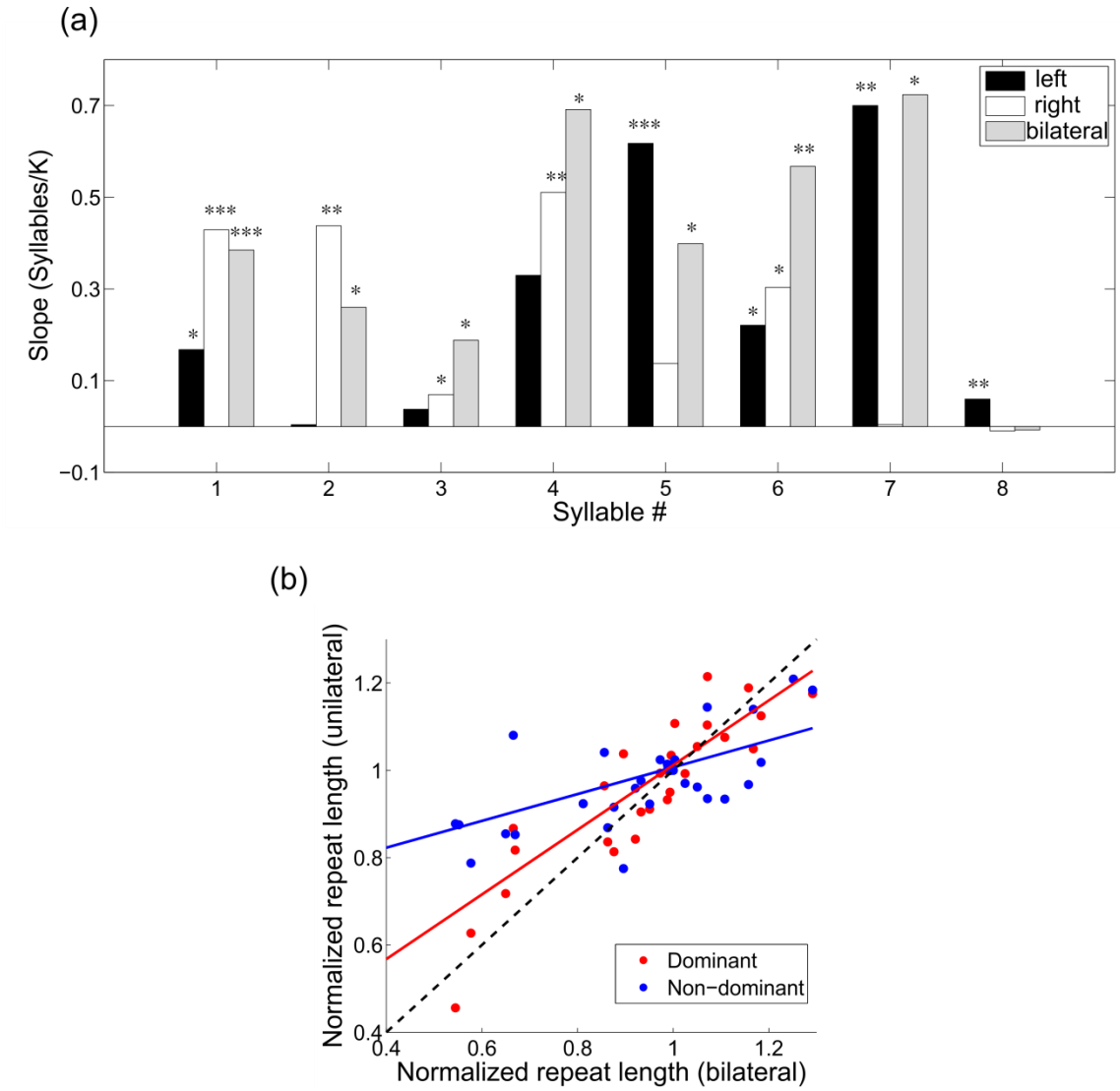


Figure 4-4. Lateralization of repeat control

(a) The unilateral cooling effect on the repeat length. Syllable # 2, 3 and 4 are dominated by the right hemisphere, syllable # 5,7 and 8 are dominated by the left hemisphere, while syllable #1 and 6 are bilaterally controlled. (b) Correlation between the bilateral cooling and unilateral cooling. The x-axis of each point is the mean repeat length at a certain temperature normalized by the mean repeat length of control; while the y-axis is the normalized repeat length of the dominant side at the same temperature. If both sides are equally contributed (syllable #6), the

average of the two hemispheres' is used. (b) The dependence of hemispheric dominance on syllable features.

We next relate the hemispheric dominance of the repeat control to the timing control. We categorize the gaps into three categories: the gap right before the repeat, the gaps within the repeat and the gap right after the repeat. We calculate the difference of the stretch between left and right cooling ($s^l - s^r$) and normalized it by the stretch effect with bilateral cooling. We find that the dominance of repeat control is positively correlated with the dominance of the gap stretch after the repeat (r=0.76, p=0.014, t-test on correlation coefficient) but not correlated with the syllable stretch (r=-0.19) or gap stretch before and within the repeat (r=0.55 and r=0.36, respectively, p>0.05, t-test on correlation coefficient, Fig. 4-5). This observation suggests that the control of repeat length might occur at the end of the repeat when the repeat-dominant hemisphere takes over.

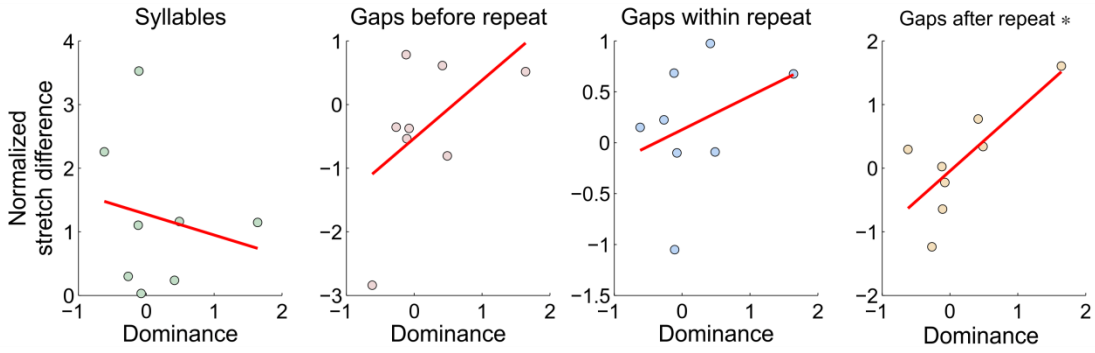


Figure 4-5. Correlation between hemispheric dominance on timing and repeats length.

x-axes are the hemispheric dominance of repeat length defined as $D_i^{repeat} = \frac{R_i^l - R_i^r}{R_i^l + R_i^r}$, where R_i^x is the slope of repeat length vs temperature of the ith repeat. y-axes are the normalized stretch difference of the syllables or gaps between left and right. The normalized stretch

difference is defined as $D_i = \frac{s_i^l - s_i^r}{s_i^{bi}}$. Therefore, if, for example, both repeat length and timing are dominant in the same hemisphere, we should expect a positive correlation and vise versa. Only the dominance of the gap stretch after the repeat has a positive correlation with the dominance of the repeat length (r=0.76).

4.3.3 Unilateral cooling effect on branching points

Previously we showed that bilateral cooling in HVC increases the sequence variability at some of the branching points. Here we want to see if the effect is due to only one of the hemispheres or both hemispheres.

Out of 20 branching points from 5 birds, 5 are significantly altered by left cooling ($\Delta T = -4K$) but not right cooling, 3 are altered by right cooling but not left cooling and 4 are affected by both ($p<0.05$, Pearson's χ^2-test or Fisher's exact test for rare transitions). There is no systematic dominance by one hemisphere, except that the control of branching points of Bird 2 seems to be dominated by the right hemisphere and the control of branching points in Bird 9 seems to be dominated by the right hemisphere. All other birds (Bird 3, 4 & 7) have effects on branching points in one or both hemispheres (Fig. 4-6 a). These data suggest that both hemispheres are involved in the transitions at the branching points.

To see in which direction unilateral cooling alters the transition probabilities, we calculated the transition entropies at the branching points for the songs cooled by either left or right HVC ($\Delta T = -4K$) and the intact songs. The difference among the three groups is marginal (p=0.07, n=20, Friedman's test), whereas there is a slight but significant increase of transition entropy if only comparing left to control and right to control separately (p=0.006, one-tailed Wilcoxon signed rank test on $\mu_{left} \leq \mu_{control}$; p=0.040, one-tailed Wilcoxon signed rank test on

$\mu_{right} \leq \mu_{control}$); and there is no difference between the effect of left and right cooling (p=0.13, two-tailed Wilcoxon signed test on $\mu_{left} = \mu_{right}$). Overall, unilateral cooling on either HVC slightly increases the transition entropies. If we take transition entropies under all temperature conditions into account, 4 branching points decrease in transition entropies with the temperature of left HVC and only one branching point has transition entropy decreasing with the temperature of right HVC (Fig. 4-6 c,d).

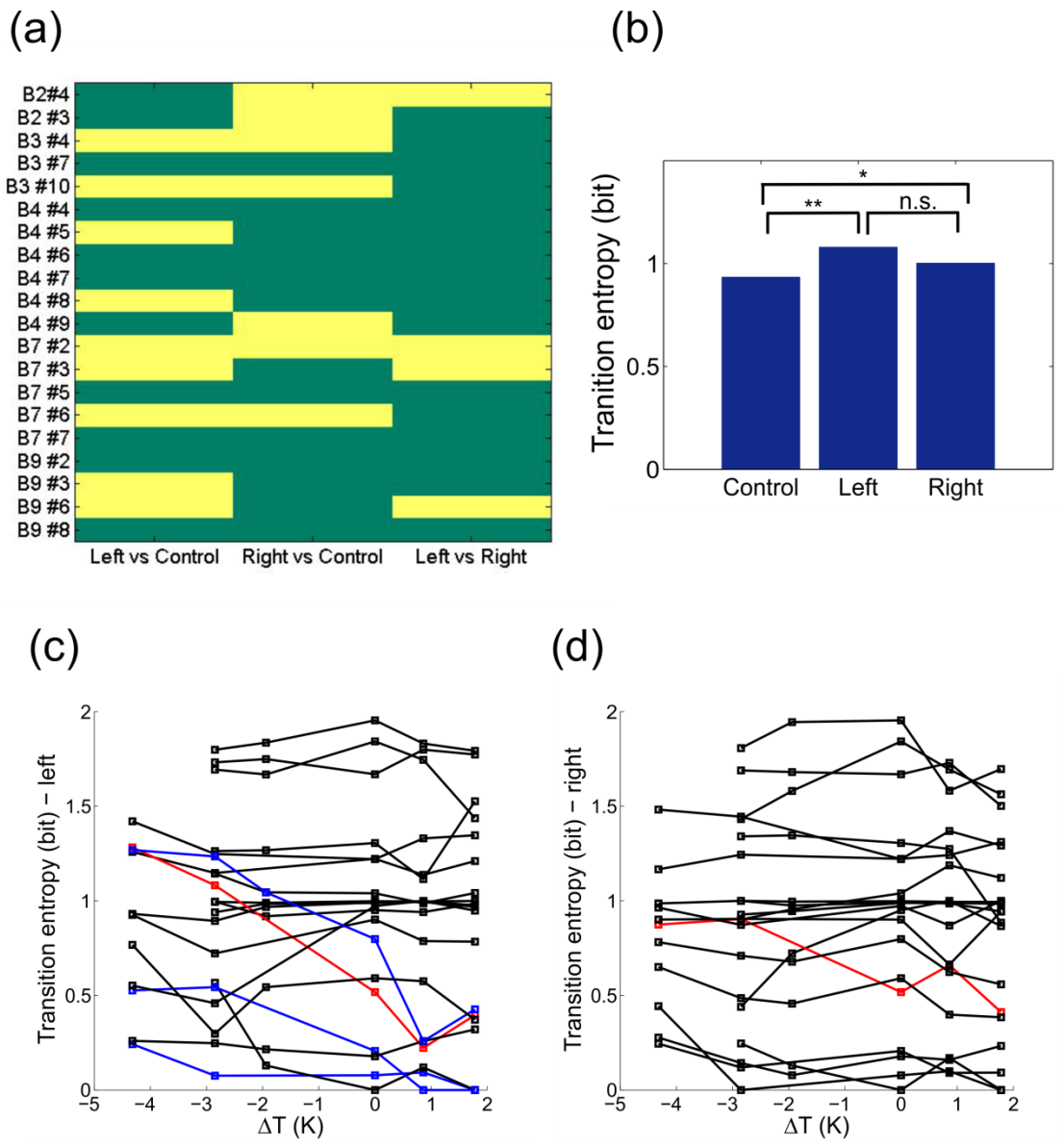


Figure 4-6. The unilateral cooling effect on the branching points

(a) A summary of the unilateral cooling effect. Each row is a branching point and each column is a comparison between unilateral cooling and control or between the left and right cooling. Yellow blocks highlight the significantly different groups. (b) Mean transition entropies in three conditions: control, cooling left HVC, and cooling right HVC. Difference compared by paired Wilcoxon signed rank test. (c, d) Transition entropies of 20 branching points with temperature. The left plot is the transition entropies with left HVC temperature; while the right plot is the transition entropies with right HVC temperature. Blue curves are the ones that significantly decrease with temperature by one side but not the other. Red lines significantly decrease with temperature of either side.

4.4 Discussion

Our results demonstrate that the timing control of the syllables of Bengalese finch songs is mostly left dominant. Both hemispheres are involved in the control of the inter-syllable gaps and the gap stretch by cooling left and right are anti-correlated, indicating a switching between the two hemispheres during the gaps. The syntax control, on the other hand, is more bilaterally involved. Most repeat lengths are controlled by one of the hemispheres. Both HVCs seem to play a role in the branching points.

We were trying to find a consistent relationship between the timing control of a syllable and the features of the syllable. We first looked at the acoustic features of the syllable and found that syllable stretch by the left cooling is more for the higher frequency and lower entropy syllables. The right cooling effect on syllable timing does not show strong correlation with any of the syllable features presumably because of its relatively weak effect on syllable duration. We pooled all the birds together, hoping to find a general relationship between the hemispheric-

specific timing control and the sound of the syllables for this species. Our result is quite consistent with Secora's finding in the peripheral study that the left side of Bengalese finch's syrinx produces higher frequency and lower entropy sounds[126], given that the two sides of the syrinx are ipsilaterally controlled by the premotor area. We then were seeking for a relationship between the syntactic role of a syllable and its stretch by unilateral cooling. We did not see that left or right HVC preferentially control a certain type of syllables in terms of their syntactic roles. Since our sample size is relatively small, this finding may not be conclusive. Also it could be that our categorization of syllables into deterministic, branching and repeated syllables is not the best way to describe the syntactic role of syllables. But overall, our observation suggests that timing effects appear to be related to certain acoustic features (fundamental frequency and entropy) but not the syntactic characteristics of the syllables.

In zebra finch, the two HVCs have synchronized activity about 40ms before the syllable onsets.[65] We found that both cooling left HVC and cooling right HVC can stretch gaps and their effects are anti-correlated, indicating that during a gap, one hemisphere is more dominant in the timing control than the other. If the synchronization is driven by a common input to both HVCs, cooling HVC will not affect when they are synchronized. The anti-correlation of the two HVCs' control in gaps suggests that synchronization might be driven by the dominant hemisphere at that moment, and cooling in one hemisphere can only lengthen the timing of some of the gaps but not others.

The change of transition probabilities at the branching points by unilateral cooling does not seem to be lateralized to one hemisphere. Particularly, some branching points are affected to a similar degree by either left or right cooling. If the syntax is encoded by the neural activity in one hemisphere, cooling in the other one should not have any effect. Therefore, at least for some of the branching points, the selection of the transition is determined by a signal that both HVCs share and that signal can be altered by changing temperature in either HVC. For example, Uva

receives ascending input from the respiratory brainstem network of both sides and projects back to HVC directly or via NIf.[66] The delay of signal caused by cooling one HVC could propagate to both sides via this feedback loop. If the selection of a transition is determined by the time when the later feedback signal arrives, one can expect that cooling on either side would have the same effect. To test this, one can record in both HVCs simultaneously during singing and see if there is correlation between the time when they are synchronized and the selection of the transition.

4.5 Conclusions

In summary, our results demonstrate that the two hemispheres have distinct control of song timing and song syntax. This study provides the first evidence for asymmetric syntax control by the two hemispheres in the same bird. Further work is needed to investigate how the neural behavior may differ between the two hemispheres in producing such functional differences. For a comprehensive description of the neural mechanisms of syntax generation, future models need to take both hemispheres into consideration.

Chapter 5

A simple miniature device for wireless stimulation of neural circuits in songbirds

To study the neural control of syntax, it is desirable to rapidly alter neural activity when the bird is singing. A useful method is electric stimulation. In this chapter, a simple, low-cost and extremely lightweight wireless neural stimulation device is demonstrated. The device has low power consumption and does not require a high-power RF preamplifier. Neural stimulation can be carried out in either a voltage source mode or a current source mode. Using the device, we carry out wireless stimulation in the premotor brain area HVC of a songbird and demonstrate that such stimulation causes rapid perturbations of the acoustic structure of the song.

5.1 Introduction

Electrical stimulation of neural circuits is an important tool in neuroscience allowing the studies of the connectivity of neural circuits[170] the relation of neural activity to behavior[125,171], and the reward mechanisms in the brain[172]. Electrical stimulation of the brain also has important therapeutic applications.[173,174] When applied to studies of the effects of neural stimulation on animal behavior, it is important that the chronically implanted stimulating device does not interfere with the animal movement and behavior. Wireless chronic stimulation systems offer a significant advantage in that the animal does not need to be tethered, thus, the behavior is minimally affected by the stimulation device. In addition, wireless stimulation allows carrying out behavioral experiments which are virtually impossible with the stimulation systems requiring a tether, for example, for studies of animal behavior in 3D environments.

Below, we describe a simple and low-cost device for wireless stimulation of neural circuits in behaving animals. The device has a simple schematic, is lightweight and is easy to make from commercially available components.

A number of wireless neural stimulation devices have been described in the literature.[175–177] Despite recent progress of the wireless stimulation techniques, the task of neural stimulation in small behaving animals (e.g., songbirds, mice) is not easily solved with existing devices. Wireless stimulation devices have to make compromises between the size/weight, the power consumption, the capabilities of the device and the ease of the device fabrication and control.[178,179] Most lightweight devices that are suitable for chronic studies of small behaving animals are based on custom-made micro-chips[180,181], thus, requiring microfabrication capabilities that are often beyond the means of neuroscience labs. Recent designs of wireless stimulating devices often use a digital mode of communication between the transmitter and the receiver.[182,183] Although sending digital commands to the wireless stimulation device is usually more reliable, digital control of the wireless device complicates the hardware and the software used for the control of the neural stimulation and, additionally, usually increases the power consumption of the device, thus, either reducing the battery life or increasing the weight of the device. There is a need for simple, low-cost and lightweight neural stimulation devices that can be built from off-the-shelf components and controlled using simple and customizable interfaces.

Below, we demonstrate a simple and low-cost device for wireless stimulation of neural circuits in small behaving animals. The device weighs only 1.4 g and is suitable for studies of songbirds that weigh only 12–15 g. The device can be made using readily available components. It can be used either as a voltage source or, with a slight design modification, as a current source. In the voltage source mode, the device has an excellent battery life. Triggering the neural stimulation pulses is carried out by an external command voltage pulse (+5 V); thus, it is easy to control the stimulation and to make the system compatible with a variety of software and

hardware platforms. The RF voltage required for neural stimulation is quite low (<10 V amplitude), so wireless neural stimulation can be carried out with a commercially available RF source and without a high-power RF amplifier.

5.2 Methods

5.2.1 Overview

The wireless stimulation system consists of two parts—an RF transmitter and a light-weight receiver/stimulator that is affixed to the head of the animal (Fig. 5-1). The transmitter supplies RF current through the primary coil that is wound around the animal cage. The receiver picks up RF signals from the transmitter via a secondary coil and generates voltage or current pulses that stimulate the brain area of interest through a bipolar stimulating electrode.

The coupling between the primary coil (transmitter) and the secondary coil (receiver) is near-field inductive: a magnetic field generated by the current in the primary coil induces the e.m.f. in the secondary coil. The voltage induced in the secondary coil is rectified, and the rectified voltage closes the FET switch, thus enabling the electrical stimulation of the brain area of interest.

The transmitter generates a pulse of RF current when turned on by an external command pulse. The command pulse (+5 V) can be generated by either a pulse generator or by a computer using an A/D card. Thus, the stimulation pulse timing and duration are controlled by simple and intuitive means and can be easily adjusted for a specific task. Below, detailed descriptions of the transmitter and receiver operation are given.

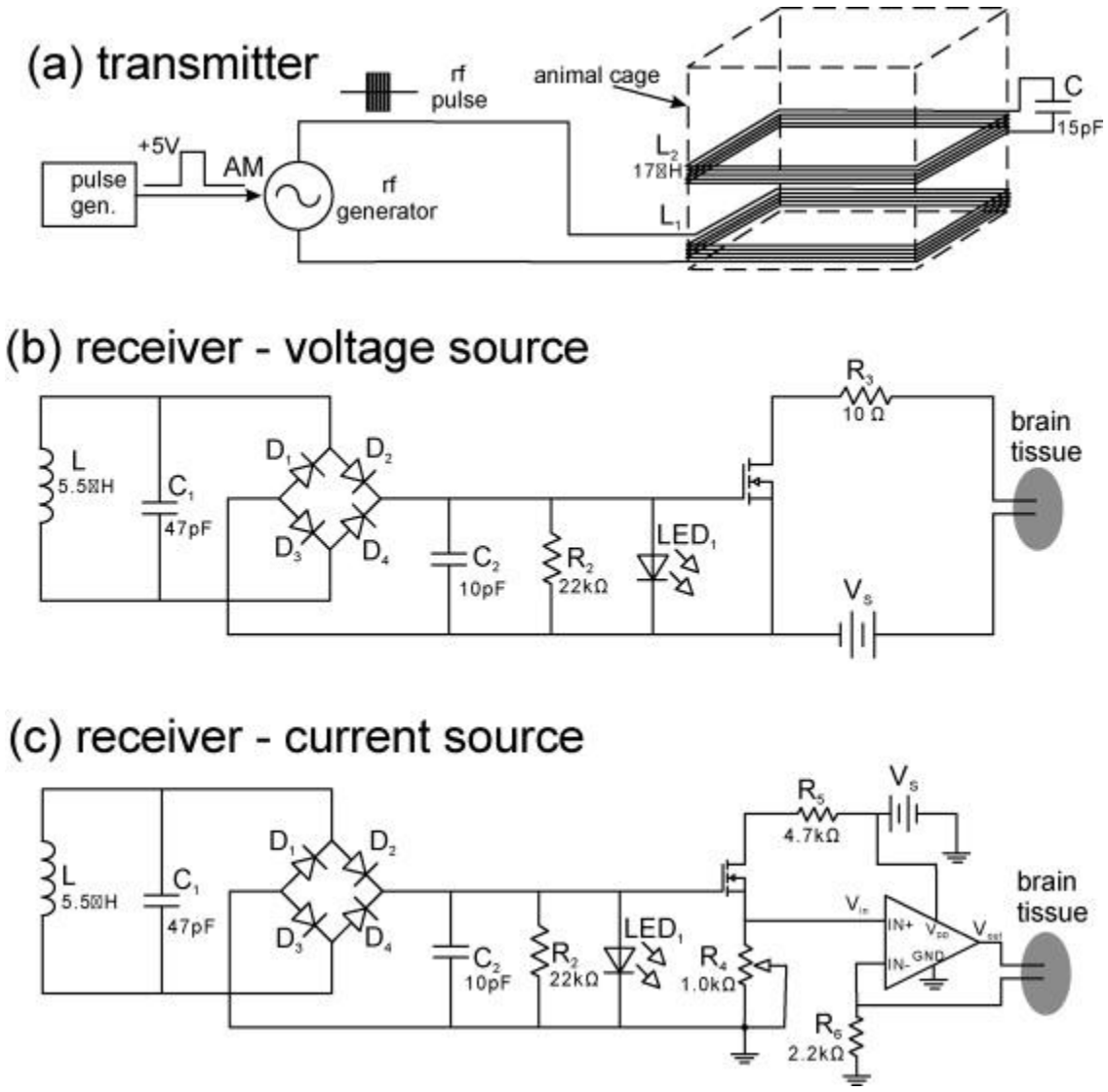


Figure 5-1. Schematic of the wireless stimulation system.

(a) Transmitter. (b) Schematic of the receiver and voltage source stimulator. (c) Schematic of the receiver and current source stimulator.

5.2.2 Transmitter

The block diagram of the transmitter is presented in Fig. 1a. The RF signal is generated by an RF generator (Stanford Research Systems, model DS345); the amplitude modulation (AM)

input of the RF generator is connected to the output of the pulse generator (AMPI, model Master-8). The command pulse (+5 V) from the pulse generator triggers the synthesizer to generate an RF pulse with the user-defined voltage amplitude up to 10 V. The output of the synthesizer is connected to a wire-wound coil L1.

The primary coil system consists of two wire-wound coils L1 and L2 that are placed around the animal cage (the cage dimensions are 8 in.×7.5 in.× 7 in., L×W×H). Coil L1 is placed near the bottom of the cage; coil L2 is placed about 1 in. above L1. Both coils are made out of 5 windings of 22 gauge insulated copper wire. Coil L1 is connected to the output of the RF generator. Primary coil L2 has an inductance of 17 μH and is connected to a capacitor C ≈ 15 pF. Thus, L2 and C form a resonant circuit having a resonant frequency of $f_{res} = \frac{1}{2\pi\sqrt{L_2 C}} \approx$ 9.9MHz. Coils L1 and L2 form an air-gap transformer: the ac current in L1 induces the ac current in L2. At the resonant frequency, the high-Q resonant circuit has very low impedance; therefore, the current in L2 is much larger than the current in L1. Using this air-gap transformer enhances the coupling from the transmitter to the receiver and enables wireless stimulation without a high-power RF preamplifier which is needed for many designs described.[125,180]

5.2.3 Receiver

The receiver circuit is assembled on a small (8 mm×8 mm) piece of thinned PC board. Two rechargeable batteries (Panasonic Model ML621 3 V battery) are used to power the device. A fully assembled receiver chip is shown in Fig. 2a. Two different receiver circuits were made and tested—a circuit that acts as a voltage source during stimulation (Fig. 5-1b) and a circuit that acts as a current source during stimulation (Fig. 5-1c). The secondary coil, the full-wave rectifier

and the MOSFET switch are common in these two circuits; the operation of these circuit elements is described below and applies to both the voltage source and the current source circuits.

A secondary coil L is made out of 0.01 in. diameter copper magnet wire (Small Parts, Inc); the coil is 12 mm in diameter and has 16 wire windings. The secondary coil has an inductance of about 5.5 μH and has Q > 200 at 10 MHz. It is connected to a capacitor C1 = 47 pF, forming a resonant circuit with a resonant frequency of 9.9 MHz (matching the resonant frequency of the transmitter). This resonant circuit picks up an ac magnetic field created by the RF current in the primary coil. It also acts like a band-pass filter making the device insensitive to the electromagnetic interference at non-resonant frequencies. The device can be used without any special shielding or electromagnetic interference protection unless large currents or powerful electromagnetic radiation sources having a frequency close to the resonant frequency (9.9 MHz) are present in the proximity of the experiment.

The ac voltage across the capacitor is rectified with a full-wave rectifier D1–D4 (two-diode arrays BAT721A and BAT721C, NXP Semiconductors); the reservoir capacitor of the rectifier C2 is chosen to be relatively small (C2 = 10 pF) and, in addition, is shunted by a resistor R2 = 22 kΩ. These values of C2 and R2 allow rapid (the RC time constant is $\tau = 0.22$ μs) changes of the rectifier output necessary for generation of brief (down to tens of microseconds) stimulating pulses. The output voltage of the rectifier is applied to the gate of the MOSFET (NTA4153NT1G, ON Semiconductor). The MOSFET has a low threshold voltage (~0.8 V) and acts as a switch—when the gate voltage is greater than 1.1 V, the drain-source resistance of the MOSFET is very low (Rds < 15 Ω at Vgs = 1.1 V), and the current can flow through the stimulating electrode and the brain tissue.

Because of the rapid charging and discharging of capacitor C2, the circuit has a very brief (microseconds) “memory” of the previous stimulation pulses and can be used for delivering high-frequency pulse trains. We have triggered the circuit with high-frequency pulse trains and found

that the amplitude and the time course of the stimulation pulses delivered by the device are independent of the command pulse frequency up to the pulse frequency of 1 kHz (data not shown). The stimulation pulses for the 1 kHz frequency of the command pulses is shown in Fig. 2(c and d).

An LED is connected in parallel with the storage capacitor C2 for easy monitoring of the receiver operation. When the LED is on, the gate voltage of the MOSFET is greater than ~1.8 V. Thus, the light of the LED is a conservative indicator that the MOSFET switch is closed and the current can flow through the stimulating electrode.

5.2.4 Voltage and current source modes of stimulation

In the voltage source mode (Fig. 5-1 b), when the switch is closed, a voltage pulse is applied to the stimulating electrode, and the stimulating voltage is the battery voltage VS = 6 V. The stimulating current is determined by the impedance of the stimulating electrode and the brain tissue. For our electrodes, Z = 10–40 kΩ, so the stimulating current is in the range 150–600 μA, depending on the electrode impedance. In the voltage source mode, the stimulating current cannot be controlled. However, the intensity of the stimulation can be controlled by changing the stimulating pulse duration—shorter pulses will be less effective in stimulating neurons.[184,185] When the stimulating pulse is not applied, the current through the battery is extremely small (it is determined by the drain-source conductance of the MOSFET when Vgs = 0 V and is less than 50 nA) and, therefore, the device has an excellent battery life (the factor limiting the battery life is battery shelf life). However, the inability to control or precisely know the stimulating current is a drawback that motivates an alternative mode of operation where the stimulation current is known and can be controlled.

The circuit in Fig. 1c converts the voltage across resistor R4 into current, thus, allowing control of the stimulating current. A micro-power operational amplifier (Texas Instruments TLV2381IDBVT) and a resistor R6 form a voltage-to-current converter. The output current of the op-amp is: Iout = Vin/R6, and the input voltage is: Vin = VS R4/(R5 + R4). Therefore, by adjusting the variable resistor R4 one can tune the stimulating current: Istim = VS R4/((R5 + R4)R6). The maximum stimulation current is limited by the maximum output current of the op-amp which is about 400 μA. The voltage compliance of the current source is equal to the battery voltage VS. Thus, for VS = 6 V, if the stimulating electrode has an impedance of 15 kΩ, the maximum current that can be delivered is 350 μA. For the circuit in Fig. 1c, the current through the battery in the absence of stimulation (quiescent current) is dominated by the supply current of the op-amp and is about 7 μA. Although this is significantly larger than the quiescent current for the voltage source circuit, this current is small enough to ensure over 2 weeks of battery lifetime.

5.2.5 Stimulating electrode

The stimulating electrode is a bipolar electrode made of teflon-insulated stainless steel wire (bare wire diameter 0.002 in., diameter of the insulation 0.0045 in., A-M Systems). The distance between the wires is 0.5 mm; the insulation is removed from the top 0.2–0.3 mm of the two wires; the wires are 7 mm long and are soldered to a 2-pin connector (Omnetics, model A9069-001). The impedance of these electrodes in phospho-buffered saline (PBS) (pH 4.0, Sigma–Aldrich) is typically between 10 and 40 kΩ at 1 kHz.

5.2.6 *In vivo* brain stimulation

An adult (>100 days post-hatch) male zebra finch (Taeneopygia guttata) was used to test the device. The details of the surgical procedure are described elsewhere (Fee and Leonardo, 2001). Briefly, the bird was anesthetized with 1–2% isofluorene. A small opening in the skull was made over right HVC, and the stimulating electrode was inserted to a depth of about 500 μm. The electrode was affixed to the skull using dental acrylic. All procedures were carried out in accordance with the locally approved IACUC protocol.

The receiver chip was attached to the animal skull in a way allowing easy removal and re-attaching (e.g., for changing the batteries). Four metal pins (P/N 52020, A-M Systems) were connected to the receiver board; four matching metal pins (P/N 52010, A-M Systems) were attached to the skull of the animal with dental acrylic. The pins on the receiver board were plugged into the pins attached to the skull, thus affixing the receiver to the skull. The receiver chip was connected to the implanted stimulated electrode with a 2-pin connector (Omnetics, model A9577-001), allowing to easily connect and disconnect the stimulating electrode and the receiver/stimulator.

Wireless electrical stimulation of HVC was carried out in the voltage source mode. Based on the measurements of the electrode impedance, the stimulating current was about 180 μA. For most trials, the pulse duration of 200 μs was used. For trials where the effect of longer stimulating pulses on the song was studied, the pulse duration of 5 ms was used. Neural stimulation during singing was carried out on interleaved songs.

5.2.7 Data acquisition and analysis

Both the acoustic signal and the output of the pulse generator were digitized at 40 kHz with a 16-bit A/D card (PCI-6251, National Instruments) and stored on a computer. Analysis of the song spectrograms was carried out offline.

Song spectrograms were computed using a Hamming window with the window duration of 12.8 ms and a window overlap of 10 ms. One unperturbed song motif was chosen as a template; four syllable templates (A, B, C, and D) were defined in this template motif (Fig. 4a). In order to assess the effects of the wireless stimulation on the song, we compared the acoustic structure of the song syllables following the stimulation pulse (syllables "perturbed by stimulation") with the "intact" song syllables (syllables in the songs when no stimulation was carried out). The presence and the magnitude of the effects of electrical stimulation in HVC on the zebra finch song vary with the temporal location of the stimulating pulse.[125] We restricted our analysis to two temporal windows of the stimulating pulse in the song motif where stimulation has caused observable perturbations of the song structure: (1) when the stimulating pulse occurred between the syllables A and B – for this case, perturbations in the acoustic structure of syllable B were analyzed – we refer to these syllables as "perturbed syllable B" and (2) when the stimulating pulse occurred during syllable B – for this case, perturbations in the acoustic structure of syllable C were analyzed – we refer to these syllables as "perturbed syllable C".

We computed a spectral similarity score between the template syllable B and each rendition of a "perturbed syllable B" and also between the template syllable B and each rendition of syllable B in the unperturbed songs. The same procedure was repeated for syllable C. The spectral similarity score between two syllables is a Pearson correlation coefficient between the spectrograms of the syllables[68]; for syllables having identical spectrograms, it has a value of 1.

The previously published procedure for computing the similarity scores was slightly modified for this study as described in Methods.

The distributions of the similarity scores for unperturbed and perturbed syllable B were compared with each other, and the stimulation was considered effective in perturbing the song if the similarity scores of perturbed syllables were significantly ($P < 0.05$, one-sided Kolmogorov–Smirnov test) lower than the similarity scores of the unperturbed syllable B. The same procedure was repeated for syllable C.

5.2.8 Electrophysiological recordings

Recordings of stimulation artifacts and neural activity were carried out using carbon fiber electrodes (Carbostar-1, Kation Scientific). The bird was anesthetized and placed into a stereotaxic instrument, and both the stimulating and recording electrodes were inserted into a premotor brain area robust nucleus of arcopallium (RA). The receiver was placed near the center of the primary coil and connected with a cable to the stimulating electrode. The receiver was triggered by RF pulses, and the stimulation artifacts and neural responses to the stimulation were recorded. The recording electrode signal was amplified ($G = 2000$) and low-pass filtered with a cutoff frequency of 5 kHz. The amplified signals were digitized using a digital oscilloscope (Tektronix 2004B).

5.3 Results

The device performance was characterized both ex vivo and in vivo. For ex vivo tests, the stimulating electrode was immersed in phosphor-buffered saline. For in vivo tests, the stimulating electrode was inserted into the brain of an anesthetized bird. We recorded the voltage and current

pulses generated by the device as well as stimulation artifacts and stimulation-evoked neural activity in an anesthetized bird. In addition, we tested the device in freely behaving small animals. The stimulating electrode was implanted in the brain area HVC of a zebra finch (HVC is a premotor area critically involved in generating the song). Using the wireless stimulator, we tested the effects of brief (0.2–5 ms) voltage pulses on the acoustic structure of the bird's song.

5.3.1 Coupling from the transmitter to the receiver

Successful operation of the device is dependent on whether enough RF power to close the MOSFET switch is coupled from the transmitter to the receiver. The MOSFET switch is closed (Rds < 15 Ω) when Vg > 1.1 V. The LED turns on when the voltage on the LED is greater than ~1.8 V; the light of the LED was used as an indicator that a sufficient gate voltage is applied to MOSFET. With the lab lights on, the LED can be easily seen with naked eye for pulse durations down to 40 μs (the shortest duration that could be achieved with our pulse generator); this provides a very convenient and noninvasive way to test that the receiver circuit is working properly while the receiver is on the behaving animal.

With an RF voltage amplitude of 8 V (into a high-impedance load) at 9.9 MHz, the LED light was on at all x–y locations inside the cage between the plane of the primary coil and 3–4 in. above the plane of the primary coil (see Supplementary Information). The coupling between the primary and the secondary coil is directional, and maximum coupling occurs when the axes of the primary and secondary coils are parallel to each other. The spatial range where the LED turns on is not significantly decreased when the orientation of the secondary coil is changed by 25 degrees or less; even larger receiver orientation changes (e.g., 45 degrees) have only a moderate effect on the spatial range of the stimulation.

The resonant frequency of the primary coil changes slightly with the vertical separation between the coils L1 and L2. The separation between L1 and L2 was adjusted to yield maximum brightness of the LED on the receiver before the stimulating electrode was connected to the receiver.

There is a delay of about 30 μs between the onset of the command (+5 V) pulse and the onset of the neural stimulation pulse. This delay is mostly due to the delay between the onset of the command pulse and the onset of the RF pulse (25 μs), there is also a smaller time lag (~5 μs) between the onset of the RF pulse and the onset of the stimulation pulse. These delay values are small compared to the pulse duration (200 μs) and to the duration of a typical neural spike, so, for most neural stimulation experiments, the neural stimulation can be considered effectively synchronous with the command pulse.

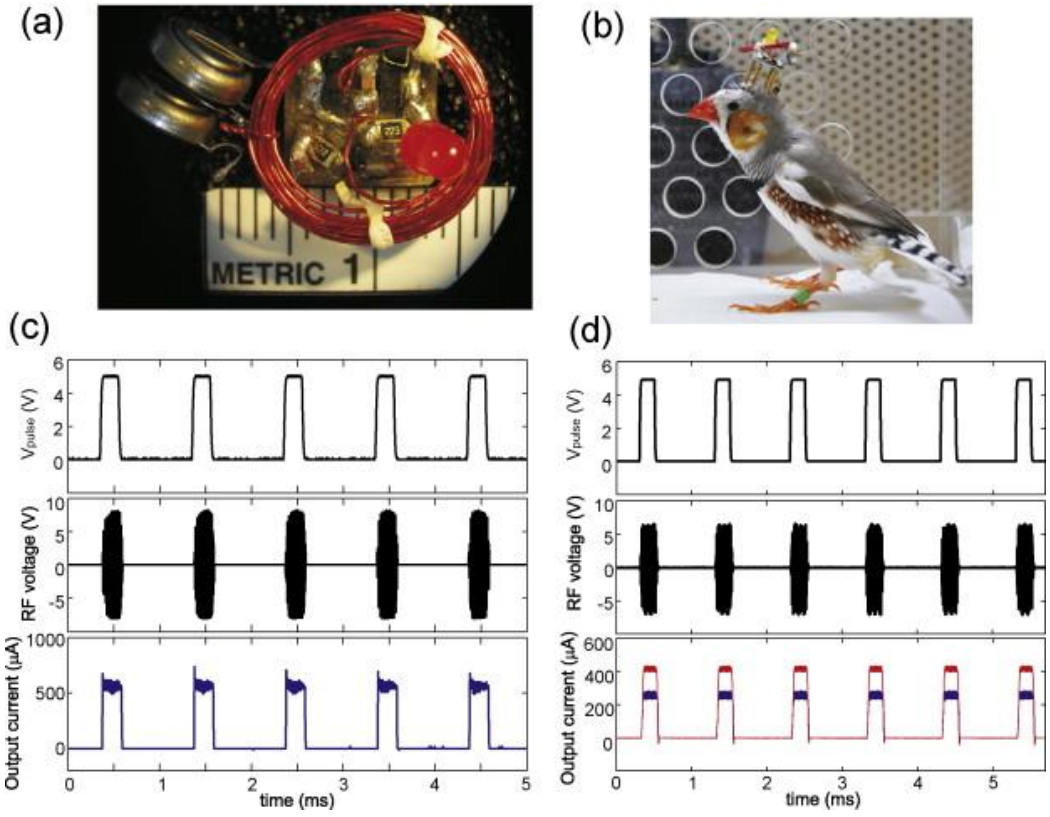


Figure 5-2. Device and in vitro test

(a) The top view of the chip with externally attached batteries, coil and LED. (b) A zebra finch implanted with the chip. (c and d) In vitro measurements of the current through the electrodes in PBS in both voltage and current source mode—top traces: external triggering pulses; middle traces: amplitude-modulated RF voltage applied to the primary coil system; bottom traces: current pulses in the voltage source mode (c) and in the current source mode for two different current set values (d). The stimulating electrode is immersed in saline.

5.3.2 Stimulator performance

Voltage source mode

In the voltage source mode, during the stimulating pulse, the voltage across the stimulating electrode is equal to the battery voltage VS (the voltage drop across the resistor R3 is negligible compared to the voltage drop across the electrode). The current is determined by the impedance of the electrode in the tissue. The stimulating current can be monitored by measuring the voltage drop across the resistor R3. Fig. 5-2c shows the time dependence of the current through the electrode in the voltage source mode when the electrode is immersed in PBS. The stimulating pulses are 200 μs long. The current is about 550 μA; the slightly larger current in the first 30 μs after the onset of the pulse is due to the charging of the double layer capacitance. The measured rise time of the current is 2 μs, the fall-time is 8 μs.

Current source mode

With the current source design, the stimulating current can be set by adjusting resistor R4. The voltage compliance of the current source is equal to the battery voltage VS = 6 V (if two batteries are used). The maximum stimulation current that can be delivered is limited by the short-circuit current of the op-amp which is about 400 μA. Fig. 5-3a shows the current pulses

delivered by the device when the stimulating electrode is in saline for currents between 25 and 245 µA. The duration of the trigger pulse is 200 µs. The rise time of the current pulses is about 35 µs; the fall time is about 20 µs.

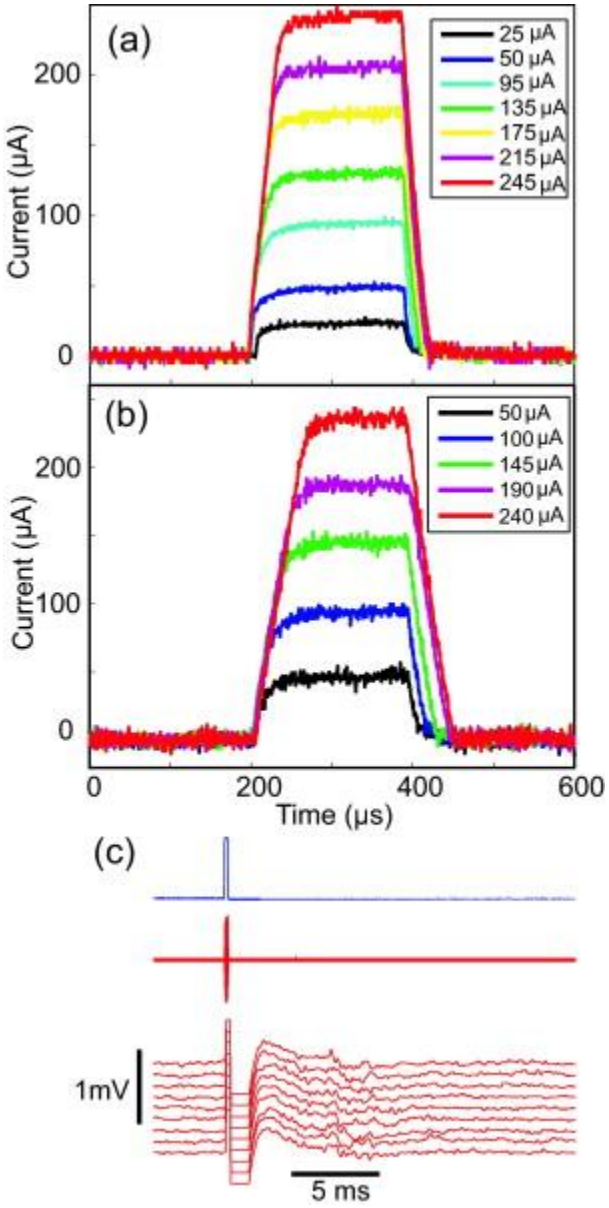


Figure 5-3. Tuning the stimulating current in the current source mode of the device.

Labels on both plots indicate the values of current set by adjusting resistor R4. (a) The current pulses of different amplitudes with the electrode immersed in saline. (b) The current pulses of different amplitudes with the electrode inserted in the brain of anesthetized bird. The supply voltage VS = 6 V for all curves except VS = 9 V for the 240 μA curve. (c) Command pulse (top trace), RF pulse (middle trace), and stimulus artifacts and stimulation-evoked multi-unit neural activity (bottom traces) in premotor brain area RA of anesthetized bird.

5.3.3 Neural activity evoked by wireless stimulation

We recorded the stimulation artifacts and the multi-unit neural activity evoked by the stimulation in brain area RA (see Section 2). Fig. 5-3c shows the stimulation pulse artifacts and the increase of the multi-unit neural activity in the time window between 3 ms and 10 ms following the onset of the stimulation pulses. The stimulating pulse duration is 200 μs; the current is 200 μA. The transient increase of neural activity following the stimulation pulse demonstrates that the device is effective in evoking neural activity in vivo.

5.3.4 Wireless stimulation of neural circuits in small behaving animals

We have implanted a stimulating electrode in the premotor area HVC (used as a proper name) of an adult zebra finch and studied the effects of brief electrical stimulation in HVC on the acoustic structure of the song. The song of a zebra finch consists of repeating sequences of vocal elements called the song motif.[18] Each song motif consists of several syllables[21]; the acoustic structure of the motif and syllables in adult bird is highly stereotypical.[18] HVC plays a critical role in the premotor coding of the bird's song[29,43,45], and electrical stimulation in HVC has been shown to perturb the acoustic structure of the song (Vu et al., 1994, Vu et al., 1998 and Wang et al.,

2002).[53,125,163] Because of high stereotypy of the song, the detection and the analysis of the song perturbations due to electrical stimulation is greatly simplified.

Fig. 4(a) shows several instances of the song motif consisting of four syllables (A, B, C, and D) in unperturbed songs (all motifs are aligned to the onset of syllable B). Fig. 5-4(b–d) shows the effects of the stimulation on the acoustic structure of the song. When the stimulating pulse (200 μs pulse duration) occurs between syllables A and B, the pulse causes distortions of the acoustic structure of syllable B (Fig. 5-4b). Stimulation during syllable B (200 μs pulse duration) leads to either a distortion of the acoustic structure of the subsequent syllable (C), or to a motif termination (Fig. 5-4c). Longer stimulating pulses lead to stronger distortions of the song: for 5 ms pulse duration, stimulation between syllables A and B and during the syllable B usually causes song termination (Fig.5-4d). Fig. 4(e and f) shows the spectral similarity scores for intact syllables B and C and for the syllables B and C perturbed by 200 μs long stimulation pulses to template syllables B and C (see Section 2). For both syllables B and C, stimulation significantly reduces the spectral similarity score of the syllable to the syllable template: the P-values for the distributions of similarity scores are $P < 3 \times 10{-}5$ for syllable B and $P < 10{-}6$ for syllable C. This indicates that the stimulation causes statistically significant perturbations of the acoustic structure of the song syllables.

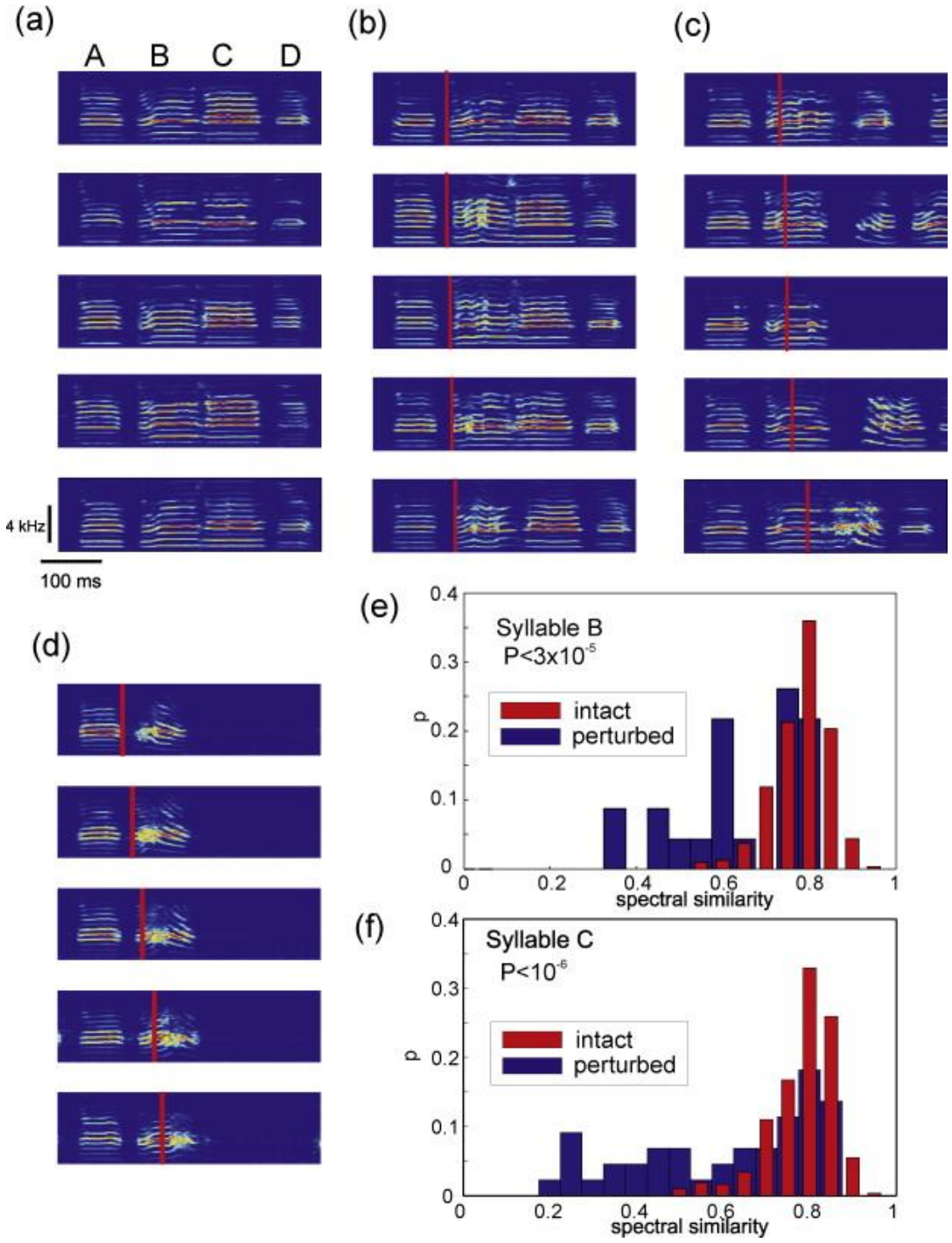


Figure 5-4. Effects of wireless stimulation in premotor brain area HVC of a zebra finch.

(a) Intact song motifs are highly stereotypical and consist of syllables A, B, C and D. (b) Stimulating between syllables A and B causes perturbations of syllable B. Vertical red lines are

temporal locations of stimulating pulses. (c) Stimulating during syllable B causes distortions of syllable C. For (b) and (c), the pulse duration is 200 μs. (d) Longer stimulating pulses (pulse duration 5 ms) cause stronger song perturbations—terminations of the song motif. (e) and (f) show the distributions of the spectral similarity scores for the intact syllables and for the syllables perturbed by the stimulation. (e) Distribution of the similarity scores of syllable B, nintact = 312, nperturbed = 23, $P<3\times10^{-5}$ (Kolmogorov–Smirnov test). (f) Distribution of the similarity scores of syllable C, nintact = 320, nperturbed = 44, $P<10^{-6}$ (Kolmogorov–Smirnov test).

5.4 Discussion

The device described above provides a simple and convenient means to carry out wireless stimulation of neural circuits in behaving animals. The device possesses several attractive features that can significantly simplify neural microstimulation experiments in small behaving animals.

Small weight allows using the device for wireless neural stimulation in small animals that cannot carry a lot of weight (e.g., songbirds that may not be able to carry more than 2 g weight on their heads). The device can be made using commercially available components and without the need to fabricate custom micro-chips. This eliminates the expense and the necessity to have access to microfabrication facilities and technical expertise to design and make micro-chips. It is easy to connect and disconnect the device from the animal.

The device can be reliably triggered by an RF pulse having a voltage amplitude of only 8 V (into a high-impedance load); this voltage is sufficient for triggering the stimulator anywhere in a physical volume of approximately 8 in.×7.5 in.×3 in. ($L \times W \times H$). Such a low RF voltage can be generated by a number of commercially available RF generators with external amplitude

modulation input, and these RF generators have become relatively inexpensive. A simple way of triggering the stimulation with a +5 V command voltage pulse make the transmitter hardware and the control of the neural stimulation simple and intuitive, and the stimulation is effectively synchronous with the command pulse.

The device has voltage compliance equal to the total voltage of the batteries (6 V if two batteries are used) and a fairly high maximum stimulation current (about 400 μA). In a current source mode, it is possible to control the output current with a single variable resistor. In the voltage source mode, the maximum stimulation current can be even higher (in the mA range) if the electrode impedance is low enough. If high stimulation currents are desired, it is necessary to use low-impedance stimulation electrodes. The compliance voltage of the stimulator is equal to the power supply voltage *V*S and can be increased if additional batteries are used.

Our device has a flexible design that can be modified straightforwardly to tailor the device to specific experimental needs. It is possible to achieve higher stimulation currents by using an op-amp with higher maximum output.1 Several design modifications can be used to increase the spatial range where wireless stimulation can be realized. The increase of the number of windings or the loop area of the secondary coil will decrease the minimum strength of RF magnetic field required to trigger neural stimulation. Alternatively, the use of charge-pumping circuits[186] will enable to trigger the stimulator using lower RF voltage from the transmitter. Both of these modifications will increase the spatial range where wireless stimulation can be realized. Such an extension of the spatial range may be desirable for behavioral experiments involving larger animal enclosures (tracks, mazes, etc.).2

Finally, we would like to note that wireless stimulation devices based on custom-made microchips utilize much more advanced technology and, therefore, can deliver better performance than the simple circuits described in this paper. However, at present, the capabilities to make and to use these high-tech devices are not widely available in a neuroscience community. The

simplicity of assembling and the ease of control of the device described in this paper make it an attractive low-cost option for wireless neural stimulation in small behaving animals. Such simple wireless stimulation devices will keep being useful to neuroscientists.

5.5 Future applications

It has been demonstrated that this device can reliably perturb the neural activity in a singing bird and consequently affect its songs. This device can be used in combination with the fast syllable detection system as described in Chapter 3 to stimulate the neural circuit at the moment when the targeted syllable is sung. The stimulating pulses can also be delivered with various time delays so that can potentially stimulate the brain at any time during a song. This tool will be particularly useful for studying the moment-to-moment singing control by the involved brain areas. By carefully adjusting the amplitude of the current pulses, a small current may be able to vary the neural activity without interrupting the song. Such microstimulation can be potentially used to locate the neural substrates that control the syntax of the song.

The device can deliver pulses to one of two sets of bipolar electrodes with the other set shorted by a jumper wire. In this way, the two hemispheres can be studied and their roles in singing control can be differentiated.

Due to the short of time, the proposed experiments were not carried out before this thesis was written. Future researchers can use our device to continue the study about the neural mechanisms of syntax generation in songbirds.

Appendix A

A semi-automatic algorithm for song segmentation and classification

Song preprocessing

Songs were recorded with a sample rate of 40kHz. The sound pressure waveform was first zero-mean normalized. The waveform was converted to a spectrogram of 512 window size and 95% overlap, yielding a 0.6 ms time resolution and 78.1 Hz frequency resolution. The natural logarithm of the power spectral density (denoted as $\log(P)$) in the frequency range between 1 kHz and 10 kHz was used for the following processing. The summation of each column of $\log(P)$, roughly equivalent to the logarithmic scaled amplitude envelope, was obtained and denoted as v_1. The spectral entropy, as

$v_2(t_i) = \sum_f \log\big(\log\big(P(t_i, f)\big)\big)\Delta f - \log(\Delta f \sum_f \log(P(t_i, f)))$, was also calculated. Both v_1 and v_2 were rescaled to be in the range of [0,1]. The maximum of v_1 and v_2 at each time point was used for segmentation ($v_m(t) = \max\{\mathrm{v_1(t)}, \mathrm{v_2(t)}\}$). Considering both features for segmentation was intended to take both sounds of relatively large amplitude and relatively large entropy into account. Values contained in v_m that were smaller than 0.001 were leveled up to 0.001. The nature logarithm was taken for this signal to decrease the variation of the amplitude so that syllables with smaller amplitude can be detected (Fig. S1). Up to here, the pre-processing of the song signal is completed.

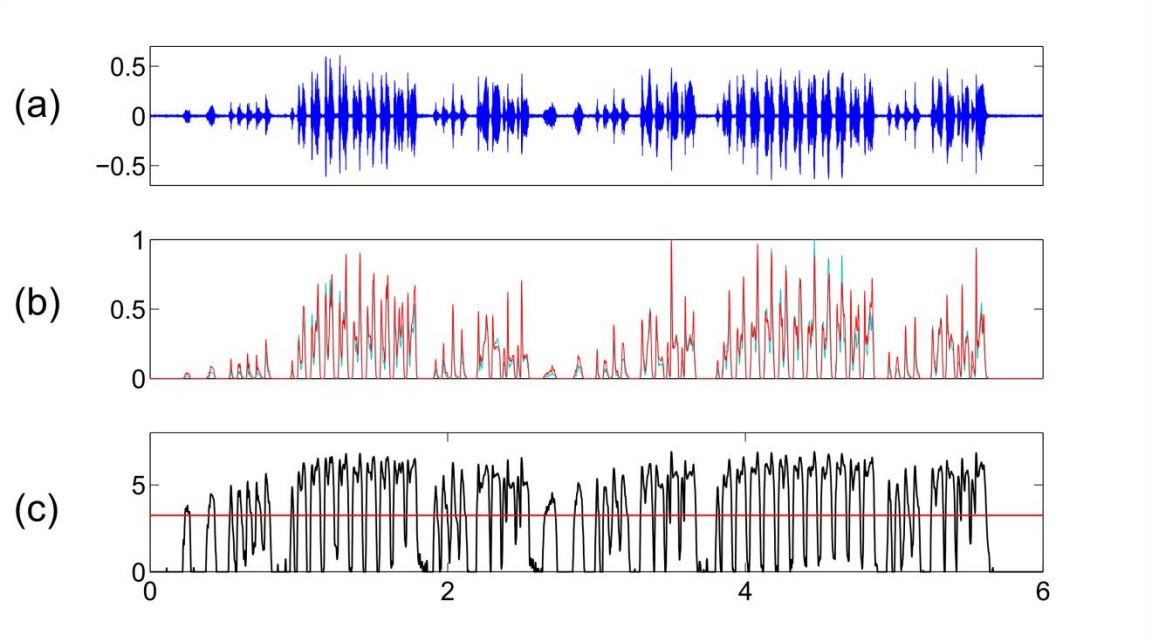


Figure S1 Signal pre-processing

(a) Sound pressure (raw signal). (b) Entropy (red) and amplitude (blue) of the signal. (c) The logarithmically scaled signal $\log(v_m)$. Red line is the mean of $\log(v_m)$ plus one half of its standard deviation.

Song segmentation

The basic idea for finding the syllable onsets and offsets is to find the local maxima and minima of the first derivative of the signal. In order to eliminate the local extremes within the syllables, amplitudes beyond a threshold as $\text{mean}(\log(v_m)) + \frac{1}{2} s.d.(\log(v_m))$ (as shown in Figure S1 c) were flatted out. Then the local maxima and minima were searched. The illegal detections were excluded, e.g., offsets before onsets were deleted, the highest peak was used if more than one peak were detected between two onsets/offsets. Then the syllables were determined based on several standards: minimal syllable duration, minimal gap duration, maximal gap duration before and after the syllable. Extremely short syllables were ignored. Very

close syllables were merged. Syllables that had large gap before and after it were considered as calls and were excluded. The segmentation result is shown in Figure S2.

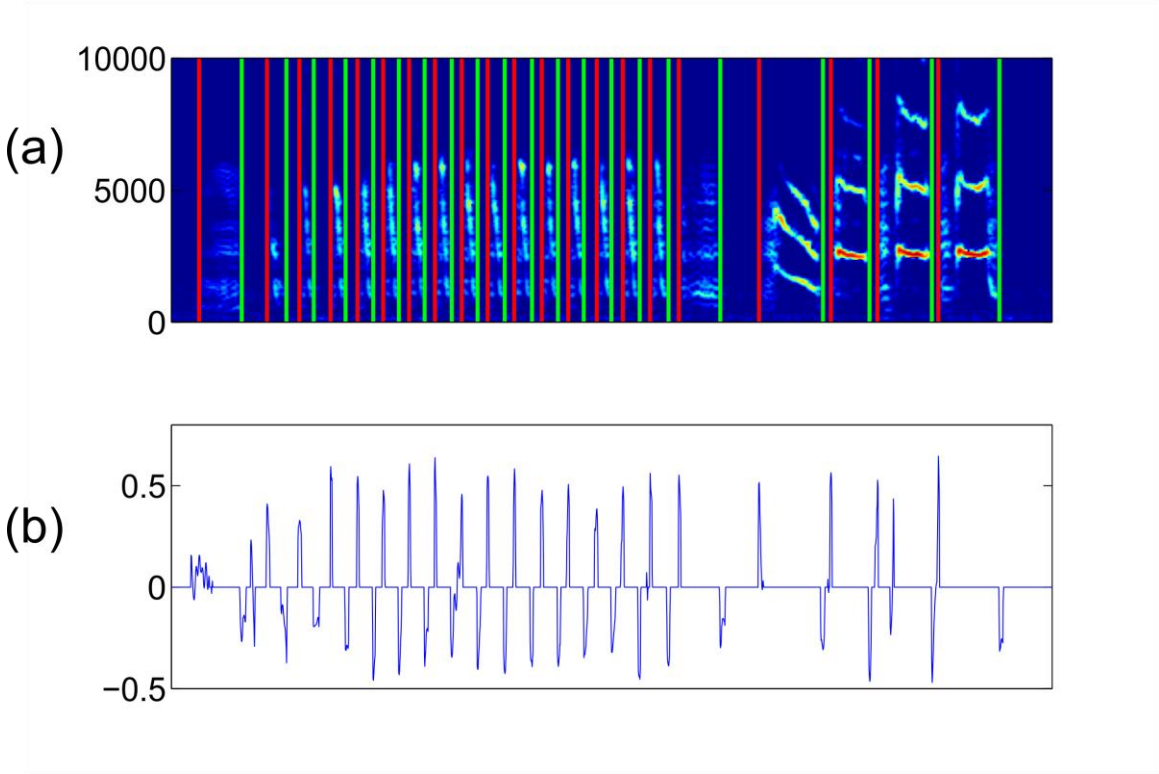


Figure S2 Song segmentation result

(a) Song spectrogram with segmentation. Red lines highlight syllable onsets and green lines highlight syllable offsets. (b) First derivative of the signal. Peaks are where the onsets are while valleys are where the offsets are.

Segmentation evaluation

This method has several pros. First, it is capable of detecting the syllables of relatively small amplitude. Second, the signal to noise ratio of the first derivative of the signal is high (Fig. S2 b), since many small variations during the gap and within the syllables have been flatted. Third, the determination of the syllable onsets and offsets does not depend on the threshold but

only the nature of the signal itself. This is good because we need a universal standard for all the songs that does not rely on human judgment.

The major con of this method is that sometimes it picks noises and mistakenly treat them as syllables. To avoid this, an extra step is needed before the signal is logarithmically scaled. A second con is that, compared to traditional threshold crossing methods, the local maxima and minima, especially the local minima, can be skipped. In this case, adjacent syllables may be joined. A potential problem of flatting the signal within the syllable by a threshold value is that the abrupt change of its first derivative may cause dips that are greater than those caused by real syllable offsets and consequently lead to false detections. To overcome this problem, one can smooth the signal within the syllables instead of setting it to one value.

Nevertheless, overall, this method solved the problem of amplitude-dependent threshold crossing and gives us syllable segments with relatively reliable onsets and offsets. It is particularly important for the study of song timing.

Syllable classification using SVM

The segmented songs were labeled manually. The syllable types were determined by eye. For each bird, about 10-20 songs were labeled manually for training a supported vector machine(SVM) for syllable classification. The syllable features used for classification were obtained as below. The nature logarithm of the power spectral density was converted to binary based on a threshold so that the spectrogram was converted to its contour. The first 65ms of the syllable(if shorter, padded with 0) was used. This 65ms binary matrix was converted to a single vector and this whole vector was used as the feature vector of one syllable. The syllable labels, together with their feature vectors, were used for training a SVM. For each song, the same training sets were used to predict the syllables in this song. This step was realized by an open source MATLAB function 'multisvm'.

Syllable classification evaluation

Correct syllable class prediction depends on correct song segmentation. The evaluation below is conditional on correct song segmentation. 118 songs containing 7857 syllables from 4 birds were used for this evaluation. The percentage correct prediction for the 4 birds are 100%, 92.2%, 98.3% and 99.5%, respectively, yielding a $97.5 \pm 3.6\%$ correct prediction on average. The performance of this method varies, depending on the syllable structures of a bird's song. Spectrally similar, but syntactically different syllables may be hard to be distinguished automatically. When manually classifying syllables, the classification involved human judgment. This error can be fixed given the context of the syllable.

Overall, this segmentation and classification method accelerated data analysis significantly.

Bibliography

1. Lashley, K. in *Cereb. Mech. Behav.* (Jeffres, L.) 112–136 (1951).
2. Botvinick, M. M. Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.* **12,** 201–208 (2008).
3. Reber, A. S. *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. (Oxford University Press, 1996).
4. Bizzi, E., Mussas-Ivaldi, F. A. & Giszter, S. Computations underlying the execution of movement : a biological perspective. *Science* **253,** 287–291
5. Pulvermüller, F. Brain embodiment of syntax and grammar: discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain Lang.* **112,** 167–179 (2010).
6. Pearson, K. G. Common Principles of Motor Control in Vertebrates and Invertebrates. *Annu. Rev. Neurosci.* **16,** 265–297 (1993).
7. Dick, T. E., Oku, Y., Romaniuk, J. R. & Cherniack, N. S. Interaction between central pattern generators for breathing and swallowing in the cat. *J. Physiol.* **465,** 715–730 (1993).
8. Syed, N. I., Bulloch, A. G. & Lukowiak, K. In vitro reconstruction of the respiratory central pattern generator of the mollusk Lymnaea. *Science* **250,** 282–285 (1990).
9. Jean, A. Brain Stem Control of Swallowing: Neuronal Network and Cellular Mechanisms. *Physiol. Rev.* **81,** 929–969 (2001).
10. Marder, E. & Calabrese, R. L. Principles of rhythmic motor pattern generation. *Physiol. Rev.* **76,** 687–717 (1996).
11. Hooper. Central pattern generators. *Curr. Biol. CB* **10,** R176 (2000).
12. Konishi, M. From central pattern generator to sensory template in the evolution of birdsong. *Brain Lang.* **115,** 18–20 (2010).
13. Konishi, M. Birdsong: From Behavior to Neuron. *Annu. Rev. Neurosci.* **8,** 125–170 (1985).
14. Fee, M. S., Kozhevnikov, A. A. & Hahnloser, R. H. r. Neural Mechanisms of Vocal Sequence Generation in the Songbird. *Ann. N. Y. Acad. Sci.* **1016,** 153–170 (2004).
15. Doupe, A. J. & Kuhl, P. K. BIRDSONG AND HUMAN SPEECH: Common Themes and Mechanisms. *Annu. Rev. Neurosci.* **22,** 567–631 (1999).
16. Marler, P. A comparative approach to vocal learning: Song development in white-crowned sparrows. *J. Comp. Physiol. Psychol.* **71,** 1–25 (1970).
17. Brainard, M. S. & Doupe, A. J. What songbirds teach us about learning. *Nature* **417,** 351–358 (2002).
18. Immelmann, K. Song development in the zebra finch and other estrildid finches. *Bird Vocalizations* **61,** (1969).
19. Thorpe, W. H. The Learning of Song Patterns by Birds, with Especial Reference to the Song of the Chaffinch Fringilla Coelebs. *Ibis* **100,** 535–570 (1958).
20. Marler, P. & Tamura, M. Culturally Transmitted Patterns of Vocal Behavior in Sparrows. *Science* **146,** 1483–1486 (1964).
21. Price, P. H. Developmental determinants of structure in zebra finch song. *J. Comp. Physiol. Psychol.* **93,** 260–277 (1979).
22. Brenowitz, E. A., Margoliash, D. & Nordeen, K. W. An introduction to birdsong and the avian song system. *J. Neurobiol.* **33,** 495–500 (1997).
23. Williams, H. Birdsong and Singing Behavior. *Ann. N. Y. Acad. Sci.* **1016,** 1–30 (2004).
24. Boughey, M. J. & Thompson, N. S. Song Variety in the Brown Thrasher (Toxostoma rufum). *Z. Für Tierpsychol.* **56,** 47–58 (1981).

25. Mooney, R., Hoese, W. & Nowicki, S. Auditory representation of the vocal repertoire in a songbird with multiple song types. *Proc. Natl. Acad. Sci.* **98,** 12778–12783 (2001).
26. Beecher, M. D. & Brenowitz, E. A. Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* **20,** 143–149 (2005).
27. Brenowitz, E. A. & Beecher, M. D. Song learning in birds: diversity and plasticity, opportunities and challenges. *Trends Neurosci.* **28,** 127–132 (2005).
28. Chi, Z. & Margoliash, D. Temporal Precision and Temporal Drift in Brain and Behavior of Zebra Finch Song. *Neuron* **32,** 899–910 (2001).
29. Nottebohm, F., Stokes, T. M. & Leonard, C. M. Central control of song in the canary,Serinus canarius. *J. Comp. Neurol.* **165,** 457–486 (1976).
30. Nottebohm, F. & Nottebohm, M. E. Relationship between Song Repertoire and Age in the Canary, Serinus canarius. *Z. Für Tierpsychol.* **46,** 298–305 (1978).
31. Nottebohm, F., Nottebohm, M. E. & Crane, L. Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song-control nuclei. *Behav. Neural Biol.* **46,** 445–471 (1986).
32. Markowitz, J. E., Ivie, E., Kligler, L. & Gardner, T. J. Long-range Order in Canary Song. *PLoS Comput Biol* **9,** e1003052 (2013).
33. Okanoya, K. The Bengalese Finch: A Window on the Behavioral Neurobiology of Birdsong Syntax. *Ann. N. Y. Acad. Sci.* **1016,** 724–735 (2004).
34. Jin, D. Z. & Kozhevnikov, A. A. A Compact Statistical Model of the Song Syntax in Bengalese Finch. *PLoS Comput Biol* **7,** e1001108 (2011).
35. Honda, E. & Okanoya, K. Acoustical and Syntactical Comparisons between Songs of the White-backed Munia (Lonchura striata) and Its Domesticated Strain, the Bengalese Finch (Lonchura striata var. domestica)(Biochemistry). *Zoolog. Sci.* **16,** 319–326 (1999).
36. Warren, T. L., Charlesworth, J. D., Tumer, E. C. & Brainard, M. S. Variable Sequencing Is Actively Maintained in a Well Learned Motor Skill. *J. Neurosci.* **32,** 15414–15425 (2012).
37. Aronov, D., Andalman, A. S. & Fee, M. S. A Specialized Forebrain Circuit for Vocal Babbling in the Juvenile Songbird. *Science* **320,** 630–634 (2008).
38. Konishi, M. The Role of Auditory Feedback in the Control of Vocalization in the White-Crowned Sparrow1. *Z. Für Tierpsychol.* **22,** 770–783 (1965).
39. Sossinka, R. & Böhner, J. Song Types in the Zebra Finch Poephila guttata castanotis1. *Z. Für Tierpsychol.* **53,** 123–132 (1980).
40. Marler, P. & Peters, S. A Sensitive Period for Song Acquisition in the Song Sparrow, Melospiza melodia: A Case of Age-limited Learning. *Ethology* **76,** 89–100 (1987).
41. Payne, R. B. & Payne, L. L. Field observations, experimental design, and the time and place of learning bird songs. *Soc. Influ. Vocal Dev.* 57–84 (1997).
42. Lachlan, R. F. & Slater, P. J. B. Song learning by chaffinches: how accurate, and from where? *Anim. Behav.* **65,** 957–969 (2003).
43. Yu, A. C. & Margoliash, D. Temporal Hierarchical Control of Singing in Birds. *Science* **273,** 1871–1875 (1996).
44. Margoliash, D. Functional organization of forebrain pathways for song production and perception. *J. Neurobiol.* **33,** 671–693 (1997).
45. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature* **419,** 65–70 (2002).
46. Long, M. A. & Fee, M. S. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature* **456,** 189–194 (2008).
47. Clayton, N. S. Song tutor choice in zebra finches. *Anim. Behav.* **35,** 714–721 (1987).
48. Clayton, N. S. Song Learning in Cross-Fostered Zebra Finches: A Re-Examination of the Sensitive Phase. *Behaviour* **102,** 67–81 (1987).

49. Nottebohm, F., Paton, J. A. & Kelley, D. B. Connections of vocal control nuclei in the canary telencephalon. *J. Comp. Neurol.* **207,** 344–357 (1982).
50. Nottebohm, F. The Neural Basis of Birdsong. *PLoS Biol* **3,** e164 (2005).
51. Vicario, D. S. Organization of the zebra finch song control system: Functional organization of outputs from nucleus robustus archistriatalis. *J. Comp. Neurol.* **309,** 486–494 (1991).
52. Wild, J. M. Neural pathways for the control of birdsong production. *J. Neurobiol.* **33,** 653–670 (1997).
53. Vu, E. T., Mazurek, M. E. & Kuo, Y. C. Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neurosci.* **14,** 6924–6934 (1994).
54. Bottjer, S. W., Miesner, E. A. & Arnold, A. P. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* **224,** 901–903 (1984).
55. Sohrabji, F., Nordeen, E. J. & Nordeen, K. W. Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behav. Neural Biol.* **53,** 51–63 (1990).
56. Williams, H. & Mehta, N. Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. *J. Neurobiol.* **39,** 14–28 (1999).
57. Brainard, M. S. & Doupe, A. J. Interruption of a basal ganglia–forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404,** 762–766 (2000).
58. Ölveczky, B. P., Andalman, A. S. & Fee, M. S. Vocal Experimentation in the Juvenile Songbird Requires a Basal Ganglia Circuit. *PLoS Biol.* **3,** e153 (2005).
59. Wild, J. M. Descending projections of the songbird nucleus robustus archistriatalis. *J. Comp. Neurol.* **338,** 225–241 (1993).
60. Reinke, H. & Wild, J. m. Identification and connections of inspiratory premotor neurons in songbirds and budgerigar. *J. Comp. Neurol.* **391,** 147–163 (1998).
61. Striedter, G. F. & Vu, E. T. Bilateral feedback projections to the forebrain in the premotor network for singing in zebra finches. *J. Neurobiol.* **34,** 27–40 (1998).
62. Cardin, J. A., Raksin, J. N. & Schmidt, M. F. Sensorimotor Nucleus NIf Is Necessary for Auditory Processing But Not Vocal Motor Output in the Avian Song System. *J. Neurophysiol.* **93,** 2157–2166 (2005).
63. Coleman, M. J., Roy, A., Wild, J. M. & Mooney, R. Thalamic Gating of Auditory Responses in Telencephalic Song Control Nuclei. *J. Neurosci.* **27,** 10024–10036 (2007).
64. Akutagawa, E. & Konishi, M. Connections of thalamic modulatory centers to the vocal control system of the zebra finch. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 14086–14091 (2005).
65. Schmidt, M. F. Pattern of Interhemispheric Synchronization in HVc During Singing Correlates With Key Transitions in the Song Pattern. *J. Neurophysiol.* **90,** 3931–3949 (2003).
66. Schmidt, M. F., Ashmore, R. C. & Vu, E. T. Bilateral Control and Interhemispheric Coordination in the Avian Song Motor System. *Ann. N. Y. Acad. Sci.* **1016,** 171–186 (2004).
67. Mooney, R. & Prather, J. F. The HVC Microcircuit: The Synaptic Basis for Interactions between Song Motor and Vocal Plasticity Pathways. *J. Neurosci.* **25,** 1952–1964 (2005).
68. Kozhevnikov, A. A. & Fee, M. S. Singing-Related Activity of Identified HVC Neurons in the Zebra Finch. *J. Neurophysiol.* **97,** 4271–4283 (2007).
69. Leonardo, A. & Fee, M. S. Ensemble Coding of Vocal Control in Birdsong. *J. Neurosci.* **25,** 652–661 (2005).
70. Roberts, T. F., Klein, M. E., Kubke, M. F., Wild, J. M. & Mooney, R. Telencephalic Neurons Monosynaptically Link Brainstem and Forebrain Premotor Networks Necessary for Song. *J. Neurosci.* **28,** 3479–3489 (2008).

71. Fujimoto, H., Hasegawa, T. & Watanabe, D. Neural Coding of Syntactic Structure in Learned Vocalizations in the Songbird. *J. Neurosci.* **31,** 10023–10033 (2011).
72. Scharff, C. & Nottebohm, F. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J. Neurosci.* **11,** 2896–2913 (1991).
73. Kao, M. H. & Brainard, M. S. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J. Neurophysiol.* **96,** 1441–1455 (2006).
74. Stepanek, L. & Doupe, A. J. Activity in a Cortical-Basal Ganglia Circuit for Song Is Required for Social Context-Dependent Vocal Variability. *J. Neurophysiol.* **104,** 2474–2486 (2010).
75. Hampton, C. M., Sakata, J. T. & Brainard, M. S. An Avian Basal Ganglia-Forebrain Circuit Contributes Differentially to Syllable Versus Sequence Variability of Adult Bengalese Finch Song. *J. Neurophysiol.* **101,** 3235–3245 (2009).
76. Hamaguchi, K. & Mooney, R. Recurrent Interactions between the Input and Output of a Songbird Cortico-Basal Ganglia Pathway Are Implicated in Vocal Sequence Variability. *J. Neurosci.* **32,** 11671–11687 (2012).
77. Ashmore, R. C., Renk, J. A. & Schmidt, M. F. Bottom-Up Activation of the Vocal Motor Forebrain by the Respiratory Brainstem. *J. Neurosci.* **28,** 2613–2623 (2008).
78. Gibb, L., Gentner, T. Q. & Abarbanel, H. D. I. Brain Stem Feedback in a Computational Model of Birdsong Sequencing. *J. Neurophysiol.* **102,** 1763–1778 (2009).
79. Lewandowski, B., Vyssotski, A., Hahnloser, R. H. R. & Schmidt, M. At the interface of the auditory and vocal motor systems: NIf and its role in vocal processing, production and learning. *J. Physiol.-Paris* **107,** 178–192 (2013).
80. Nottebohm, F. Auditory Experience and Song Development in the Chaffinch Fringilla Coelebs. *Ibis* **110,** 549–568 (1968).
81. Kroodsma, D. E. & Konishi, M. A suboscine bird (eastern phoebe, Sayornis phoebe) develops normal song without auditory feedback. *Anim. Behav.* **42,** 477–487 (1991).
82. Nordeen, K. W. & Nordeen, E. J. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav. Neural Biol.* **57,** 58–66 (1992).
83. Okanoya, K. & Yamaguchi, A. Adult bengalese finches (Lonchura striata var. domestica) require real-time auditory feedback to produce normal song syntax. *J. Neurobiol.* **33,** 343–356 (1997).
84. Woolley, S. M. N. & Rubel, E. W. Bengalese Finches Lonchura Striata Domestica Depend upon Auditory Feedback for the Maintenance of Adult Song. *J. Neurosci.* **17,** 6380–6390 (1997).
85. Woolley, S. M. N. & Rubel, E. W. High-Frequency Auditory Feedback Is Not Required for Adult Song Maintenance in Bengalese Finches. *J. Neurosci.* **19,** 358–371 (1999).
86. Sakata, J. T. & Brainard, M. S. Real-Time Contributions of Auditory Feedback to Avian Vocal Motor Control. *J. Neurosci.* **26,** 9619–9628 (2006).
87. Sakata, J. T. & Brainard, M. S. Online Contributions of Auditory Feedback to Neural Activity in Avian Song Control Circuitry. *J. Neurosci.* **28,** 11378–11390 (2008).
88. Hosino, T. & Okanoya, K. Lesion of a higher-order song nucleus disrupts phrase level complexity in Bengalese finches. [Miscellaneous Article]. *Neuroreport July 14 2000* **11,** 2091–2095 (2000).
89. Skaggs, W. E., McNaughton, B. L., Wilson, M. A. & Barnes, C. A. Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* **6,** 149–172 (1996).
90. Buzsáki, G. & Draguhn, A. Neuronal Oscillations in Cortical Networks. *Science* **304,** 1926–1929 (2004).

91. Foster, D. J. & Wilson, M. A. Hippocampal theta sequences. *Hippocampus* **17,** 1093–1099 (2007).
92. Diesmann, M., Gewaltig, M.-O. & Aertsen, A. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **402,** 529–533 (1999).
93. Ikegaya, Y. *et al.* Synfire Chains and Cortical Songs: Temporal Modules of Cortical Activity. *Science* **304,** 559–564 (2004).
94. Li, M. & Greenside, H. Stable propagation of a burst through a one-dimensional homogeneous excitatory chain model of songbird nucleus HVC. *Phys. Rev. E* **74,** 011918 (2006).
95. Jin, D. Z., Ramazanoğlu, F. M. & Seung, H. S. Intrinsic bursting enhances the robustness of a neural network model of sequence generation by avian brain area HVC. *J. Comput. Neurosci.* **23,** 283–299 (2007).
96. Solis, M. M. & Perkel, D. J. Rhythmic Activity in a Forebrain Vocal Control Nucleus In Vitro. *J. Neurosci.* **25,** 2811–2822 (2005).
97. Long, M. A., Jin, D. Z. & Fee, M. S. Support for a synaptic chain model of neuronal sequence generation. *Nature* **468,** 394–399 (2010).
98. Jun, J. K. & Jin, D. Z. Development of Neural Circuitry for Precise Temporal Sequences through Spontaneous Activity, Axon Remodeling, and Synaptic Plasticity. *PLoS ONE* **2,** e723 (2007).
99. Katahira, K., Okanoya, K. & Okada, M. A neural network model for generating complex birdsong syntax. *Biol. Cybern.* **97,** 441–448 (2007).
100. Chang, W. & Jin, D. Z. Spike propagation in driven chain networks with dominant global inhibition. *Phys. Rev. E* **79,** 051917 (2009).
101. Jin, D. Z. Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. *Phys. Rev. E* **80,** 051902 (2009).
102. Yamashita, Y. *et al.* Developmental learning of complex syntactical song in the Bengalese finch: A neural network model. *Neural Netw.* **21,** 1224–1231 (2008).
103. Hanuschkin, A., Diesmann, M. & Morrison, A. A Reafferent and Feed-forward Model of Song Syntax Generation in the Bengalese Finch. *J Comput Neurosci* **31,** 509–532 (2011).
104. Patterson, M. M. *Bioelectric recording techniques*. **1,** (Academic Press, 1973).
105. Supèr, H. & Roelfsema, P. R. Chronic multiunit recordings in behaving animals: advantages and limitations. *Prog. Brain Res.* **147,** 263–282 (2005).
106. Denk, W., Strickler, J. H. & Webb, W. W. Two-photon laser scanning fluorescence microscopy. *Science* **248,** 73–76 (1990).
107. Svoboda, K. & Yasuda, R. Principles of Two-Photon Excitation Microscopy and Its Applications to Neuroscience. *Neuron* **50,** 823–839 (2006).
108. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat. Methods* **2,** 932–940 (2005).
109. Ogawa, S., Lee, T.-M., Nayak, A. S. & Glynn, P. Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Reson. Med.* **14,** 68–78 (1990).
110. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412,** 150–157 (2001).
111. Wilkes, K. V. More Brain Lesions. *Philosophy* **55,** 455–470 (1980).
112. Penfield, W. & Roberts, L. *Speech and brain mechanisms*. (Princeton University Press, 1959).
113. Ojemann, G., Ojemann, J., Lettich, E. & Berger, M. Cortical language localization in left, dominant hemisphere. *Spec. Suppl.* **112,** 316–326 (2009).

114. Choi, D. W., Koh, J. Y. & Peters, S. Pharmacology of glutamate neurotoxicity in cortical cell culture: attenuation by NMDA antagonists. *J. Neurosci.* **8,** 185–196 (1988).
115. Konopka, R. J. & Benzer, S. Clock Mutants of Drosophila melanogaster. *Proc. Natl. Acad. Sci.* **68,** 2112–2116 (1971).
116. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci.* **8,** 1263–1268 (2005).
117. Arbas, E. A. & Calabrese, R. L. Rate modification in the heartbeat central pattern generator of the medicinal leech. *J. Comp. Physiol. A* **155,** 783–794 (1984).
118. Thompson, S. M., Masukawa, L. M. & Prince, D. A. Temperature dependence of intrinsic membrane properties and synaptic potentials in hippocampal CA1 neurons in vitro. *J. Neurosci.* **5,** 817–824 (1985).
119. Volgushev, M., Vidyasagar, T. R., Chistiakova, M. & Eysel, U. T. Synaptic transmission in the neocortex during reversible cooling. *Neuroscience* **98,** 9–22 (2000).
120. Volgushev, M., Vidyasagar, T. R., Chistiakova, M., Yousef, T. & Eysel, U. T. Membrane properties and spike generation in rat visual cortical cells during reversible cooling. *J. Physiol.* **522,** 59–76 (2000).
121. Katz, P. S., Sakurai, A., Clemens, S. & Davis, D. Cycle period of a network oscillator is independent of membrane potential and spiking activity in individual central pattern generator neurons. *J. Neurophysiol.* **92,** 1904–1917
122. Andalman, A. S., Foerster, J. N. & Fee, M. S. Control of Vocal and Respiratory Patterns in Birdsong: Dissection of Forebrain and Brainstem Mechanisms Using Temperature. *PLoS ONE* **6,** e25461 (2011).
123. Aronov, D. & Fee, M. S. Analyzing the dynamics of brain circuits with temperature: Design and implementation of a miniature thermoelectric device. *J. Neurosci. Methods* **197,** 32–47 (2011).
124. Zucker, R. S. & Regehr, W. G. Short-Term Synaptic Plasticity. *Annu. Rev. Physiol.* **64,** 355–405 (2002).
125. Wang, C. Z. H., Herbst, J. A., Keller, G. B. & Hahnloser, R. H. R. Rapid Interhemispheric Switching during Vocal Production in a Songbird. *PLoS Biol* **6,** e250 (2008).
126. Secora, K. R. *et al.* Syringeal Specialization of Frequency Control during Song Production in the Bengalese Finch (Lonchura striata domestica). *PLoS ONE* **7,** e34135 (2012).
127. Chomsky, N. *Aspects of the Theory of Syntax*. (MIT Press, 1965).
128. Patel, A. D. Language, music, syntax and the brain. *Nat. Neurosci.* **6,** 674–681 (2003).
129. Berwick, R. C., Okanoya, K., Beckers, G. J. L. & Bolhuis, J. J. Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* **15,** 113–121 (2011).
130. Kershenbaum, A., Ilany, A., Blaustein, L. & Geffen, E. Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proc. R. Soc. B Biol. Sci.* **279,** 2974–2981 (2012).
131. Cromwell, H. C. & Berridge, K. C. Implementation of Action Sequences by a Neostriatal Site: A Lesion Mapping Study of Grooming Syntax. *J. Neurosci.* **16,** 3444–3458 (1996).
132. Fedorenko, E., Nieto-Castañón, A. & Kanwisher, N. Syntactic processing in the human brain: What we know, what we don't know, and a suggestion for how to proceed. *Brain Lang.* **120,** 187–207 (2012).
133. Indefrey, P. *et al.* A neural correlate of syntactic encoding during speech production. *Proc. Natl. Acad. Sci.* **98,** 5933–5936 (2001).
134. Broca, P. Remarks on the seat of the faculty of articulate language, followed by an observation of aphemia.
135. Borovsky, A., Saygin, A. P., Bates, E. & Dronkers, N. Lesion correlates of conversational speech production deficits. *Neuropsychologia* **45,** 2525–2533 (2007).

136. Suthers, R. A. & Margoliash, D. Motor control of birdsong. *Curr. Opin. Neurobiol.* **12,** 684–690 (2002).
137. Mooney, R. Neural mechanisms for learned birdsong. *Learn. Mem.* **16,** 655–669 (2009).
138. Doupe, A. J., Perkel, D. J., Reiner, A. & Stern, E. A. Birdbrains could teach basal ganglia research a new song. *Trends Neurosci.* **28,** 353–363 (2005).
139. Lipkind, D. & Tchernichovski, O. Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. *Proc. Natl. Acad. Sci.* **108,** 15572–15579 (2011).
140. Sakata, J. T. & Brainard, M. S. Social Context Rapidly Modulates the Influence of Auditory Feedback on Avian Vocal Motor Control. *J. Neurophysiol.* **102,** 2485–2497 (2009).
141. Sakata, J. T., Hampton, C. M. & Brainard, M. S. Social Modulation of Sequence and Syllable Variability in Adult Birdsong. *J. Neurophysiol.* **99,** 1700–1711 (2008).
142. Yamada, H. & Okanoya, K. Song syntax changes in Bengalese finches singing in a helium atmosphere. *Neuroreport* **14,** 1725–1729 (2003).
143. Warren, T. L., Tumer, E. C., Charlesworth, J. D. & Brainard, M. S. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J. Neurophysiol.* **106,** 1806–1821 (2011).
144. Wohlgemuth, M. J., Sober, S. J. & Brainard, M. S. Linked Control of Syllable Sequence and Phonology in Birdsong. *J. Neurosci.* **30,** 12936–12949 (2010).
145. Fisher, R. A. On the Interpretation of $\chi 2$ from Contingency Tables, and the Calculation of P. *J. R. Stat. Soc.* **85,** 87–94 (1922).
146. Wallenstein, S. & Wittes, J. The Power of the Mantel-Haenszel Test for Grouped Failure Time Data. *Biometrics* **49,** 1077–1087 (1993).
147. Katahira, K., Suzuki, K., Okanoya, K. & Okada, M. Complex Sequencing Rules of Birdsong Can be Explained by Simple Hidden Markov Processes. *PLoS ONE* **6,** e24516 (2011).
148. Marler, P. & Sherman, V. Song structure without auditory feedback: emendations of the auditory template hypothesis. *J. Neurosci.* **3,** 517–531 (1983).
149. Bottjer, S. W. & Arnold, A. P. The role of feedback from the vocal organ. I. Maintenance of stereotypical vocalizations by adult zebra finches. *J. Neurosci.* **4,** 2387–2396 (1984).
150. Schmidt, M. F. & Konishi, M. Gating of auditory responses in the vocal control system of awake songbirds. *Nat. Neurosci.* **1,** 513–518 (1998).
151. Leonardo, A. Experimental test of the birdsong error-correction model. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 16935–16940 (2004).
152. Konishi, M. The Role of Auditory Feedback in Birdsong. *Ann. N. Y. Acad. Sci.* **1016,** 463–475 (2004).
153. Prather, J. F., Peters, S., Nowicki, S. & Mooney, R. Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature* **451,** 305–310 (2008).
154. Margoliash, D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci.* **3,** 1039–1057 (1983).
155. Margoliash, D. & Konishi, M. Auditory representation of autogenous song in the song system of white-crowned sparrows. *Proc. Natl. Acad. Sci.* **82,** 5997–6000 (1985).
156. Margoliash, D. Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *J. Neurosci.* **6,** 1643–1661 (1986).
157. Sakata, J. T. & Brainard, M. S. Online Contributions of Auditory Feedback to Neural Activity in Avian Song Control Circuitry. *J. Neurosci.* **28,** 11378–11390 (2008).
158. Leonardo, A. & Konishi, M. Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* **399,** 466–470 (1999).
159. Skocik, M. & Kozhevnikov, A. Real-time system for studies of the effects of acoustic feedback on animal vocalizations. *Front. Neural Circuits* **6,** (2013).

160. Paton, J. A., Manogue, K. R. & Nottebohm, F. Bilateral organization of the vocal control pathway in the budgerigar, Melopsittacus undulatus. *J. Neurosci.* **1,** 1279–1288 (1981).
161. Ocklenburg, S., Ströckens, F. & Güntürkün, O. Lateralisation of conspecific vocalisation in non-human vertebrates. *Laterality Asymmetries Body Brain Cogn.* **18,** 1–31 (2013).
162. Laje, R. & Mindlin, G. B. *The Physics of Birdsong*. (Springer-Verlag Berlin Heidelberg, 2005).
163. Vu, E. T., Schmidt, M. F. & Mazurek, M. E. Interhemispheric Coordination of Premotor Neural Activity during Singing in Adult Zebra Finches. *J. Neurosci.* **18,** 9088–9098 (1998).
164. Trevisan, M. A., Cooper, B., Goller, F. & Mindlin, G. B. Lateralization as a symmetry breaking process in birdsong. *Phys. Rev. E* **75,** 031908 (2007).
165. Rasmussen, T. & Milner, B. The Role of Early Left-Brain Injury in Determining Lateralization of Cerebral Speech Functions. *Ann. N. Y. Acad. Sci.* **299,** 355–369 (1977).
166. Nottebohm, F. & Nottebohm, M. E. Left hypoglossal dominance in the control of canary and white-crowned sparrow song. *J. Comp. Physiol.* **108,** 171–192 (1976).
167. Halle, F., Gahr, M. & Kreutzer, M. Effects of unilateral lesions of HVC on song patterns of male domesticated canaries. *J. Neurobiol.* **56,** 303–314 (2003).
168. Okanoya, K. & Watanabe, S. Asymmetric effects of left-and right-HVC lesion on song perception and production in Bengalese finches. in *Nerv. Syst. Behav. Proc. 4th Int. Congr. Neuroethol.* 323 (Thieme Stuttgart,, Germany, 1994).
169. Okanoya, K., Ikebuchi, M., Uno, H. & Watanabe, S. Left-side dominance for song discrimination in Bengalese finches (Lonchura striata var. domestica). *Anim. Cogn.* **4,** 241–245 (2001).
170. Swadlow, H. A. Neocortical efferent neurons with very slowly conducting axons: strategies for reliable antidromic identification. *J. Neurosci. Methods* **79,** 131–141 (1998).
171. Houweling, A. R. & Brecht, M. Behavioural report of single neuron stimulation in somatosensory cortex. *Nature* **451,** 65–68 (2008).
172. Olds, M. E. & Fobes, J. L. The Central Basis of Motivation: Intracranial Self-Stimulation Studies. *Annu. Rev. Psychol.* **32,** 523–574 (1981).
173. Perlmutter, J. S. & Mink, J. W. Deep Brain Stimulation. *Annu. Rev. Neurosci.* **29,** 229–257 (2006).
174. Tarsy, D., Vitek, J. L., Starr, P. & Okun, M. *Deep brain stimulation in neurological and psychiatric disorders*. (Springer, 2008).
175. Wise, K. D., Anderson, D. J., Hetke, J. F., Kipke, D. R. & Najafi, K. Wireless implantable microsystems: high-density electronic interfaces to the nervous system. *Proc. IEEE* **92,** 76–97 (2004).
176. Peng, C.-W. *et al.* High frequency block of selected axons using an implantable microstimulator. *J. Neurosci. Methods* **134,** 81–90 (2004).
177. Pinkwart, C. & Borchers, H.-W. Miniature three-function transmitting system for single neuron recording, wireless brain stimulation and marking. *J. Neurosci. Methods* **20,** 341–352 (1987).
178. Millard, R. E. & Shepherd, R. K. A fully implantable stimulator for use in small laboratory animals. *J. Neurosci. Methods* **166,** 168–177 (2007).
179. Ativanichayaphong, T., He, J. W., Hagains, C. E., Peng, Y. B. & Chiao, J.-C. A combined wireless neural stimulating and recording system for study of pain processing. *J. Neurosci. Methods* **170,** 25–34 (2008).
180. Arfin, S. K., Long, M. A., Fee, M. S. & Sarpeshkar, R. Wireless Neural Stimulation in Freely Behaving Small Animals. *J. Neurophysiol.* **102,** 598–605 (2009).

181. Ghovanloo, M. & Najafi, K. A Wireless Implantable Multichannel Microstimulating System-on-a-Chip With Modular Architecture. *IEEE Trans. Neural Syst. Rehabil. Eng.* **15,** 449–457 (2007).
182. Ye, X. *et al.* A portable telemetry system for brain stimulation and neuronal activity recording in freely behaving small animals. *J. Neurosci. Methods* **174,** 186–193 (2008).
183. Xu, S., Talwar, S. K., Hawley, E. S., Li, L. & Chapin, J. K. A multi-channel telemetry system for brain microstimulation in freely roaming animals. *J. Neurosci. Methods* **133,** 57–63 (2004).
184. Yeomans, J. S. *Principles of brain stimulation*. (Oxford University Press New York, 1990).
185. Tehovnik, E. J. Electrical stimulation of neural tissue to evoke behavioral responses. *J. Neurosci. Methods* **65,** 1–17 (1996).
186. Palumbo, G. & Pappalardo, D. Charge Pump Circuits: An Overview on Design Strategies and Topologies. *IEEE Circuits Syst. Mag.* **10,** 31–45 (2010).

VITA

Yisi Zhang

Yisi Zhang was born in Beijing, People's Republic of China on January 6th, 1986. She entered Yuzhong Primary School in Xicheng District, Beijing from 1992 to 1998. From the September of 1998, she started her junior high and high school study at Beijing No. 2 Middle School in Dongcheng District. She received her Bachelor of Science from Tsinghua University in 2008 after accomplishing a four-year program in Physics. In the summer of 2008, she came to Penn State University to pursue a Ph.D. in Physics. She worked with Professor Alexay Kozhevnikov to study the neural mechanisms of songbird singing till this thesis was written.