Neural Dynamics
Underlying Complex Behavior in a Songbird

Thesis by

Anthony Leonardo

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Computation and Neural Systems Program

California Institute of Technology

Pasadena, California

2002

(Defended 2 May 2002)

Abstract

Zebra finches memorize a tutor song as juveniles, and then faithfully reproduce this song for the remainder of their lives. This thesis investigates how the intricate and concerted activity of ensembles of neurons in the zebra finch brain gives rise to this complex behavior.

The first section of the thesis demonstrates how neurons in RA, a pre-motor song control nucleus, coordinate their activity to generate the instantaneous spectral and temporal structure seen in the song. Similar sounds are produced by entirely different ensembles of RA neurons. Additionally, rapidly changing patterns of RA neural activity are capable of generating constant acoustic outputs. This is a new form of neural coding, previously unobserved in any other system. This unusual neural representation provides a concise account of an unexplained aspect of song learning.

The second part of the thesis develops a new, noninvasive method of altering the auditory feedback heard by singing zebra finches. This type of feedback perturbation causes the normally stable songs of adult birds to deteriorate, indicating that auditory feedback is essential to song control throughout life. Restoration of the normal auditory feedback heard by these birds enables the gradual recovery of their original songs, demonstrating that a memory of the original song persists in the song control system despite the destabilization of the behavior. The song is thus maintained by an active feedback control process.

The final project investigates the role of nucleus LMAN in real-time song control via auditory feedback. Neurons in nucleus LMAN were previously believed to compare auditory feedback to a memorized song template. The result of this comparison was thought to be used as an error-correction signal to update the motor control program. By recording from individual LMAN neurons while simultaneously manipulating the auditory feedback heard by the singing bird, it is shown that during singing, rather than auditory feedback, LMAN processes an efference copy of the motor commands used to generate the bird's song. This suggests a different model of the system, in which the efference copy is the basis of the real-time error-correction signal.

Taken together, these three projects significantly expand our knowledge of the neural mechanisms responsible for the generation of birdsong.

Acknowledgements

Science is a collective enterprise, and the research contained in this thesis could not have succeeded without the help, advice, and encouragement of many people. Above all, I am indebted to my advisor, Mark Konishi. Mark provided exactly the right balance of guidance and freedom that was essential for my growth as a scientist. His uncanny ability to see to the heart of a problem and distill it into its essential elements is a quality that I have striven to emulate. I must further thank my undergraduate advisor, Herb Simon, who provided my initial training as a cognitive scientist, and then encouraged me to pursue my interests in the links between neurons and behavior at Caltech.

The work contained in Chapter 2 was done with Michale Fee at Bell Labs, in one of the most fun and rewarding collaborative science experiences I have ever had. My committee members, Gilles Laurent, Jerry Pine, John Allman, and Pietro Perona, provided me with sound advice and asked many great questions which led to improvements in this thesis. The members of the Konishi Lab, especially Roian Egnor, Ben Arthur, Chris Malek, and Teresa Nick, were a continuous source of excellent feedback and discussions. Kourosh Saberi introduced me to Golay codes, which proved critical for implementing the active sound cancellation system used in Chapter 4. Gene Akutagawa taught me the basics of histology, and Marc Schmidt provided my initial education on zebra finch electrophysiology. Partha Mitra and Bijan Pesaran introduced me to the tools and techniques of spectral analysis. Mike Walsh and Herb Adams provided vital assistance during the design of various instruments used in the different experiments. Mark Konishi, Roian Egnor, and the institute proofreader read through the chapters of this thesis and alerted me to numerous inconsistencies and points which needed further clarification.

During my visits to Harvard, Marc Hauser graciously provided me with office space and a fun place to work. Hideo Mabuchi introduced me to the wonderful world of Eiswein, and then went above and beyond the call of duty by allowing me to live in his house for 3 months while much of the analysis in this thesis was completed. Hideo, Paul Rothemund, Reid Harrison, Fidel Santamaria, Michael Gibson, and many others made Caltech and the Computation and Neural Systems program a remarkably exciting and fulfilling place to do science.

My parents, Martin and Bernadette Leonardo, my sister Teresa, and my good friend Alex Hawkinson have been a source of unwavering support through out this journey.

Finally, for Roian, there are not enough words on all the pages ever written; thank you, for making me happy.

Contents

Abstract iii

Acknowledgements iv

1. Introduction : Neural dynamics underlying complex behavior in a songbird 1
   - 1.1 A brief history of birdsong 3
   - 1.2 Summary of work 7

2. Neuronal ensemble coding of birdsong 10
   - 2.1 Introduction..... 10
   - 2.2 Results..... 12
   - 2.3 Discussion..... 22
   - 2.4 Methods..... 25

3. Decrystallization of adult birdsong by perturbation of auditory feedback 29
   - 3.1 Introduction..... 29
   - 3.2 Results..... 32
   - 3.3 Discussion..... 39
   - 3.4 Methods..... 40

4. An efference copy may be used to maintain the stability of adult birdsong 45
   - 4.1 Introduction..... 45
   - 4.2 Results..... 51
   - 4.3 Discussion..... 61
   - 4.4 Methods..... 63

Appendix 1 : Miniature motorized microdrive construction protocol 71

References 82

List of Figures

1.1 A zebra finch and his song..... 3
1.2 Organization of the song control system..... 5

2.1 Miniature motorized microdrive for chronic recording in small animals..... 12
2.2 Simultaneous recording of three RA neurons in the singing bird..... 13
2.3 Inter-spike-interval and burst-width distribution for RA neurons..... 14
2.4 Effect of acoustic time-warping on song motifs and neural recordings..... 15
2.5 RA neuronal ensemble activity during singing (35 cells)..... 16
2.6 Neural and song correlation matrices..... 17
2.7 Neural ensemble auto-correlation..... 18
2.8 Conditional probability distribution for the neural correlation matrix..... 19
2.9 Rapidly and slowly changing sounds in zebra finch song..... 21
2.10 Constant acoustic structure is produced by rapidly changing neural ensembles.. 21

3.1 The normal structure of adult birdsong..... 30
3.2 Diagram of auditory feedback perturbation system..... 31
3.3 Stuttering of syllables in decrystallized birdsong..... 33
3.4 Addition and deletion of syllables in decrystallized birdsong..... 34
3.5 Spectral distortion in decrystallized song syllables..... 35
3.6 Decrystallized songs contain increased variability in syllable sequencing..... 36
3.7 Localized feedback perturbation produces localized song changes..... 37
3.8 Time course of decrystallization and song recovery..... 38

4.1 Zebra finch song control system and error-correction model..... 46
4.2 Song selectivity of anesthetized LMAN neurons..... 48
4.3 Histological reconstruction of motordrive recording sites in LMAN..... 49
4.4 Auditory feedback perturbation system and active sound cancellation system.... 50

4.5 Recording from a single LMAN neuron during two motifs of singing..... 51
4.6 Song and spike train alignment for two LMAN neurons..... 52
4.7 LMAN neurons fire spikes precisely timed to the bird's son..... 53
4.8 Time-varying Kullback-Leibler information for an LMAN neuron..... 57
4.9 RA neuron activity during normal and perturbed auditory feedback singing..... 58
4.10 LMAN neural activity during directed and undirected song..... 60

A.1 Overview of motorized microdrive..... 74
A.2 Photographs of microdrive subassemblies..... 75
A.3 Motor connections to Sutter MP-285 microdrive controller..... 80

The truth shall make you free

Introduction: Neural dynamics underlying complex behavior in a songbird

--from the Edwin Smith Surgical Papyrus, (Egypt, 1700 BC)
the oldest written record of the word "brain"

The brain. It has fascinated man throughout history, but its secrets have remained largely untapped, and its mysteries unsolved. From 1700 BC to 1500 AD, the first written records about the nervous system (Edwin Smith Surgical Papyrus, ancient Egypt) were expanded into a coarse description of its basic anatomy (e.g., 1561 AD, publication of the first description of the cranial nerves, G. Falloppio). Despite the many accomplishments over this span of 3000 years, the tools needed to begin understanding the physical mechanisms that underlie thoughts and behaviors would require nearly another 500 years of development. It was not until almost the twentieth century that nerve cells were discovered to be independent elements (Ramon y Cajal, 1889), the fundamental computational units of brain function. Progress since then has been rapid, and our understanding of the brain has advanced tremendously. Yet the core questions that have haunted us from the beginning remain unanswered; the neural mechanisms by which thoughts and behaviors emerge are nearly as opaque to us as they were 3000 years ago. The brain remains a riddle that will never lose its luster, because it represents everything that we are, and all that we can ever become.

After the discovery of independent nerve cells by Cajal, much of neuroscience research in the first half of the twentieth century was characterized by studies of the electrical properties of neurons in anesthetized animals. The culmination of this work was the collective realization, made by many researchers and in a number of species, that neurons could be thought of as detectors of behaviorally relevant stimuli. It is worthwhile to consider

two representative examples which clarify why these discoveries revolutionized the study of brain and behavior. Hubel and Wiesel (1959) found that neurons in the primary visual cortex of the anesthetized cat discharged maximally when presented with edges of a particular orientation. They further realized that the complex receptive fields of these neurons could be accounted for as a sum of the simpler receptive fields of neurons in the lateral geniculate nucleus, which sends input to the primary visual cortex. At nearly the same time, Lettvin et al. (1959) showed that visual ganglion cells in the frog's retina were tuned to small dark objects moving intermittently in their receptive fields – they were effectively bug-detectors. The implications of these discoveries were staggering: the visual scene was being decomposed, sliced into a series of smaller pieces which could be manipulated, transformed, and used for computation by the rest of the brain.

Remarkable observations such as those described in the previous paragraph demonstrated that the information processing capabilities of the brain could be measured, understood, and correlated with behavior. Over time, however, it became clear that the response properties of neurons in behaving animals were often dramatically different and more complex than those in anesthetized animals. Furthermore, behaviors are generated not by individual cells, but rather by the activity of large numbers of interacting neurons. These collective network dynamics are not easily understood when considering the activity of single neurons. In the last twenty years, the measurement of this network activity and its relation to behavior has become of a problem of paramount importance in systems neuroscience. This shift, from the study of single cells in anesthetized animals to many cells in behaving animals, has brought with it a number of challenges. First, technology is needed to allow the measurement of the activity of large ensembles of neurons in a manner that is robust to the movements of a freely behaving animal. Second, techniques are needed to observe the behavior of an animal, to manipulate that behavior in a controlled manner, and to compensate for behavioral variability. Finally, analytical methods are required to quantify the dynamic activity of many neurons, and to correlate this activity with the behavior of the animal.

The work described in this thesis has touched on all of these problems, but in particular has focused on the question of how the activity of large numbers of interacting neurons is involved in the generation of behavior. The animal in which I have done this

research is the zebra finch (*Taeniopygia guttata*), a small Australian songbird (Figure 1.1). As juveniles, these birds memorize a tutor's song and produce an accurate copy of it by the time they reach adulthood (Immelman, 1969; Price, 1979). Once learned, a bird will sing only this song for the remainder of his life. The song contains spectral and temporal structure over a wide range of time scales and is thus a complex learned behavior. It may at first seem misleading, to hope that the lessons learned from the brain of a bird can be generalized in any manner to the brain of a man. But the differences in complexity between humans and bird are no more than that – differences, variations on a common theme. Evolution has generated a single class of neural information processing systems, and while the details of each individual system are unique, the general principles that govern their function appear to be the same across countless species, from insects to primates. Thus by studying the neural control system underlying birdsong, one hopes to gain insight into the general types of neural dynamics evolved by the brain to produce complex behaviors. In the sections which follow, I briefly review the relevant behavioral and anatomical literature on the generation of birdsong, and then summarize the progress I have made towards understanding the neural dynamics underlying complex behavior in a songbird.

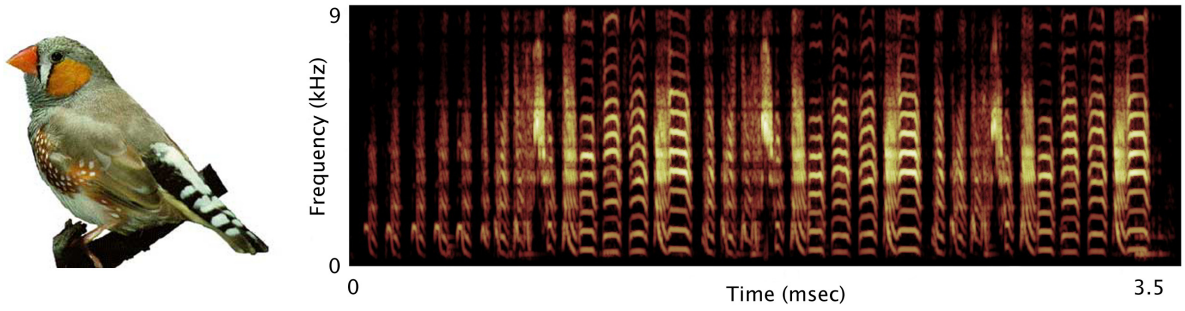

Figure 1.1. A zebra finch and his song.

1.1 A brief history of birdsong

Zebra finches are small, weighing only 10-15 grams; the zebra finch brain is on the order of 500 mm^3. Of this, the combined volume of all of the nuclei responsible for generating birdsong is probably less than 20 mm^3. The individual brain areas studied in this thesis range in size from 400 um to 700 um spheres, each containing a few thousand neurons which are roughly 10 um in diameter. The total number of neurons in the song control system is

probably less than 500,000, and is certainly less than 1,000,000. Yet despite these seemingly formidable size constraints, the zebra finch does something remarkable. When he is very young, only 20-30 days old, a male zebra finch will memorize the song of a tutor. The song is a complex pattern of different sounds, arranged in a sequence of about 1 second in length (Immelman, 1969). At around day 45 he begins to sing, first producing scratchy, ill-formed squawks, and then later gradually refining them into a nearly perfect copy of the tutor song. Finally, at roughly 90 days of age, the learning process is complete and the bird will sing only this single crystallized song for the remainder of his life. This sophisticated behavior, central in the life of zebra finches as a courtship mechanism and a mode of territorial defense, is generated by a tiny collection of neural circuits, a biological machine so small it can only be seen with a microscope.

Equally remarkable is how the behavior responds to perturbations of the normal song learning process. Song has been known to be a learned behavior, dependent on hearing the songs of adult conspecifics, since the work of Thorpe (1958), who showed that young birds isolated from tutors develop highly abnormal songs. Interestingly, the songs of isolate birds do contain some recognizable species-specific components, most notably the number and length of individual sound units (syllables), and the separation between them (Marler and Sherman, 1985). Thus some species-typical components of the song are innate. In contrast, the complex modulations of spectral structure, and the stereotyped temporal patterning of different sounds are all components of the song that must be learned. The amount of tutoring required for memorization of the song can be incredibly brief. In some species, hearing the tutor song approximately 10 times is sufficient for accurate reproduction when the bird reaches adulthood (Todt and Hultsch, 1985, Konishi, 1985). Immelman (1969) showed that birds tutored by males of a different species will still attempt to copy the song of the tutor, and will generally be fairly successful at this, despite the unusual sounds that they are generating (see also Price, 1979). After the template song has been memorized, but before the young bird begins to sing, the tutor may be removed without any effect on quality of song eventually produced by the adult bird (Marler, 1976).

Auditory input is critical to song learning in two ways. First, as discussed above, it is vital for the process of memorizing the tutor's song.. The second role of audition in song learning was first demonstrated by Konishi (1965), who showed that birds deafened after

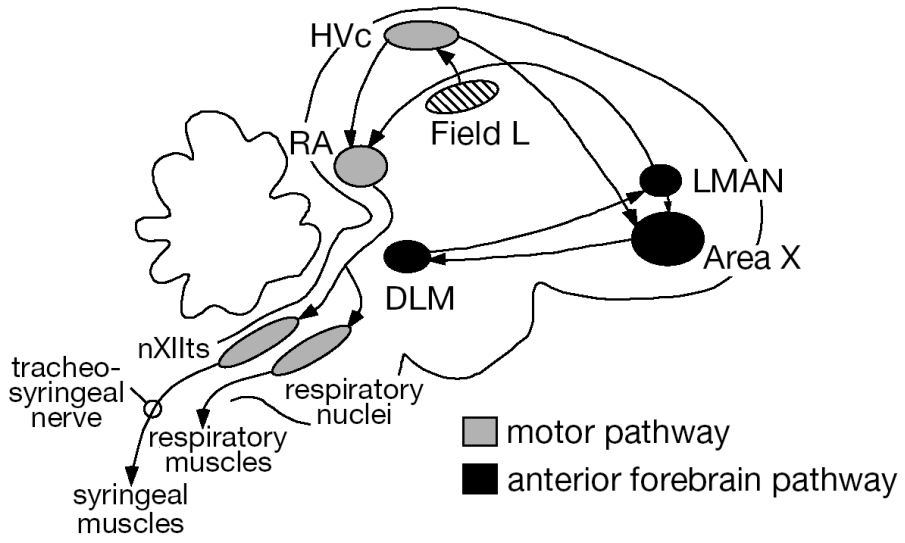

Figure 1.2. Schematic diagram of the zebra finch brain. The motor control pathway, from HVc to the syrinx, is shown in gray. The anterior forebrain pathway is shown in black. *Figure kindly provided by Allison Doupe, UCSF.*

they memorized a tutor song but before they began to sing developed highly abnormal songs and never successfully reproduced the tutor song. Auditory feedback therefore plays a second role in learning, during the sensorimotor process of transforming the bird's initially noisy vocalizations into an accurate copy of the tutor song. It was proposed very early on that in this second stage of learning, auditory feedback is used in an error-correction process to refine the bird's own vocalizations into a progressively more accurate copy of the tutor song. Attempts to identify the neural locus of the song template, and the neural circuits whereby the error-correction process takes place, continue to this day and form a significant part of this thesis.

The zebra finch has become a model system for studying of the neural basis of natural behaviors (i.e., neuroethology), because the richness of the singing behavior is generated by a discrete set of brain nuclei which are amenable to neurophysiological investigation. Nottebohm et al. (1976) first identified the major brain nuclei responsible for the control of song. HVc (the high vocal center) sends a projection to RA (robust nucleus of the archistriatum), which in turn projects to the tracheosyringeal portion of the hypoglossal nucleus, and from there to the muscles of the syrinx, the vocal organ (see Figure 3). Subsequent work by Vicario (1991) demonstrated that the pathway from RA to the syrinx is

myotopically organized, with different subregions of RA innervating different syringeal muscles. There are also additional projections from RA to the midbrain, and from there to hindbrain respiratory centers which are also intimately involved in the production of song (Wild, 1993). The syrinx is composed of several pairs muscles, which are intricately controlled during song generation (Suthers, 1990; Goller and Suthers, 1996). There are further intrinsic nonlinearities to the syrinx, such that simple variations in a few muscular control parameters can produce highly complex sounds (Fee et al., 1998). Lesions of HVc or RA result in immediate loss of singing ability or in substantial degradation of song structure (Nottebohm et al., 1976). This first song control pathway, from HVc through RA and to the vocal organ, is thus involved in the production of song.

Nottebohm et al. (1976) also observed a second pathway from HVc to Area X in the anterior neostriatum. Work by Arnold et al. (1976), Bottjer et al. (1989), and others demonstrated that this anterior forebrain pathway traveled from Area X, to the thalamic nucleus DLM, and then back to the anterior neostriatum to nucleus LMAN (lateral magnocellular nucleus of the anterior neostriatum). LMAN projects directly to RA, returning to the original song control pathway described by Nottebohm. A feedback pathway further links LMAN back to Area X (Vates and Nottebohm, 1995). Lesions of any of the nuclei in the anterior forebrain pathway in the juvenile bird have devastating effects on song learning (Scharff and Nottebohm, 1991). Lesions of these nuclei in the adult bird have few consequences for song production (Scharff and Nottebohm, 1991; Bottjer et al., 1984, Nordeen and Nordeen, 1993) unless a further manipulation is done which artificially induces the song to change. In these cases, lesions of nucleus LMAN prevent song deterioration (Williams and Mehta, 1999; Brainard and Doupe, 2000b). Finally, LMAN has been found to make NMDA synapses onto RA neurons, and disruption of these connections during song learning by infusing LMAN or RA with NMDA blockers results in the bird developing a highly abnormal song (Basham et al., 1996; Aamodt et al., 1996). Based on these observations, the second song control pathway, from HVc through the anterior forebrain and back to nucleus RA, is thus involved in song learning and song maintenance.

1.2 Summary of work

This thesis is composed of three major chapters; the first is concerned with the neural mechanisms underlying the generation of song structure, and the subsequent two with use of feedback by the neural control system to maintain song stability.

Chapter 2

In many models of neural population coding, similar sensory or motor states are represented in the brain by similar neuronal ensembles (Georgopoulos et al., 1999; Lewis and Kristan, 1998; Wilson and McNaughton, 1993). In Chapter 2, I explored this issue by measuring the activity of large numbers of single neurons in the pre-motor nucleus RA of the singing zebra finch. During singing, individual RA neurons generate precise bursts of action potential spikes (Yu and Margoliash, 1996). The active set of RA neurons at any time point in the song constitutes a neural ensemble. I found that highly similar song elements were typically produced by stereotyped but uncorrelated ensembles of RA neurons. Small changes in acoustic structure between two syllables were frequently subserved by entirely different patterns of RA neural activity, despite the strong correlations that are known to be present in the vocal muscles during the production of similar sounds. This provides an explanation for the behavioral observation that zebra finches learn syllable structure and sequence simultaneously (Tchernikovski et al, 2001), instead of first learning syllable primitives and then syllable order. There is no neural representation of syllable primitives in RA, such that similar sounds are generated by similar sets of neurons. Furthermore, the neural activity in RA often changed on a time scale an order of magnitude faster than the acoustic structure in the song, causing constant acoustic outputs to be produced by sequences of rapidly changing neural ensembles. Finally, the addition of neurons to the RA neuronal ensemble did not yield a steady increase in predictability about what type of sound was being generated. Most models of neural population coding follow a sum-of-the-parts principle, in which the individual neurons provide considerable information about what is being represented, and additional neurons provide progressively more information. The neural code used within RA is substantially different from these models, and individual neurons provide little unambiguous information about song structure. These observations suggest that within the neural control system of RA, network interactions are the dominant feature, and it is only by

reading out sufficient activity to represent the entire state of the network that decoding is possible. The spike patterns observed in RA thus represent a new type of neural code underlying complex learned behaviors.

Chapter 3

Young birds learn to sing by using auditory feedback to compare their own vocalizations to a memorized or innate song pattern; if they are deafened as juveniles they will not develop normal songs (Konishi, 1965; Nottebohm, 1968). The completion of song development is called crystallization. After this stage, song shows little variation in its temporal or spectral properties. However, the mechanisms underlying this stability are largely unknown. In Chapter 3, I present evidence that auditory feedback is actively used in adulthood to maintain the stability of song structure. I found that perturbing auditory feedback heard during singing in adult zebra finches caused a slow deterioration of song. This "decrystallization" consisted of a dramatic loss of the spectral and temporal stereotypy seen in crystallized song, and included stuttering, creation, deletion and distortion of song syllables. After normal feedback was restored, these deviations gradually disappeared and the original song was recovered, indicating that a memory of the original song pattern remained stable, despite the destabilization of the behavior. Thus, the brains of adult birds that do not learn new songs nevertheless retain a significant amount of plasticity. Despite its seeming stability, song is maintained throughout the bird's life by an active control process, and somewhere in the song control system there must be neurons that are sensitive to perturbations of auditory feedback.

Chapter 4

Nucleus LMAN of the anterior neostriatum is believed to be crucial for the auditory feedback-based plasticity seen in birdsong, and is thought to convey an error-correction signal to the motor control system based on the degree of match between the bird's vocalizations and a memorized song template (Brainard and Doupe, 2000a). In Chapter 4, I measured the activity of individual LMAN neurons while simultaneously manipulating the auditory feedback that birds heard during singing, thus controlling the level of error they could detect in their songs. LMAN neurons were found to produce spikes locked with

millisecond precision to specific acoustic features in individual song syllables. This timing precision is comparable to that seen the motor control neurons that generate the song itself (Chi and Margoliash, 2001). Furthermore, using information theoretic methods, I show that perturbation of the auditory feedback heard by singing birds had no effect on LMAN spike patterns. There are four possible signals that could produce the song-locked spike trains observed in LMAN: an auditory feedback signal, a proprioceptive feedback signal from the vocal muscles, a motor command signal, or an efference copy. An efference copy is a record of the commands used to generate a motor output (Sperry, 1950; von Holst and Mittelstaedt, 1950; Bridgeman, 1995). Our results clearly show that auditory feedback does not generate the spike patterns in LMAN. There is no known afference from the syringeal muscles back to the song control system (Bottjer and Arnold, 1984), eliminating the possibility that LMAN spike patterns were driven by proprioceptive feedback. LMAN can be lesioned without effecting the production of song (Bottjer et al., 1984), indicating that it does not generate a motor command. Therefore, the only hypothesis that is consistent our measurements is that LMAN activity during singing is driven by an efference copy of the bird's song. No other signal could cause LMAN activity to have such a high degree of correlation to song structure while being immune to changes in auditory feedback. This conclusion represents a significant departure from the classical view of LMAN being primarily a processor of auditory feedback. The neural mechanisms of song learning and maintenance are substantially more complex than we have realized. Rather than being used in real-time, the influence of auditory feedback on song plasticity appears to occur offline, when the bird is not singing.

Taken together, these three projects significantly expand our knowledge of the neural mechanisms responsible for the generation of birdsong.

Chapter 2 : Neuronal ensemble coding of birdsong

The work described in this chapter was done in collaboration with Michale Fee (Bell Labs)

In many models of neural population coding, similar sensory or motor states are represented in the brain by similar neural ensembles (Georgopoulos et al., 1999; Lewis and Kristan, 1998; Wilson and McNaughton, 1993). To explore this issue, we measured the activity of large numbers of single neurons in the pre-motor nucleus RA of the singing zebra finch. During singing, individual RA neurons generate precise bursts of action potential spikes (Yu and Margoliash, 1996). We found that highly similar song elements were typically produced by stereotyped but uncorrelated ensembles of RA neurons. Small changes in acoustic structure between two syllables were frequently subserved by entirely different ensembles of RA neurons. Furthermore, the neural activity in RA often changed on a time scale an order of magnitude faster than the acoustic structure in the song, causing constant acoustic outputs to be produced by sequences of rapidly changing neural ensembles. Thus there are dynamics internal to the song control system that occur reliably and precisely yet are not directly correlated with structure in the song itself. These data represent a new type of neural code underlying complex learned behaviors.

2.1 Introduction

Zebra finch song contains spectral and temporal structure over a wide range of time scales. The individual sound elements of the song (syllables) contain distinct spectral features and are produced in a fixed sequence known as a motif. Each time the zebra finch sings, the motif is repeated a variable number of times. This learned behavior is remarkable in its stability and precision; once the song is mastered, the bird will generate it identically for the remainder of his life (Marler, 1970). A set of discrete brain nuclei known collectively as the song system underlie this complex behavior, and within this system, the forebrain pre-motor

nucleus RA (robust nucleus of the archistriatum) controls the instantaneous temporal and spectral structure in the song (Vu, 1994, Chi and Margoliash, 2001). In this chapter we explore the relation between the activity of different RA neurons and their coordination in the production of song structure.

Early recordings in RA demonstrated that neural activity increased substantially before singing began, indicating that RA has a pre-motor role in song production (McCasland, 1987). More detailed experiments in singing birds by Yu and Margoliash (1996) demonstrated that individual RA neurons fire sequences of rapid and highly stereotyped bursts of action potentials at specific time points in the song, and these bursts are timed with submillisecond precision. Stimulation in RA of the awake nonsinging bird generates sounds that contains learned acoustic structure (Vicario and Simpson, 1995), and stimulation of RA during song produces distortions in the syllable being generated at the moment of stimulation (Vu et al., 1994). RA is thus a forebrain nucleus specialized for complex, learned vocal control.

RA is approximately 700 um in diameter, and contains two classes of neurons: projection neurons, whose axons exit the nucleus, and interneurons, that only make local GABAergic connections within RA (Spiro et al., 1999). The projection neurons receive input from HVc (Nottebohm et al., 1976), and both classes of neurons receive input from the anterior forebrain pathway known to be involved in song learning and maintenance. The interneurons provide fast inhibition to projection neurons and are capable of synchronizing projection neurons that are not directly connected to each other (Spiro et al., 1999). RA is organized myotopically, with respect to the syringeal muscles (Vicario and Nottebohm, 1988). It projects to the tracheosyringeal portion of the hypoglossal nucleus, whose motorneurons project directly to the syrinx, as well as to dorsal regions of the thalamus involved in respiratory control (Vicario, 1991). This anatomical information, and the physiological data discussed in the previous paragraph, summarize our current understanding of RA's place in the neural control system used to generate song. Many pieces of the puzzle are clearly missing. In particular, the neural code used by ensembles of interacting neurons in RA to generate specific types of acoustic structure has not yet been explored. To address this question, we have made the first measurements of the detailed structure of the RA neuronal population activity during singing.

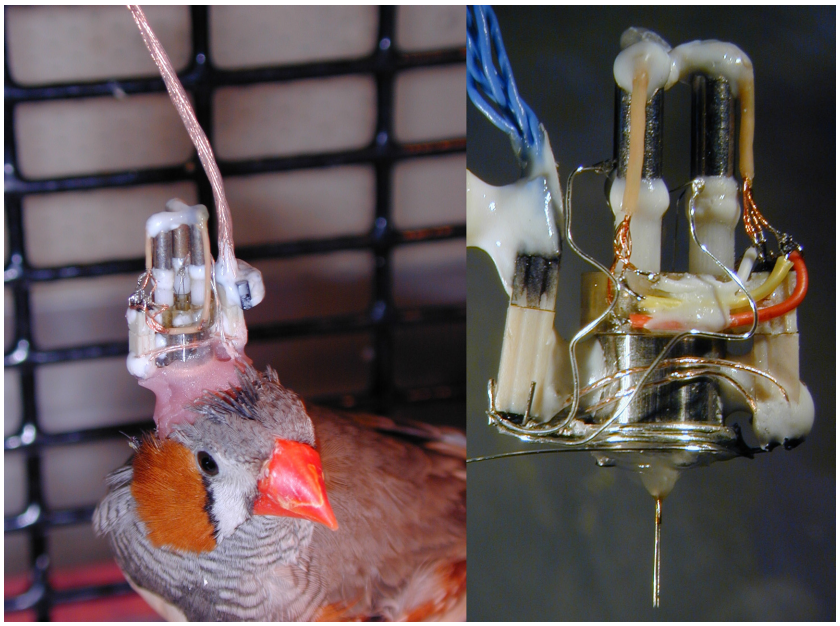

Figure 2.1. Miniature motorized microdrive for chronic recording in small animals. The microdrive weighs 1.5 g and contains three independent, motorized electrodes which can be positioned remotely.

2.2 Results

Using a newly developed 3-channel motorized microdrive (Fee and Leonardo, 2001; Figure 2.1), approximately 90 single neurons were recorded in four singing zebra finches. Pre-song, these neurons had periodic firing rates of ~20-30 spikes/sec, but were not synchronized with each other. During each motif of song, the baseline firing patterns were replaced by a bursting regime in which each neuron fired a stereotyped sequence of spike bursts (Figure 2.2). As has been shown previously, these bursts were locked to acoustic structure in the song with millisecond precision (Chi and Margoliash, 2001), and were locked to each other with submillisecond precision (Yu and Margoliash, 1996). We calculated the inter-spike-interval (ISI) histograms obtained during song for each neuron, and found they were narrowly peaked at 2.5 msec (Figure 2.3a). This sharp peak in the ISI histogram reflects the separation between inter-spike-intervals within bursts and inter-spike-intervals between bursts. The transition point between these two modes of activity can be seen from the ISI histogram to occur at ~8 msec. Using this as a threshold, we decomposed each spike train into a series of bursts (ISIs < 8 msec) and intervals between bursts (ISIs >= 8 msec). Based

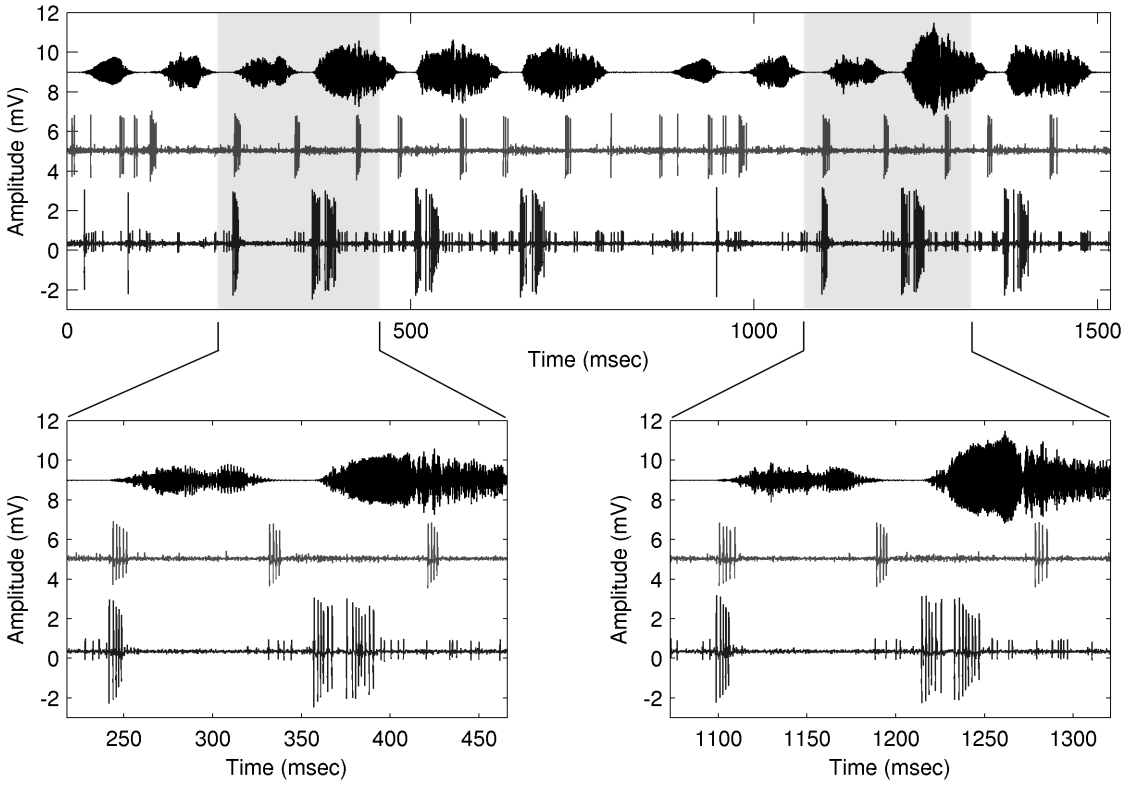

Figure 2.2. Simultaneous recording of three RA neurons in the singing bird. Top panel: two motifs of song are shown along with three simultaneously recorded RA neurons (one neuron on the first electrode, and two separable neurons on the second electrode. Bottom panel: RA neurons generate action potential bursts precisely locked to each other and to structure in the song. Each time the same syllables are repeated from motif to motif of song, virtually identical spike patterns are generated by the neurons.

on this transformation, we found that the average length of a spike burst in RA was 10 msec. This reflects the number of short ISIs that tended to occur in series.

The precise burst timing seen in RA spike trains suggests that patterns of RA neuronal activity may play an important role in the generation of song. Although most of the neurons in the data set were recorded as simultaneous pairs or triplets on different electrodes, these groups of neurons were too small to make any inferences about the concerted neural activity of the entire nucleus. However, the spectral structure in zebra finch song is reproduced very accurately each time the bird sings, and this suggested a method for aligning the firing patterns of RA neurons that were not recorded simultaneously. By lining up common spectral features produced each time the bird sang, we were able to align the spike trains of neurons recorded at different times onto a common time axis. The lengths of the

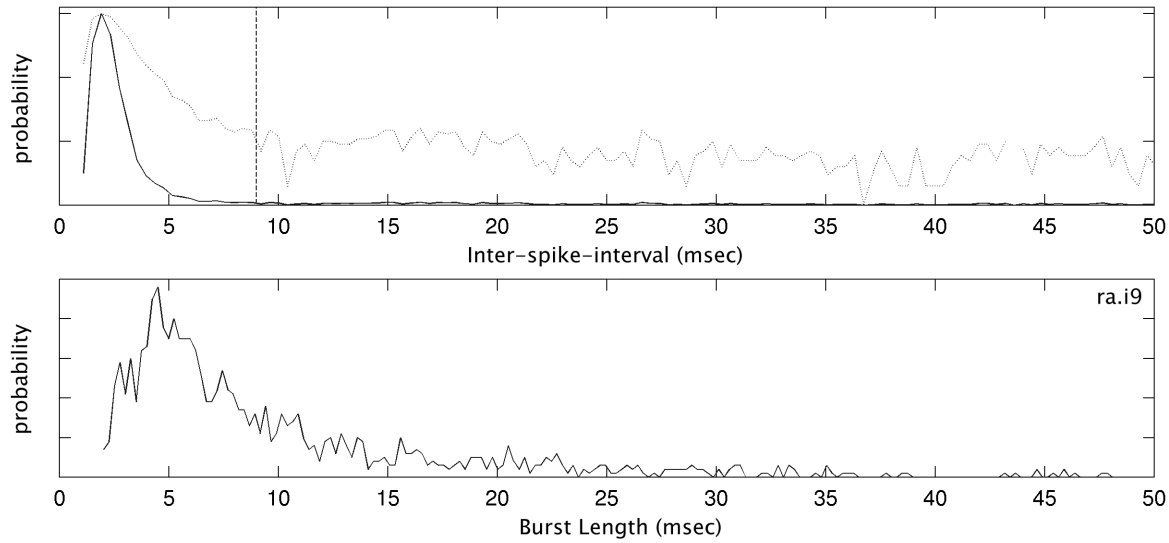

Figure 2.3. Inter-spike-interval and burst-width distribution for RA neurons. The frequent bursts generated by RA neurons produce a sharp peak at ~2.5 msec in the ISI histogram (top panel; dotted line is the log probability). At ~8 msec, there is a transition between the two types of ISIs produced by RA neurons, those within bursts and those between bursts. Using this value as a threshold, each spike train was decomposed into a sequence of bursts and burst intervals. From this, a histogram of burst lengths was estimated for the population of RA neurons recorded in each bird. The burst length distribution is peaked at ~6 msec, with a mean length of 10 msec (bottom panel).

individual syllables varied independently of each other from motif to motif of song, and were randomly stretched and compressed by approximately 5% (Figure 2.4). By estimating the magnitude of this acoustic time-warping and then compensating for it appropriately in each syllable that the bird sang (see Methods), we were able to remove a large portion of the temporal variability inherent in zebra finch song. The spike trains of the simultaneously recorded neurons were then projected onto the aligned time axes of their associated song syllables. Because all of the songs were aligned to a common reference, this procedure mapped the spike trains of all of the sequentially recorded neurons onto a common time axis, allowing them to be analyzed as a group. Because the spike trains of individual RA neurons were time locked to the bird's song (Chi and Margoliash, 2001), the time locked structure between different neurons was revealed by aligning them on a time axis based on the location of precisely reproduced acoustic features in the song (Figure 2.5). As a consequence, we

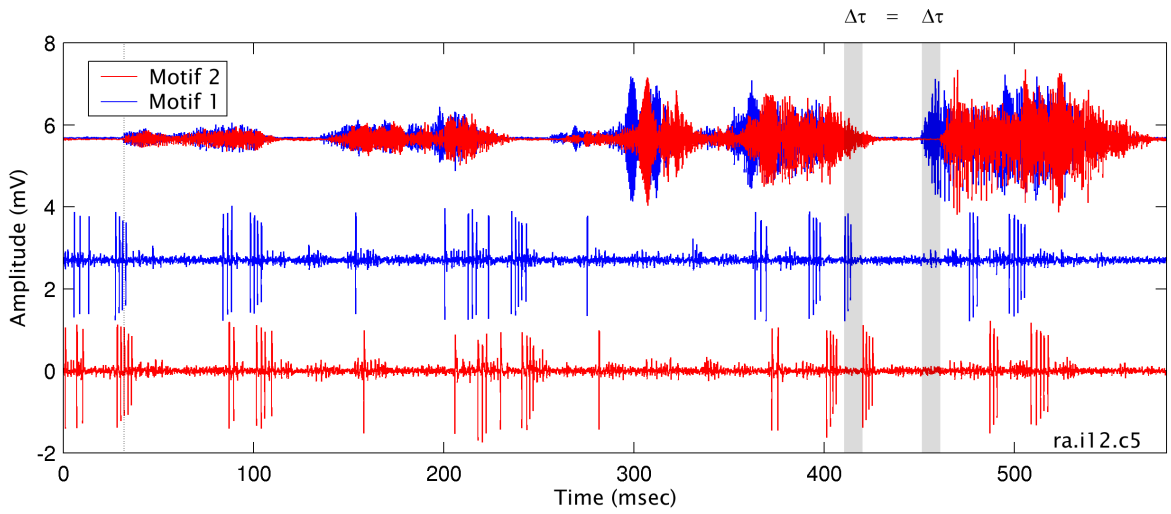

Figure 2.4. Effect of acoustic time-warping on song motifs and neural recordings. Top trace: two motifs of song, aligned to the onset of the first syllable. Middle and bottom traces: activity of a simultaneously recorded RA neuron during each song motif. The lengths of individual syllables vary from motif to motif of song, making alignment to a single reference point inaccurate; at time points far from the reference there is considerable alignment jitter. This jitter is exactly matched by variability in the onsets of bursts in the neural recordings.

were able to analyze the collective activity of 20-40 neurons, instead of the 2-3 that we could record simultaneously.

Examination of the song-aligned neural data for each bird revealed that each time point in the bird's song could be characterized by approximately 15% of the population of recorded neurons that were currently bursting. Each pattern of neural activity formed a short-lived neural ensemble. Every time the bird sang, the neurons in RA sequenced through the same set of patterns (Figure 2.5), and the neural ensemble consequently evolved in a predictable manner. How similar were the patterns of neural activity that were generated at different time points in the bird's song? We addressed this question by constructing a two-time correlation matrix, such that any point [t1,t2] in the matrix was defined as the correlation between the patterns of neural activity occurring at times t1 and t2 in the song (see Methods). The correlation was normalized to be 1 if the ensembles were identical in membership and firing rate. Figure 2.6a shows the correlation matrix for the RA population shown in Figure 2.5. The matrix diagonal represents the correlation of a neural ensemble to itself, and is always one. Off-diagonal values of the matrix are symmetric, ranging from 1

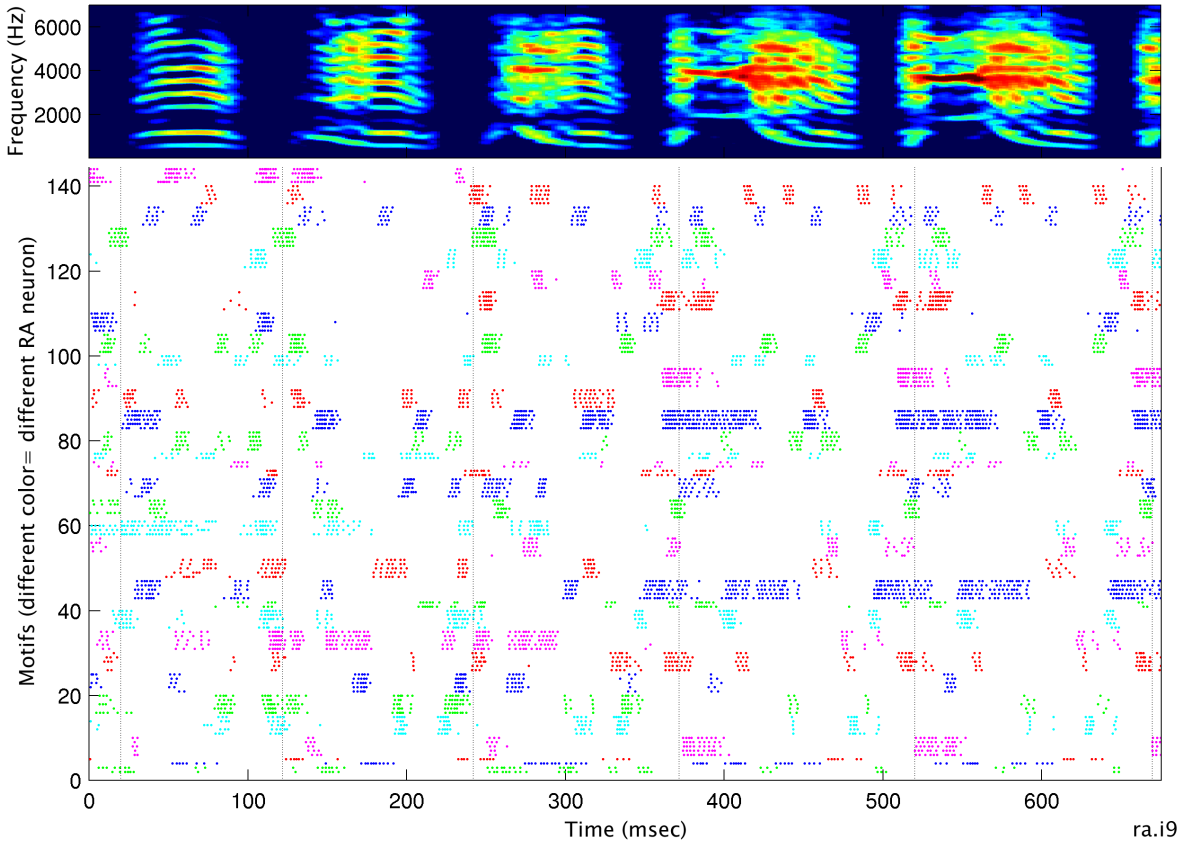

Figure 2.5. RA neuronal ensemble activity during singing (35 neurons). The activities of the individual neurons have been converted into spike train rasters. Each row represents one motif of the song shown in the spectrogram at the top of the figure. Different colored spike trains represent the activities of different RA neurons which have all been song-aligned to a common time axis. Each RA neuron reliably produces a unique pattern of bursts. Over the course of the entire motif, different neuronal ensembles turn on and off in an intricate pattern.

(identical neural ensembles) to zero (uncorrelated neural ensembles), and occasionally to values less than zero. Negative correlations indicate two neural ensembles that shared a common set of neurons that increased their average firing rate in one group and decreased it in the second group. Repeated sequences of neural activity across the population of RA neurons appear in the matrix as continuous lines of high correlation. The average of the neural correlation matrix along its rows represents the average correlation of each neural ensemble to all other neural ensembles, and measures how similar different neural ensembles were to each other. Interestingly, we found the average correlation between different neural

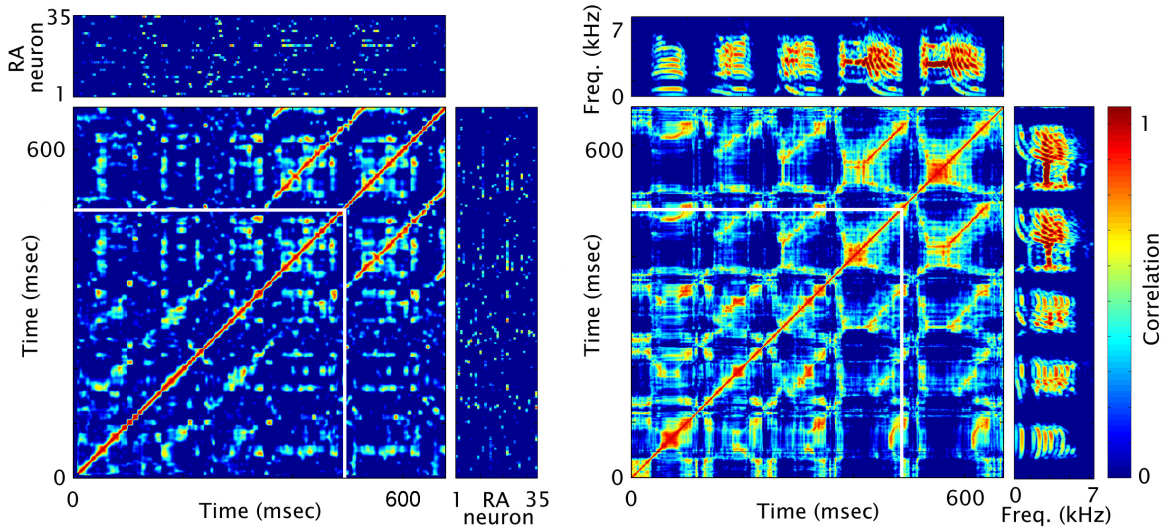

Figure 2.6. Neural and song correlation matrices. Left panel: Each point in the neural correlation matrix represents the correlation between two patterns of RA neural activity. The matrix of average firing rate activity for each of the 35 neurons in the recorded population is shown along each axis of the neural correlation matrix. Right panel: The song correlation matrix. The time-frequency spectrogram, from which sound correlations were calculated, is shown along each axis of the song correlation matrix. The white lines in the song and neural correlation matrix separate the unique part of the song from a repeated syllable which the bird stuttered. During the stuttering of this syllable, a repeated sequence of patterns of neural activity can be seen to occur in the neural correlation matrix (the red line parallel to the diagonal). A similar track of high correlation activity in the song correlation matrix shows the time course of the repeated sound.

ensembles in RA to be zero. Each pattern of neural activity in RA tended to be unique, and occurred at only one time point in the song. Thus over the course of the song, different sub-populations of neurons transiently turned on and off in an intricate pattern.

We measured how long patterns of RA neural activity lasted by averaging the correlation matrix along its paradiagonals. This is the neural ensemble auto-correlation over the entire duration of the song (Figure 2.7), and is a measure of the persistence time each neural ensemble before it was replaced by the next pattern. The auto-correlation had a 10 msec width, indicating that patterns of neural activity tended to be stable for 10 msec. By adding random time offsets to the firing profile of each neuron, and then recomputing the neural correlation matrix and auto-correlation, we were able to measure how much the precise alignment of different neurons relative to each other contributed to the persistence time of a neural ensemble. For example, if bursts tended to line up with each other (onset to

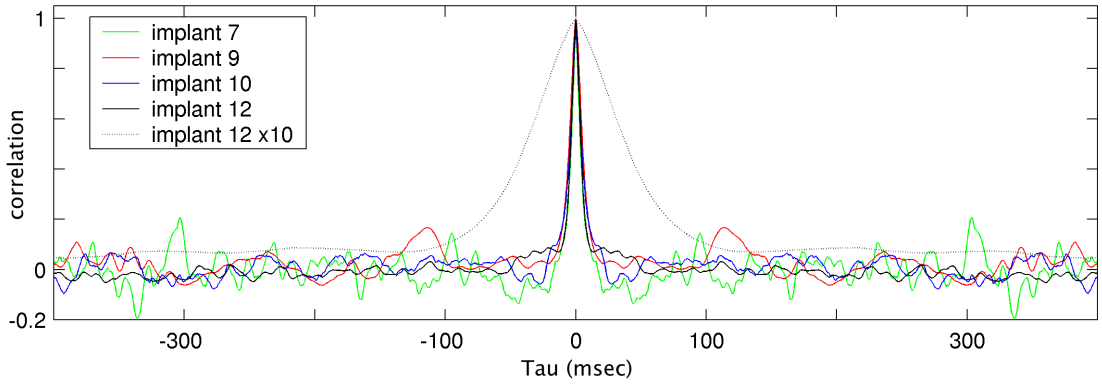

Figure 2.7. Neural ensemble auto-correlation. Patterns of neural activity tended to last, on average, 10 msec, as can be seen from the width of the auto-correlation peak.

onset and offset to offset), this could substantially increase the correlation time of that pattern of neural activity. The auto-correlation of the randomized neural data should then be significantly narrower than that of the unrandomized neural data, since the synchronization between neurons has been removed. However, we found that the auto-correlation width of the randomized neural data was also 10 msec. Thus, the patterns of neural activity in RA last for 10 msec only because of the distribution of burst widths for RA neurons (which have a mean of 10 msec), and not because of any substantial degree of synchronization between RA neurons.

The importance of the time scale of activity in RA and the similarity between different patterns of neural activity can only be fully appreciated by considering their relation to the spectral structure they generated in the song. Each segment of song that occurred in a small window of time was decomposed into its sound frequency spectrum. A song correlation matrix was then constructed in a manner similar to that of the neural correlation matrix, so that each element in the matrix represented the correlation between the sounds produced at two different time points. The song correlation matrix resembles that of the neural correlation matrix (Figure 2.6a), but it contains noticeably more areas of high correlation (Figure 2.6b). This observation was confirmed by visual inspection of the time-frequency spectrogram of the bird's song: although in general individual syllables are not repeated within a single motif, there are frequently repeated syllable-subsequences, stretches of sound 20-50 msec in length that recur multiple times.

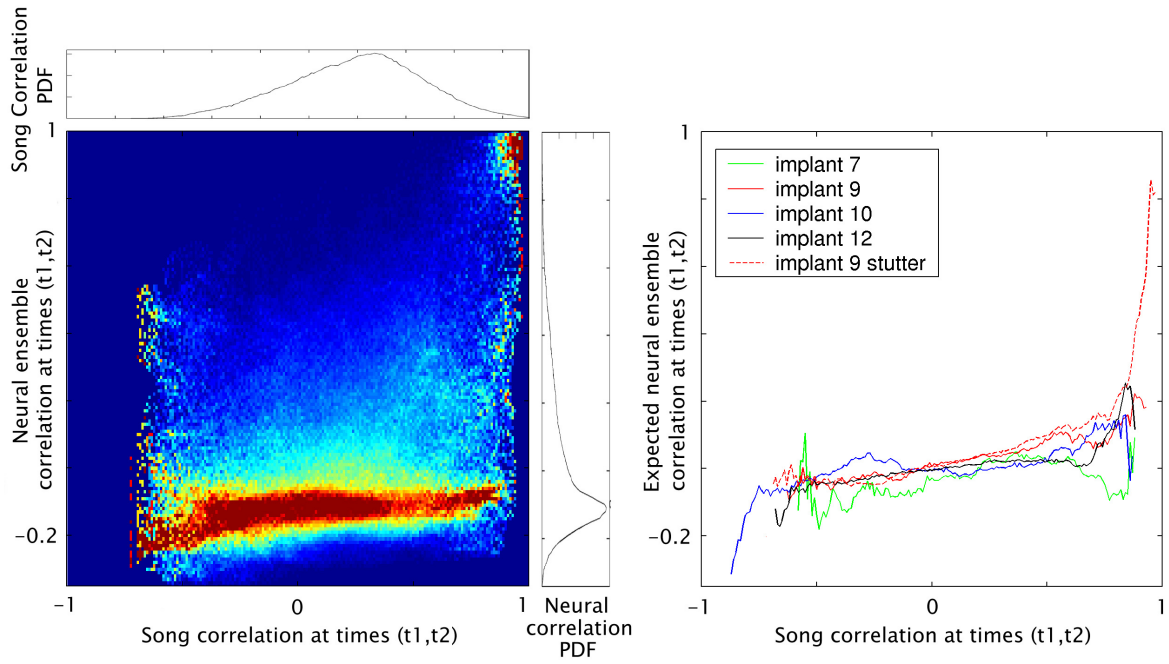

Figure 2.8. Conditional probability distribution for neural ensemble correlation. Left panel: each column represents the distribution of neural ensemble correlations, given the level of song correlation indicated on the x-axis. There is an abrupt transition, in which sounds that are nearly identical (correlations > 0.90) are generated by highly similar patterns of neural activity, but below this threshold there is no longer any relation between the similarity of songs and the similarity of the neural ensembles which generated them. The bright spot in the top right corner of the matrix represents the correlation of neural ensembles generating the stuttered syllable shown in Figure 2.6. Right panel: expected neural ensemble correlation given the level of song correlation indicated on the x-axis. Note that in general, sounds with similar spectral structure are represented by completely different, uncorrelated patterns of neural activity, unless the two sounds are virtually identical as in the stuttering of a syllable (red dashed line).

Given two highly similar sounds, the similarity of the patterns of neural activity that generated them may be examined by finding the corresponding locations in the song and neural correlation matrices (after a latency correction; see Methods). We quantified the relation between neural pattern structure and song spectral structure by estimating the conditional probability distribution for the correlation between two neural ensembles, given the level of correlation between the two sounds they generated. For example, consider all of the locations in the song correlation matrix that have a correlation between 0.95 and 0.96, indicating that they are highly similar sound pairs. The corresponding locations in the neural correlation matrix form a distribution of neural ensemble correlations that reflects how

similar the patterns of neural activity were that generated these similar sounds. A two-dimensional histogram of conditional probability distributions was obtained by calculating each neural ensemble correlation distribution as the song correlation stepped from –1 to 1 (Figure 2.8a). Each column of the image shows the distribution of neural ensemble correlations, given a song correlation in a particular interval (e.g., 0.95 to 0.96). Figure 2.8b shows the means of the two-dimensional histogram; that is, the expected neural ensemble correlation as a function of song correlation. For virtually identical sounds, such as the stuttering of a syllable (Figure 2a, b) or the repetition of syllables in a second motif of song, the neural ensembles tended to be highly correlated. However, for song correlations less than 0.9, the neural ensembles became completely uncorrelated, despite the strong similarities that were still present between the sounds. In other words, although the bird was producing many similar sounds at different time points in the song, he generated each of them using entirely different ensembles of RA neurons. It is quite surprising that small variations in the spectral similarity of two sounds were not accompanied by small variations in the RA neural activity that generated them.

We used the correlation matrices to quantify the relation between the time scale of the spectral structure in the song and the time scale of the neural ensemble activity in RA. Some syllables, like chirps, contain rapidly changing time-dependent spectral structure, whereas others, like harmonic stacks, contain relatively constant spectral structure for their entire duration (Figure 2.9). Measurements of the properties of the vocal organs and their associated muscles (Fee et al., 1998; Goller and Suthers, 1995; Goller and Suthers, 1996) have shown that, in general, a constant sound like a harmonic stack is produced by a fixed syringeal configuration, whereas a rapidly changing sound is produced by a changing syringeal configuration. We analyzed the time scale of correlations in syllable spectral structure and neural ensemble structure by estimating the width of the correlation matrix diagonals (see Methods). The width of the correlation matrix diagonal reflects the length of time over which a sound or pattern of neural activity was constant. (Figure 2.10a). A scatter plot of neural ensemble widths as a function of corresponding song widths (Figure 2.10b) shows that song widths occurred over a range of values a factor of ten larger than those of the neural ensembles. A linear regression across the combined data sets of all four birds showed a correlation value of 0.04, and a slope of 0.08 ($p < 0.0001$). The time scale of the neural

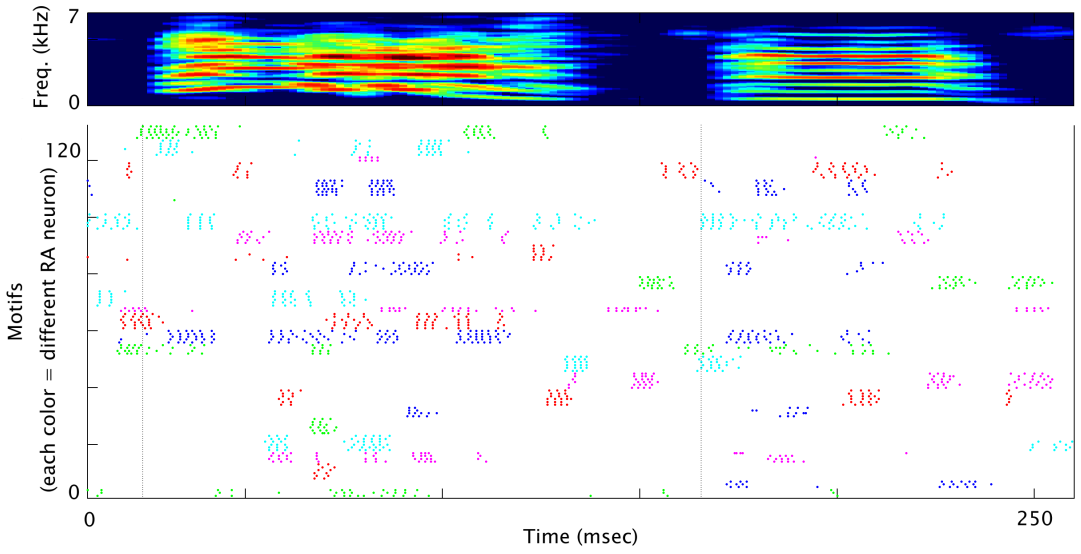

Figure 2.9. Rapidly and slowly changing sounds in zebra finch song. Top panel: time-frequency spectrogram for a complex syllable with fast transitions in spectral structure, and simple syllable with relatively constant spectral structure. Bottom panel: aligned spike train rasters; as in figure 2.5, different colored spike trains denote different RA neurons.

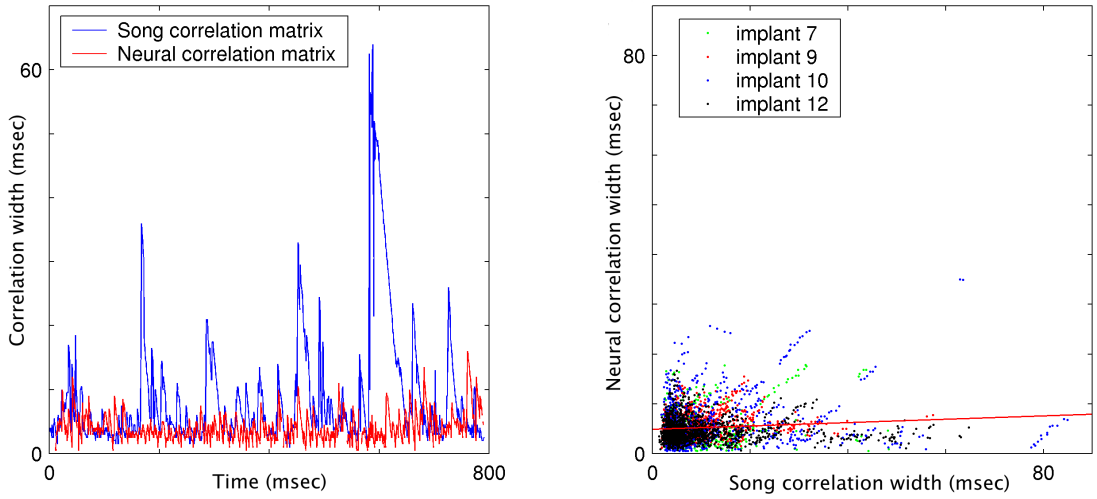

Figure 2.10. Constant acoustic structure is produced by rapidly changing neural ensembles. The width at which the correlation matrix diagonals dropped below a threshold was used to estimate the time scale of constant activity in the neural and song correlation matrices. Left panel: Width of the song correlation matrix (blue) and neural correlation matrix (red) as a function of time in the song, using a correlation threshold of 0.6. Right panel: scatter plot of neural correlation widths as a function of song correlation widths. Although the time scale of constant acoustic structure in the song was as long as 100 msec, the time scale over which patterns of neural activity in RA remained constant was only ~10 msec.

activity in RA was therefore not well correlated with that of the spectral structure of individual syllables, which ranged from ~10-150 msec, but was instead constant at a width of approximately 10 msec. This has considerable implications for the downstream decoding mechanisms that convert RA neural activity into sound.

2.3 Discussion

How could rapidly changing, uncorrelated neural activity in RA produce constant, correlated vocal outputs? The anatomy of the song control system makes such a transformation conceptually straightforward. Motorneurons, which receive input directly from RA, could integrate the short burst sequences into analog control signals of arbitrary length and amplitude. In this model of the system, burst timing between neurons is the dominant feature. Consistent with this, the results of our analysis are unchanged regardless of whether we use analog or binary firing rates to define the RA neural ensemble activity. There is a tremendous convergence between RA and the vocal organs: approximately 10000 RA neurons [ref], of which 15% are active at any time, project, via the tracheosyringeal nerve, to approximately 10 syringeal muscles (Goller and Suthers, 1996). This suggests there is substantial redundancy between RA and the syrinx. Our analysis is consistent with this idea, as many different RA subpopulations are capable of producing the same vocal output. RA is organized myotopically, with respect to the syringeal muscles (Vicario and Nottebohm, 1988).

Recent experimental results in HVc (Hahnloser et al, 2002) have shed considerable light on the upstream mechanisms that may drive the uncorrelated patterns of neural activity that we observe in RA. During singing, the HVc neurons that project to RA fire at only a single time point in each motif. Each RA neuron, in contrast, tends to fire at 10-15 specific time points in the song, because it receives input from many HVc neurons, as well as from other RA projection neurons. Given that each HVc neuron fires only once during each motif of song, the random projections from HVc neurons to different subpopulations of RA neurons could account for the lack of correlation between RA neural ensembles at different time points in the song. Because there is no correlation in HVc activity over the course of a single motif of song, there is no correlation in RA activity. As the HVc projection neurons

also generate brief bursts of spikes, there is a short time scale of neural activity present in HVc that could drive the short time scale of activity we observe in RA.

Our description of the patterns of neural activity generated in RA during singing also provides an explanation for some recent behavioral observations on the process of song learning. Tchernikovski et al. (1999) have shown that rather than learning a few sound primitives (like harmonic stacks, frequency downsweeps, and whistles) and then refining these into mature syllables occurring in a particular sequence, extremely different syllables can in fact emerge from the same primitives. For example, consider a scenario in which the zebra finch is learning three sequential syllables, a harmonic stack (H), followed by a frequency downsweep (F), followed by a harmonic stack (H). When the bird begins the song learning process, he generates a sequence of sounds that randomly begin as F-H-H. The bird must learn both the detailed spectral structure of the three syllables and their proper order (H-F-H). Now, in one model of vocal control, the zebra finch could refine the spectra of each syllable and then switch their orders so they followed the proper sequence. In this scheme, each syllable is generated from a primitive that is similar to its final form. What Tchernikovski et al. (1999) observed, however, was that the zebra finch slowly transformed each initial syllable into its final target, simultaneously learning both the spectral structure and the sequence order. There was no direct correspondence between the final structure of a syllable and its starting point. Our results clarify the neural basis of the results of this learning process. There is no representation of syllable primitives in RA, such that different sounds of the same type (like harmonic stacks with different fundamental frequencies) are generated by correlated patterns of neural activity with subtle differences between them. Instead, the zebra finch learns to activate a different RA neural ensemble at each time point in the bird's song, independent of what acoustic structure is being produced. These different neural ensembles are slowly refined until they generate the desired spectral structure. Similar sounds are thus subserved by entirely different patterns of neural activity.

The tuning of single neuron responses to sensory or motor features (Georgopoulos et al., 1986; Lewis and Kristan, 1998; Wilson and McNaughton, 1993; Bialek et al., 1991) has dominated the classical view of neural coding. Within these systems, smooth variations in a stimulus feature typically produce smooth variations in the neural activity that represents it. For example, the direction of reach of a monkey can be decoded by a vector average of the

activities of individual M1 neurons with respect to their preferred directional tuning (Georgopoulos et al., 1986). A small change in the direction of reach will, according to the model, produce a small change in the activity and membership of the M1 neural ensemble. A second property of these types of neural representation is that adding more neurons, appropriately distributed across the encoded feature space, to the population estimate of the stimulus will generally produce a steady increase in the precision of the estimate. For example, in rat hippocampus, increasing the number of place cells used to estimate the rats spatial position produces a steady increase in the spatial resolution of the estimate (Wilson and McNaughton, 1993).

Our results describe a very different type of neural representation, one in which there is no systematic mapping between the distributed activity of the network and the vocal output of the bird. In contrast to neurons in the systems described above, RA neurons show no consistent tuning to a single acoustic feature across the set of time points in the song when they are active. Furthermore, small changes in spectral structure are subserved by the activation of completely different subpopulations of RA neurons, despite the strong correlations that are known to be present in the vocal muscles during the production of similar sounds. Finally, the addition of neurons to the RA neuronal ensemble does not yield a steady increase in predictability about what type of sound is being generated. One neuron may produce ten bursts during the song motif, each timed to a completely different type of sound. The knowledge that this neuron was active reduces the uncertainty about what sound is being produced to a set of ten different possibilities. Adding a second active neuron to the estimate reduces the uncertainty about what sound is being generated to the occasions when the two neurons fired together. This may reduce the initial ten locations to three, but there is not a corresponding increase in correlation between the remaining sound types. It is only when the neural ensemble has sufficient neurons that it becomes uniquely defined that accurate prediction of song structure is possible. Most models of neural population coding follow a sum-of-the-parts principle, in which the individual neurons provide considerable information about what is being represented, and additional neurons provide progressively more information. The neural code used within RA is substantially different from these models, and individual neurons provide little unambiguous information about song structure. These observations suggest that within the neural control system of RA, network interactions

are the dominant feature, and it is only by reading out sufficient activity to represent the entire state of the network that decoding is possible. The spike patterns we observe in RA thus represent a new type of neural code underlying complex learned behaviors

2.4 Methods

Birds were housed in custom designed plexiglass cages, and had unlimited access to food and water. All birds used in the experiment were male zebra finches (adults with crystallized songs). Females zebra finches were housed in similar, separate plexiglass cages, and were presented to the males upon isolation of one or more RA neurons (so-called directed song).

Neurophysiology. Birds were anesthetized with 1-2% isoflurane and nucleus RA was identified with an extracellular targeting electrode based on stereotaxic coordinates and electrophysiological activity. Upon identification of the center of RA, a three electrode miniature motorized microdrive was cemented onto the skull using the procedure described in Fee and Leonardo (2001). The microdrive weighs 1.5 grams, and contains three independently controlled motors, allowing each electrode to be remotely positioned extracellularly with 0.5 um spatial resolution. The electrode tips were implanted ~700 um above RA. Electrodes were made from 80 um tungsten wire, insulated with parylene, and had ~3 MOhm impedance (5-10 um tips; Microprobe, Inc.). Birds were allowed to recover for sufficient time so that they were singing reliably upon presentation of a female bird (~1-2 days). During each day of recording, a custom modified Sutter MP-285 microdrive controller was used to position the electrodes in nucleus RA and record their depths. Singing was normal in birds with implanted drives; songs produced after the implant are identical in structure to those produced before the implant.

Extracellular recordings of single neurons of up to 10 mV in amplitude were isolated on individual electrodes; isolation quality was equivalent to that seen in a head-fixed anesthetized animal. Single neuron recordings were verified with two methods. First, individual spike waveforms of 1.5 msec in length were extracted from each raw electrode signal using a 3x RMS threshold, and then interpolated by a factor of 10 to remove sampling jitter. This produced a matrix of aligned spike waveforms. We calculated the singular value decomposition (SVD) of this matrix. The eigenvectors associated with the largest two eigenvalues of the SVD represent a subspace of the matrix, which contains most of the

variability of the spike waveforms. We projected the spike waveforms onto this two-dimensional subspace (a great reduction from the original 60 dimensional space) and looked for well-defined clusters of points. These clusters represent well-isolated neurons. As we are isolating single neurons, this clustering process was not used for spike-sorting, but rather as an reliable denoising mechanism to automatically remove the occasional contamination from a second neuron or electrical artifact. Single neuron isolations were further verified by confirming the presence of a spike refractory period in the inter-spike-interval histogram. The locations of the electrodes within RA were verified histologically at the conclusion of the experiment. Neural and acoustic data were collected by custom designed software written in Labview (National Instruments).

RA is organized myotopically, with respect to the syringeal muscles (Vicario and Nottebohm, 1988). It is possible that uncorrelated sets of neurons within a single myotopic projection (from RA to the hypoglossal nucleus) generate similar vocal outputs. However, the electrodes in our microdrive were spaced sufficiently that they sampled, relatively uniformly, the entire volume of nucleus RA. Within the population of neurons we recorded in each bird, some cells came from the same area of RA, and others from different areas. Because of this, the redundancy of any single subregion of RA cannot account for our observation that uncorrelated neural ensembles generate similar vocal outputs.

Song alignment. Variability from acoustic time-warping can be removed from the data by using highly resolved spectral features in different song syllables as clock points, and then linearly stretching and compressing common segments of sound in different motifs to be the same length. This piecewise-linear time-warping is based only on the song, and is completely independent of the spike trains. The partial derivative of the song spectrogram over time contains sharply peaked transitions at the points when the spectrum of the song changes abruptly (e.g., the beginnings and ends of syllables), that are suitable for use as alignment points. For each motif, we construct a vector of times corresponding to the occurrences of common spectral features, based on the alignment of the spectral derivatives of each syllable to a set of syllable templates. Adjacent time points in this vector define an interval of sound in the song. By stretching and compressing the corresponding time intervals in different motifs to be a common length, the variability from acoustic time-warping can be normalized out of all of the recorded motifs. Each spike train is then

projected onto the warped time-axis of its corresponding acoustic signal. The alignment templates are representative features for each syllable.

Neural ensemble correlations. A neural ensemble was defined as the spatial pattern of activity across the recorded population of neurons at a particular time point in the song. Each neural ensemble was represented by a vector of instantaneous analog firing rates for each neuron in the recorded population. Let N_i be the neural ensemble activity vector observed at time t_i, and N_j be the neural ensemble activity vector observed at time t_j. We defined the correlation between the two ensembles as

$$C(t_i, t_j) = \frac{N_i \cdot N_j}{\| N_i \| \cdot \| N_j \|}$$

The mean was subtracted from each neural ensemble activity vector prior to calculating its correlation to other ensembles, to ensure the measured correlations reflect coordinated fluctuations in the two ensembles and not common DC offsets. Correlations between two null ensembles (times at which no neurons in the population fired) were excluded from the analysis as they contained no data. We also used the Hamming distance to measure the neural ensemble correlations, without change in results. The neural ensemble activity was estimated in a 0.5 msec window in 0.5 msec steps.

Song correlations. Direct multitapered estimates of the power spectrum of each song were calculated with a sliding 8 msec window, using a step size of 0.5 msec, and a time-bandwidth product of NW = 2 (Thomson, 1982; Tchernichovski et al., 2000). Sound correlations were calculated with the dot-product metric described above. However, strong correlations between different time points in the song were produced by the average power-spectrum of the syrinx and the oscillatory temporal pattern of syllables and silent intervals. These correlations obscured the fluctuations in fine spectral structure. Because of this, we first normalized out of the song the dominant spectral and temporal modes (essentially the average power spectra and the envelope).

Conditional distributions. In order to compare the song and neural correlation matrices, we first compensated for the 10 msec latency in the neural-to-song response (A. Kozhevnikov and M. Fee, unpublished measurements), and then smoothed the neural data with the same window function used to estimate the time-frequency spectrum of the song (8 msec). This ensured that the neural and acoustic data had the same spatial and temporal resolution. This

analysis was restricted to nonlocal effects – that is, events that occurred sufficiently off-diagonal that they were not influenced by the average song or neural ensemble width. Local syllable correlations were considered in the correlation width analysis.

The weakness of any analysis relating sound correlations to neural activity in the song control system is that there is no simple linear relation between acoustic structure in the song and the control parameters used by the bird to generate that structure. For example, the syrinx is known to be highly nonlinear in its response to linear control signals (Fee et al., 1998). Further, it is likely that the control parameters in RA are represented in motor coordinates, rather than auditory ones, since RA projects to motorneurons. However, despite these considerations, there are clear correlations in the activity of the syringeal muscles when similar sounds are generated (Goller and Suthers, 1996). Thus, regardless of the exact nature of the control parameters represented in RA, the main point of our analysis is that there is no correlation in the neural ensemble activity over time, despite the strong correlations in activity in the vocal control areas downstream of RA.

Correlation width analysis. For each song correlation matrix, we calculated a contour line with respect to the matrix diagonal. This contour line represents how much time elapsed before a particular sound at time t dropped below a correlation threshold. The same threshold was then used to calculate the time-varying width of the neural ensemble correlation matrix. A correlation threshold of 0.6 was found empirically to accurately reflect the length of stationary acoustic structure in the bird's song. The value of the minimum correlation widths seen in the scatter plot (Figure 4c) is a function of this correlation threshold. Large thresholds produce smaller correlation widths as the contour line becomes a tighter fit to the matrix diagonal (having zero width at a threshold of 1). However, our results (the slope of the line relating neural and song widths) are robust over a large range of threshold values, and are not affected by changes in the y-intercept of the scatter plot.

Chapter 3 : Decrystallization of adult birdsong by perturbation of auditory feedback

Young birds learn to sing by using auditory feedback to compare their own vocalizations to a memorized or innate song pattern; if they are deafened as juveniles they will not develop normal songs (Konishi, 1965; Nottebohm, 1968). The completion of song development is called crystallization. After this stage, song shows little variation in its temporal or spectral properties. However, the mechanisms underlying this stability are largely unknown. Here we present evidence that auditory feedback is actively used in adulthood to maintain the stability of song structure. We found that perturbing auditory feedback during singing in adult zebra finches caused a slow deterioration of song. This "decrystallization" consisted of a dramatic loss of the spectral and temporal stereotypy seen in crystallized song, and included stuttering, creation, deletion and distortion of song syllables. After normal feedback was restored, these deviations gradually disappeared and the original song was recovered. Thus, the brains of adult birds that do not learn new songs nonetheless retain a significant amount of plasticity.

3.1 Introduction

The song of zebra finches consists of three levels of organization: syllables, which are individual sound components of the song separated by silent intervals; motifs, which are sequences of syllables; and bouts, which are sequences of motifs (Sossinka and Bohner, 1980). The spectral structure of the syllables is unstable in the juvenile bird. Similarly, motifs and bouts are organized differently from song to song. However, as the song gradually assumes its adult form, these different levels of organization become highly stereotyped, the variability of the spectral structure of the syllables becomes extremely small, and the bird sings these syllables in a highly predictable order (Marler, 1970). At this stage the song is referred to as crystallized (Figure 3.1). Some birds, like the zebra finch, maintain

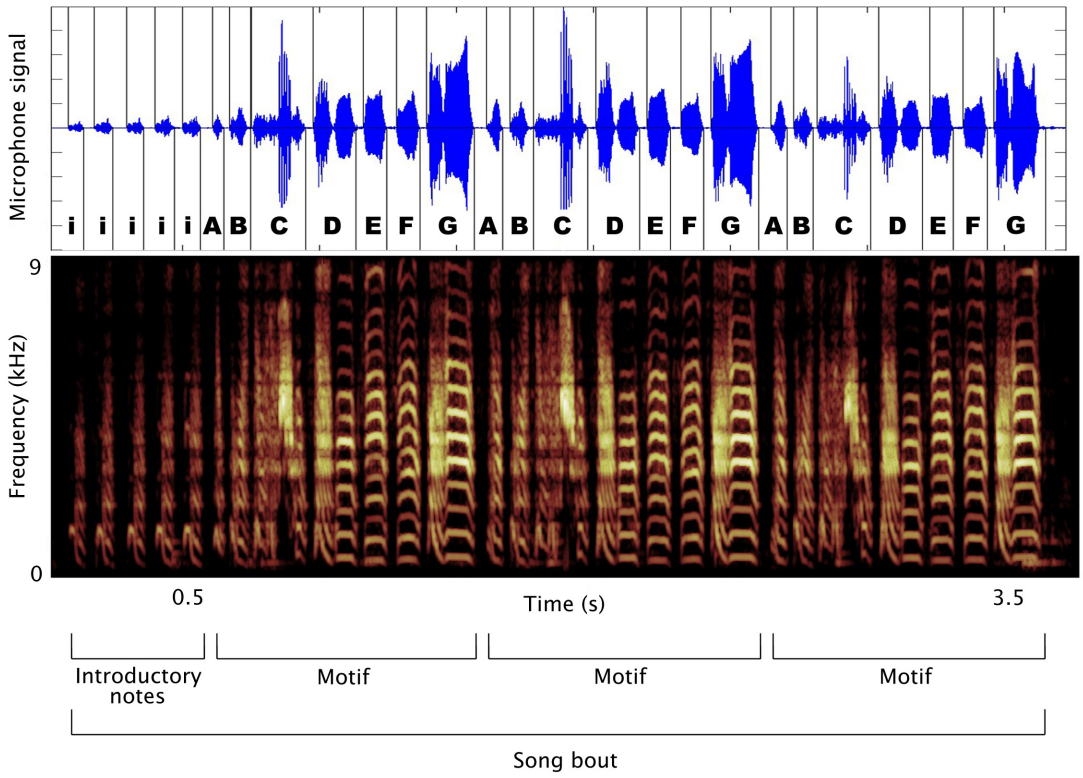

Figure 3.1. The normal structure of adult zebra finch song. Top trace shows the raw microphone signal, parsed into discrete bursts of sound (syllable). Bottom trace shows the time-frequency spectrogram of the song. Syllables are produced in a repeated sequence called a motif. During a bout of singing, the motif is repeated a variable number of times.

their crystallized song throughout adulthood and are called age-limited learners. In contrast, open-ended learners, like canaries, can learn new songs in adulthood (Marler and Peters, 1987).

The stability of song in age-limited learners was previously thought to be maintained without auditory feedback (Konishi, 1965; Nottebohm, 1968). Recent reports, however, show that deafening these birds after crystallization causes changes in song, suggesting that some auditory feedback is important for song maintenance throughout life (Nordeen and Nordeen, 1992; Okanoya and Yamaguchi, 1997; Wooley and Rubel, 1997). Six to eight weeks after deafening, adult zebra finch song undergoes a deterioration involving the addition and deletion of syllables, the abnormal repetition of syllables (stuttering), and modified syllable sequences. By opening the auditory feedback loop, deafening shows how well the song pattern generator can maintain its original output without this signal. However,

to learn how the song control system works, it is necessary to manipulate auditory feedback without disabling either the auditory or vocal control system. The ability to modify feedback signals dynamically allows particular spectral and temporal components of the song to be chosen for manipulation. Furthermore, the effects of restoring normal auditory feedback after exposure to abnormal feedback can be observed. We show here how these manipulations cause the adult song to undergo dramatic changes.

We developed a computer-controlled system to perturb auditory feedback. We designed two feedback paradigms (see Figure 3.2). In both methods, the computer detected singing and then played back a feedback signal to the bird. Because the playback signals were delivered through an overhead speaker, the birds heard a superposition of natural and artificial sounds. In the first paradigm, the adaptive protocol (n = 3 birds), a computer alternated between recording vocalizations and playing back the last vocalization to the bird. The position of the feedback varied depending on the exact timing of the bird's song, and the feedback changed as the song changed (see Methods). In order to explore the effects of feedback on a more local level, we designed a second syllable-triggered protocol (n = 2 birds) in which we perturbed the feedback of only a single target syllable. The computer recognized and played back a stored copy of this syllable each time the bird produced it. The timing of the feedback signal was fairly constant across different song deliveries.

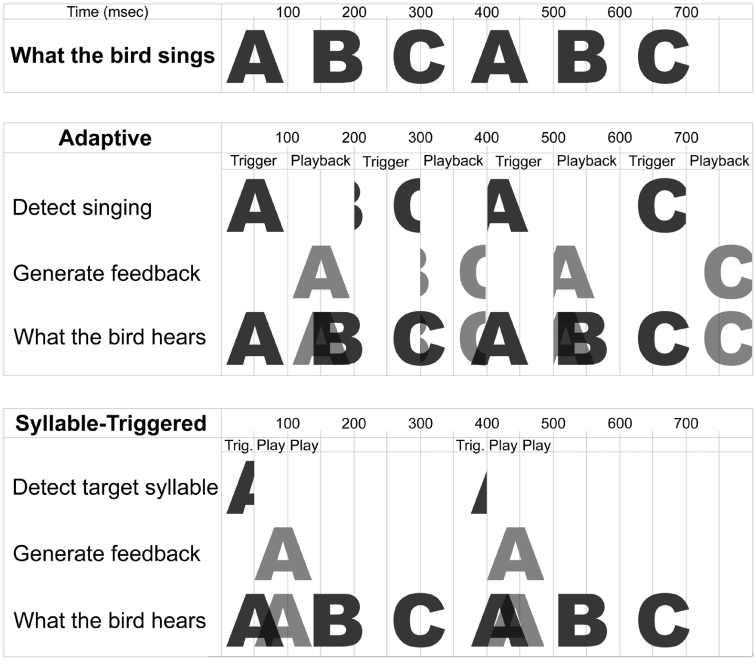

Figure 3.2. Protocols for constructing the feedback signals. In the adaptive protocol, the computer alternated between recording vocalizations and playing the last vocalization back to the bird. In the syllable-triggered protocol, the computer monitored the bird's song in 50 msec bins and detected the production of a single target syllable. In both protocols, the bird heard the superposition of his own vocalizations and the computer generated feedback.

3.2 Results

We recorded the songs of five adult male zebra finches in separate, sound attenuated chambers for several weeks before beginning the experiment. This procedure allowed us to ascertain that the variability of their songs was well within that found in crystallized songs. The birds were then placed in the feedback system. After a period of 1-4 months in this environment, four of the five birds showed dramatic changes from their original songs. To study the progression of these changes, we allowed some of the birds to sing without artificial feedback on a randomly chosen 10-15% of their song deliveries one day per week. All of the analyzed data consisted of recordings of the bird singing by himself, with no artificial feedback being produced at the time. After the feedback was permanently stopped, we tracked all the birds for another 8-16 months to determine if they were able to recover their original songs.

Decrystallization of the song occurred in both global song organization and local spectral structure, and consisted of the emergence of new spectral and temporal properties and the increased occurrence of properties that were rare in the baseline song. We use the term decrystallization to refer collectively to all of the perturbation-induced changes from the original quantitative and statistical structure of the song; this does not necessarily imply a return to a juvenile state of song structure. These changes include stuttering, creation, deletion and distortion of song syllables. As the decrystallization progressed, the proportion of normal songs decreased and that of abnormal ones increased. However, baseline songs were still produced with low probability even during the peak of song degradation.

The changes seen in the three adaptive protocol birds were very similar to those seen in deafened birds (Nordeen and Nordeen, 1992). Stuttering occurred in all three adaptive protocol birds and was the most dramatic change in song organization. Both complex syllables (Figure 3.3 a, b) and modified introductory notes (Figure 3.3 c, d) were stuttered. A secondary effect of stuttering was a substantial increase in the maximum observed song length. Other changes to song organization induced by the feedback were the addition of new syllables to the song and, infrequently, the deletion of old syllables from the song (Figure 3.4). Many zebra finch syllables contain sets of frequencies that are integer multiples of a common fundamental frequency; such sets are called harmonic stacks (Figure 3.1, syllable D). The same three birds also showed spectral distortion in their song syllables

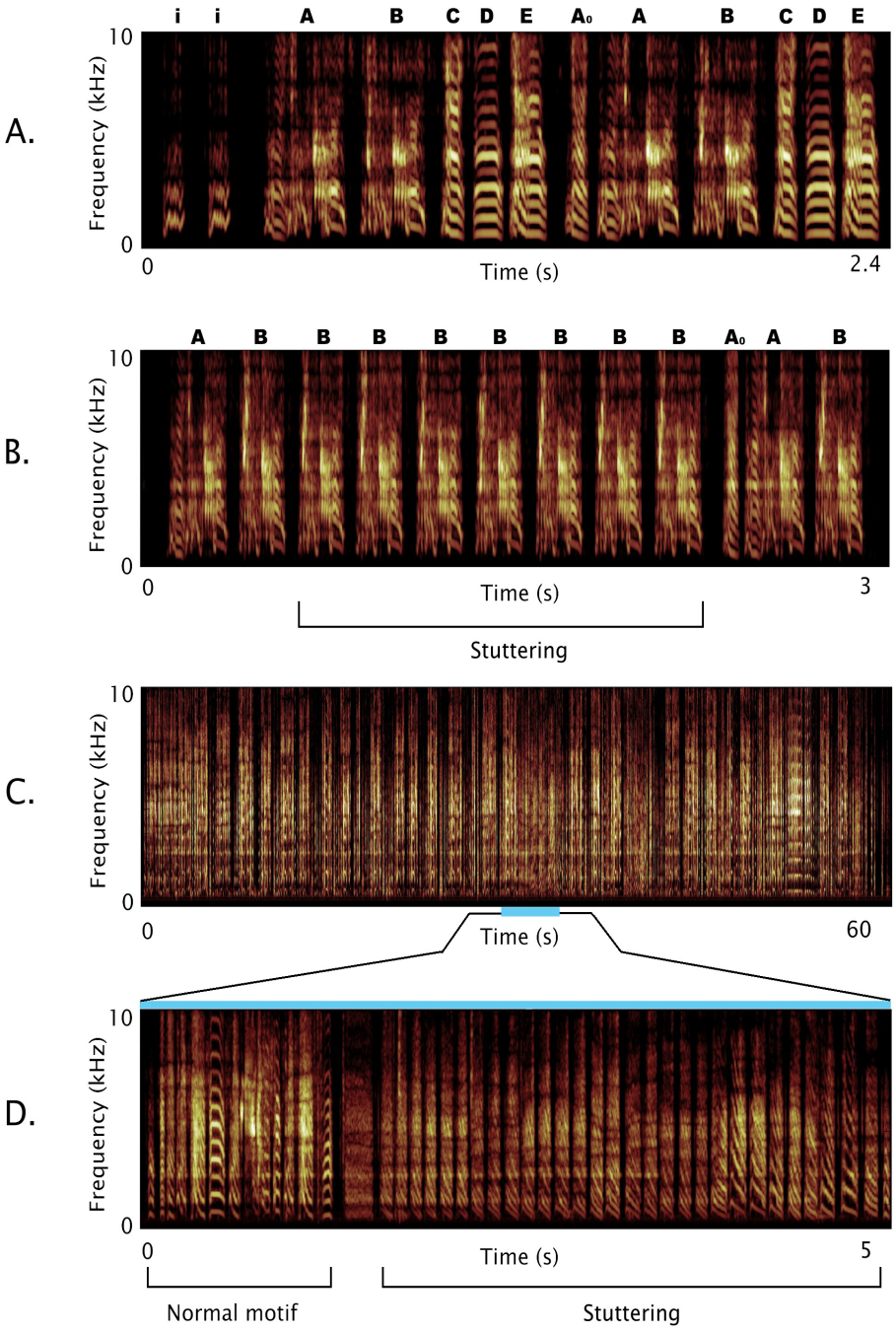

Figure 3.3. Stuttering of syllables in decrystallized birdsong. Normal (A) and decrystallized (B) songs of bird d6, who frequently stuttered a complex song syllable after two weeks of exposure to perturbed auditory feedback. Bird d5 (C, D) showed a dramatic increase in average song bout length due to the stuttering of an introductory note. Note that these abnormal songs were produced by the birds after the removal of feedback and were not simply an instantaneous effect of auditory feedback perturbation.

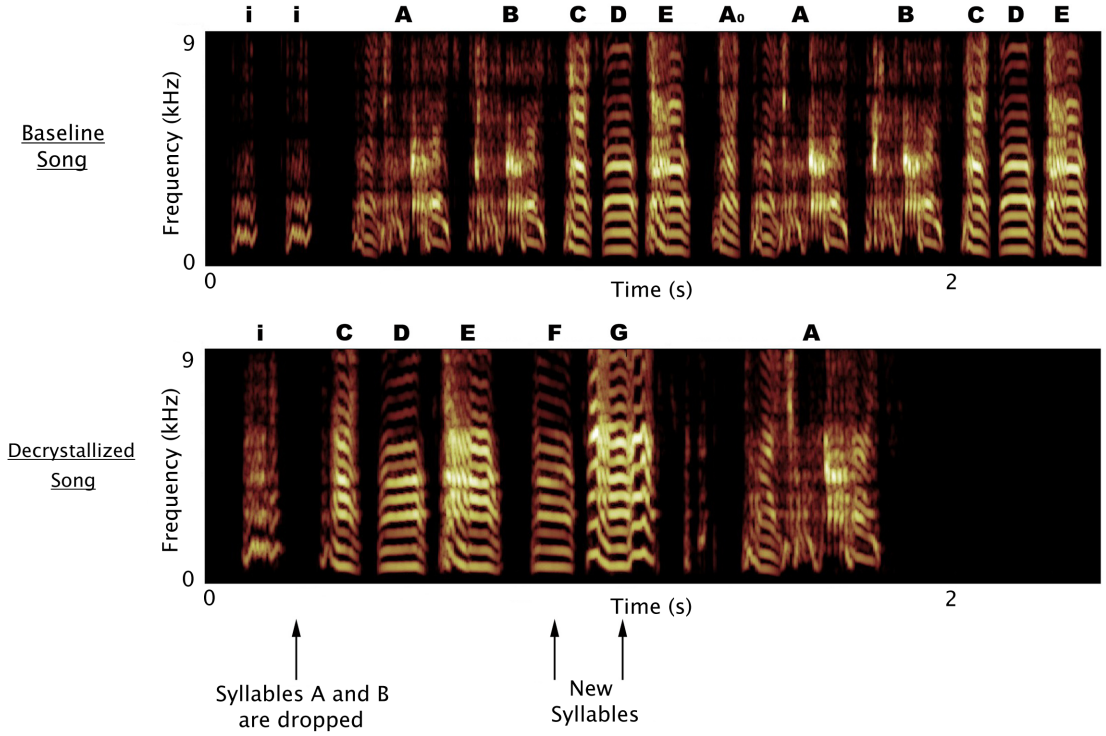

Figure 3.4. Addition and deletion of syllables in decrystallized birdsong. Top panel: normal song of bird d6. Bottom panel: decrystallized song of bird d6, showing the deletion of syllables A and B, and the addition of new syllables F and G.

(Figure 3.5) including wobbles in the harmonic structure of simple notes and the production of two superimposed harmonic stacks, suggesting a loss of precise control over the vocal organ (the syrinx; Suthers, 1990; Fee et al., 1998). There was considerable variability in the magnitude and time course of the changes between different birds, but significant changes were generally seen within six weeks. Finally, all the changes in song structure described above could occur within different motifs in the same bout (Figure 3.6). This is significant because one of the hallmarks of crystallized song is its robust temporal stereotypy - an identical syllable sequence is maintained within all the motifs of a bout. Decrystallized song, in contrast, lacks this stereotypy.

Two birds received feedback in the syllable-triggered protocol. In one of the birds, significant changes in the spectrum of the target syllable appeared in less than a week and increased in magnitude for the remainder of the feedback period. No changes were seen in the spectrum or the ordering of any of the other syllables in the song. The changes to the

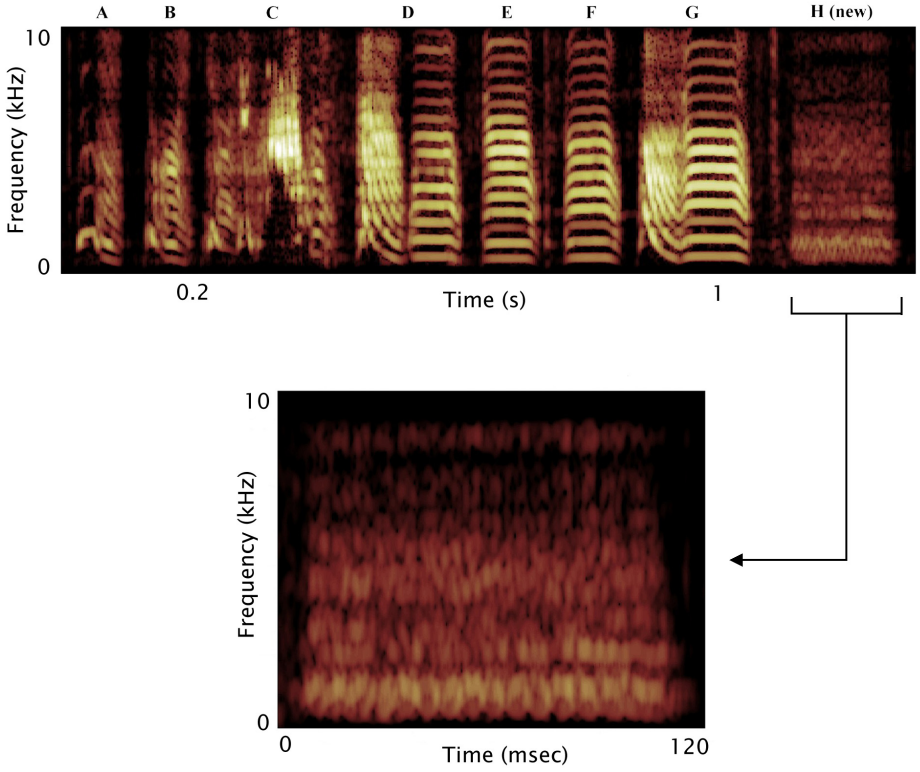

Figure 3.5. Spectral distortion in decrystallized song syllables. Syllable H appeared in the song of bird d4 after decrystallization. It has a highly abnormal sound spectrum, containing no discernable harmonics or well controlled spectral structure. Juvenile birds produce sounds very similar to syllable H before they have learned to control the syrinx properly.

target syllable consisted of the appearance of harmonic frequencies around previously single frequency portions of the syllable's spectrum (see Figure 3.7). The presence of these additional harmonics grew more frequent with time, until eventually this initially tonal syllable was sometimes produced as a distorted harmonic stack. The ability to alter the spectral structure of a sound in response to altered auditory feedback has also been observed in humans (Houde and Jordan, 1998). The second bird used in this protocol showed no changes in his song after 11 weeks of feedback perturbation.

Decrystallization essentially consisted of a large increase in the variability of the song. In order to quantify the changes associated with the arrangement and variability of syllable sequences, we developed an automatic method to sort the data acquired on a given day into syllable types by classifying the different spectral and temporal features of each

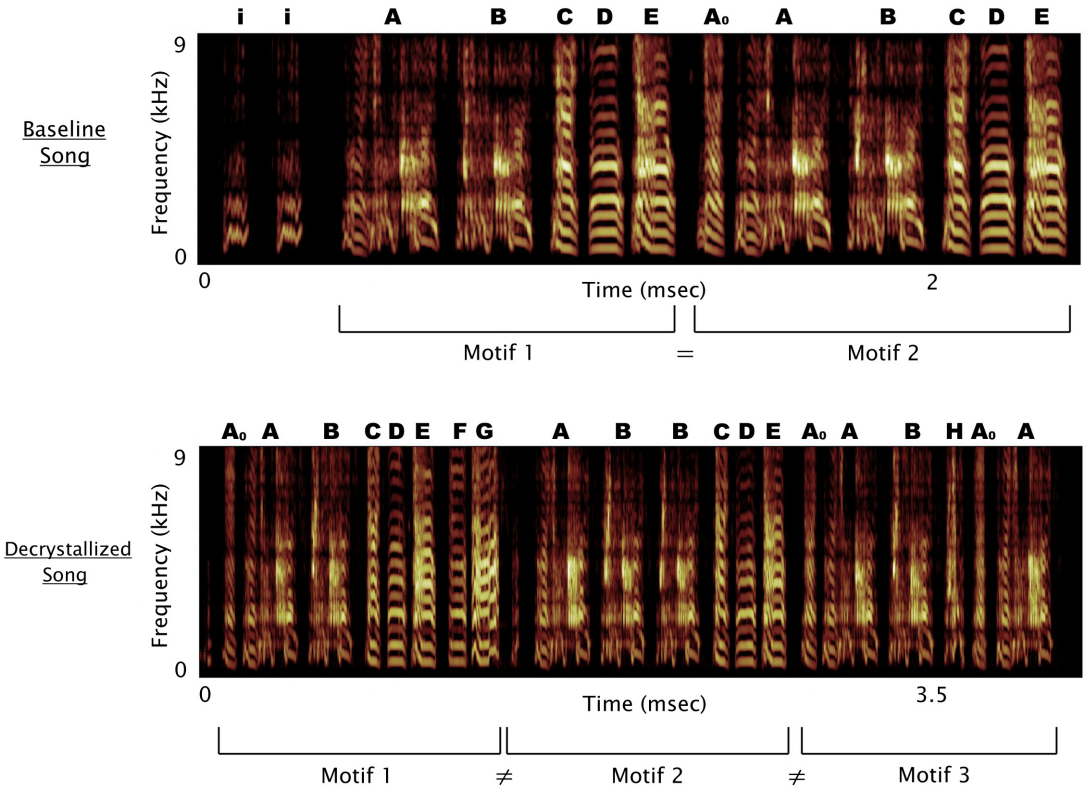

Figure 3.6. Decrystallized songs contained increased variability in syllable sequencing. A hallmark of normal zebra finch song is that song bouts are comprised of a single repeated motif. After decrystallization, in contrast, birds frequently produced song bouts comprised of different motifs.

syllable (see Methods). This transforms the raw voltage waveforms recorded from the microphone into a series of syllable strings, such as "A B C D E", from which we then calculated the probability of different syllable sequences. We defined the baseline motif as the most probable sequence of syllables that the bird repeated in a bout before the feedback period began. Figure 3.8a shows the probability of singing the baseline motif as a function of time for one bird. After receiving one month of perturbed feedback, the probability of the baseline motif for this bird decreased by a factor of six.

The probabilistic sequencing of syllables by the bird on a particular day can be fully characterized as a Markov chain (Feller, 1968), in which the likelihood of singing a particular syllable depended only on the occurrence of the last syllable produced, and not on any prior syllables. The bird's song over the course of the experiment can then be described as a

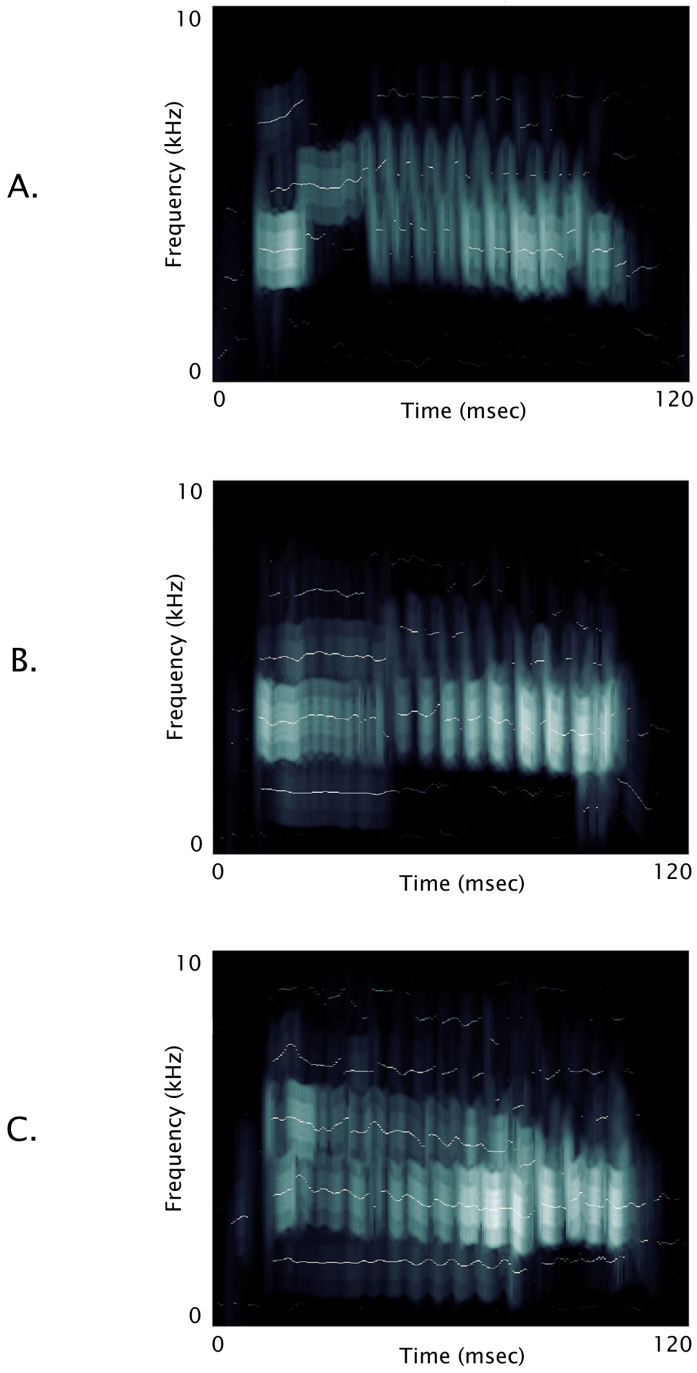

Figure 3.7. Decrystallization of a single syllable. A: baseline (pre-feedback) version of the syllable. B: After one week of syllable-triggered feedback. Harmonic frequencies (white lines) have appeared around the single frequency in the early portion of the syllable (at t = ~ 30 msec). C: After one month of feedback, additional harmonic frequencies now stretch throughout the duration of the syllable.

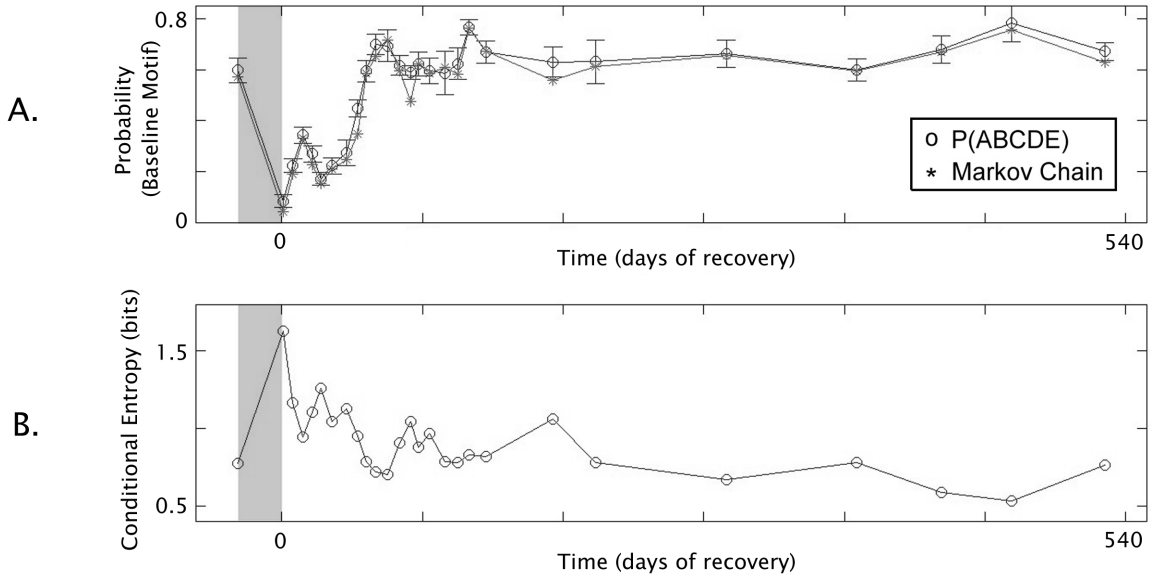

Figure 3.8. Time course of decrystallization and recovery. A) The probability of one bird singing his baseline motif as a function of time. The shaded area represents the feedback interval. Circles represent the baseline motif probability calculated from the entire sequence of syllables (Prob {A B C D E}), while asterisks represent the baseline motif probability calculated from the syllable to syllable transition probabilities of the Markov chain for the song (Prob{A} * Prob{B|A} * Prob{C|B} * Prob{D|C} * Prob{E|D}). The two curves are identical to within 4% error. B) The conditional entropy of the song, which is a measure of song variability, for the same bird.

Markov chain which evolves in time (see Figure 3.8a), and the conditional entropy of the Markov chain can be used as an estimate of the variability of the song on a particular day (Figure 3.8b; Cover and Thomas, 1991). The conditional entropy measures the uncertainty in observing syllable B, given that syllable A was just produced. If each syllable is followed only by a single other syllable (only one type of syllable sequence is produced), then the conditional entropy will be low, whereas if any syllable can follow any other syllable (many different sequences are produced), the conditional entropy will be high. We examined the data for all three adaptive protocol birds using this method and found that birds in a decrystallized state had a significantly higher conditional entropy than they did in their baseline state. Taken together, the time course of the baseline motif probability and the conditional entropy shows that the bird in Figure 3.8 went from being a low variance singer

with essentially a single motif, to being a high variance singer who produced a large number of different motifs.

After the removal of the artificial feedback, the temporal and spectral variability in the songs of all three adaptive protocol birds slowly decreased over the following weeks and months. The songs eventually recrystallized and were again stable. Both the probability of the baseline motif and the conditional entropy of the song returned to their baseline levels. The occurrences of stuttering, abnormal sequencing of syllables, and modified spectral organization gradually became infrequent and they were replaced by the temporal and spectral organizations characteristic of the original song. A complete recovery took about two to four months. The syllable-triggered bird made a partial recovery by 8 months after the cessation of feedback. The slow progression of these changes suggests that auditory feedback does not exert a great deal of instantaneous control over the production of song, but instead has a cumulative effect on song maintenance. The recrystallized songs appeared to remain stable indefinitely. We tracked one bird for a year after his recovery and saw no departures from the baseline song structure.

3.3 Discussion

Our results demonstrate that zebra finches need auditory feedback to maintain their songs in adulthood. This species, which does not modify its song or learn new songs after crystallization, appears to retain a great deal of plasticity in its auditory-vocal control system. This plasticity is sufficient to produce modifications in both the temporal pattern of song organization and the spectral structure of individual syllables. This finding is not consistent with the classical depiction of song development in which a dynamic learning period in youth ends in a static maintenance period in adulthood. Thus, the distinction between age-limited and open-ended learners may not be as sharp as these names imply. Furthermore, the recrystallization of the bird's original song shows clearly that, despite the destabilization of the behavior, a memory of the original song remains which may be recovered upon restoration of normal auditory feedback. This indicates that the song is maintained by an active control process, and therefore, somewhere in the song control system, we would expect to find neurons which are sensitive to perturbations of auditory feedback.

3.4 Methods

Birds were housed in custom designed plexiglass cages and were paired with a female who lived in a separate partition of the cage. The playback speaker's output was calibrated to be approximately the same as the sound level of the bird's vocalizations (80-90 dB SPL in the bird's ear). At the start of the baseline song recordings, the birds ranged from 130 to 300 days old (mean age was 200 days; zebra finches reach adulthood at 90 days). The singing rates of the different birds ranged from tens to thousands of motifs per night. However, there was no apparent correlation of age or singing volume with the magnitude of the effects we observed. For the two birds who did not sing frequently, we collected no baseline data during the feedback period.

Adaptive Feedback Perturbation. Microphone data were sampled at 40 kHz, after being low-pass filtered (10 kHz cutoff, 7-pole anti-aliasing filter). As is shown in Figure 3.2, the computer continuously acquired data from the microphone in segments of 100 msec. Each of these "bins" of data was passed through a software-based infinite-impulse response (IIR) filter (the "trigger filter"), that was used to detect song vocalizations while avoiding noise artifacts (such as pecking, wing flaps, and low amplitude calls). The playback of sound occurred when the output of the trigger filter exceeded an RMS threshold. The artificial feedback was thus produced with a 100 msec delay, and coincided with a silent interval in the bird's song or with the following syllable, depending on the timing of the song with respect to the bin borders created by the computer. The alternation between recording vocalizations and playing sounds prevented a positive feedback loop from developing. The playback stimulus was constructed from the 100 msec sound that had caused the trigger event, and was narrowband filtered before being played back to the bird. A narrowband filter was chosen to make the playback sounds difficult to localize. For one bird a wideband IIR filter was used. No significant differences in results were observed between this bird and the two narrowband birds. Delayed feedback was played to the birds, rather than other interfering stimuli, in order to replicate the spectral structure of auditory feedback that these birds normally hear. The effects of white noise will be examined in future work.

The birds exposed to the adaptive protocol showed changes in their songs similar to those of deafened birds. Because deafening can be caused by prolonged exposure to

excessively loud sounds (Wooley and Rubel, 1997), we used a behavioral test to demonstrate that the conditions used in our experiment did not cause deafness. Zebra finches respond to sounds by vocalizing. At the beginning and the end of the feedback period, we examined differences in the calling probability of two of the adaptive protocol birds to the playback of quiet sounds versus no sounds. The sounds were a variety of natural stimuli including conspecific calls and songs. Both birds produced significantly more calls during the presentation of quiet sounds ($p < 0.001$, generalized-likelihood-ratio test for Bernoulli random variables), which indicates that the birds could hear the sounds.

Syllable-Triggered Perturbation. Recognition of the target syllable was achieved by using a series of IIR filters in conjunction with each other to perform a logical operation (e.g., power in band X and not in band Y). The triggering was based on a small segment of the time-varying spectrum of the target syllable, which was unique to that syllable. An original copy of the crystallized trigger syllable was used as the playback stimulus for the duration of the experiment. Typical zebra finch syllables are 80-150 msec in length. The triggering resolution was 50 msec, which was short enough to ensure that the feedback always overlapped the trigger syllable itself (and partially on the following syllable). The other syllables of the song received no feedback.

Spectral Analysis. We calculated the time-frequency spectrogram for each song with a sliding window (5-8 msec), in which each time point consisted of the direct multitaper estimate of the power spectrum (with a time-bandwidth product $NW = 3$ or 4; Ho et al., 1998). The spectrograms shown in all of the figures were analyzed in this manner. For the analysis associated with Figure 3.7, a harmonic analysis was then used to determine the location and magnitude of the jumps in the discrete spectrum of each syllable (Thomson, 1982) based on the multitapered spectral estimates. This analysis revealed additional statistically significant harmonic frequencies in the target syllable after the feedback was presented to the bird ($p < 0.01$). An example of these results is shown in Figure 3.7.

Syllable Classification. For each syllable, we extracted the length and a number of time-varying parameters based on the spectral analysis described above. A modified K-means clustering algorithm (Selim and Ismail, 1982) was then used to partition the syllables produced on a given day into subsets. These subsets were labeled by the experimenter (syllable A, syllable B, etc.; labeling was done blind to the day on which the data were

acquired). For each day of data, approximately 1000 syllables were analyzed. The standard deviations that are shown as error bars in Figure 3.8 were obtained by bootstrapping the probability estimates from the data (Bradley, 1993). The steps underlying this process are described in detail in the following two paragraphs.

For each day of data, an automatic algorithm was used to extract the individual syllables from all of the song files. Syllables were identified in the time-domain as events whose power crossed an RMS threshold (with respect to background activity). For each syllable, we then calculated the length and a number of time-varying parameters (envelope, peak frequency, pitch, goodness-of-pitch, and Wiener entropy) based on the spectral analysis techniques described above. Many previous classification methods for birdsong syllables and songs have been based on direct comparisons (e.g., cross-correlations) of the time-frequency spectrograms of the signals in question. However, the entire spectrum contains so much information that it is difficult to make accurate comparisons of two signals in a reasonable amount of computation time. Furthermore, it is highly unlikely that the bird is actively controlling every single feature contained in the song. Recent experimental work (Fee et al., 1998) has shown that by modulating only a few control parameters, the syrinx can produce an enormous range of vocalizations. This result forms the basis of our analytical methods. We represented each syllable of the bird's song with a few time-varying parameters, which were intended to be analogous to features of the song that were under direct control by the bird. These features are briefly described in the next paragraph.

It is well known the dorsal muscles of the syrinx can rapidly close and open the syrinx during song, producing an amplitude modulation of the signal. This is reflected in the envelope of each syllable, which we calculated via the Hilbert Transform, and then lowpass filtered to 250 Hz (an upper bound on the maximum rate at which the bird could add AM modulation to the signal). The ventral muscles of the syrinx have a role in frequency modulation and in controlling the fundamental frequency of the sound being produced by the syrinx. We estimated the pitch of each syllable using the peak of the cepstrum. Cepstral analysis is a standard technique in the speech community for extracting the pitch of a signal (Kadambe and Boudreauxbartels, 1992). The cepstrum is defined at the spectrum of the log spectrum of a signal. In brief, if a signal has a lot of pitch (like a harmonic stack), frequencies that are spaced at integer multiples of the fundamental frequency will contain

more power than the rest of the signal. The cepstrum of this signal will have a peak at the fundamental frequency of the harmonic stack. The magnitude of this peak can be used as an estimate of the goodness-of-pitch, that is, how "pitchy" the sound really was. The Wiener entropy is a measure of the predictability of the signal (Ho et al., 1998). Syllables that are very tonal will have a very low Wiener entropy (highly predictable) whereas those that are white will have a very high Wiener entropy (highly unpredictable). The Wiener entropy most likely reflects a number of parameters that are under control of the bird. Like the other time-varying features discussed, it has a very characteristic waveform for each syllable that the bird produces.

Based on these features, a modified K-means clustering algorithm was then used to partition the syllables produced on a given day into subsets by iteratively minimizing the distance between all syllables and their respective cluster centers. K cluster centers were chosen randomly from the entire set of syllables. Each syllable was then assigned to the nearest cluster based on its distance from the cluster centers. Because the lengths of the syllables we observed typically varied from 30 to 250 msec, the lengths of the different feature vectors also varied. Because of this, the Euclidean distance was not a suitable metric for comparing two syllables. Instead, we used a symmetric dynamic time-warping (DTW) algorithm to compute the distance between two syllables (Casacuberta et al., 1987). The DTW algorithm is another standard tool in the speech community and finds the optimal nonlinear alignment between the two vocalizations (Myers et al., 1980). This compensated for differences in length between different syllables and differences in phase between different iterations of the same syllable (which would wreak havoc with the Euclidean distance). Once each syllable was assigned to a cluster, the cluster centroids were recomputed (based on the new cluster memberships) and the process was started over. Iteration continued until the system of clusters converged and the centroids became stable. The resultant syllable clusters were then labeled by the experimenter, and converted into text strings from which the various probabilities were be computed (e.g., the probability of the bird singing his baseline syllable sequence).

Markov Chain and Conditional Entropy. A Markov chain for the bird's song was constructed for each day of data from the K-means classified syllables. Each state of the chain represented a single syllable in the song, and the bird's song over the course of the

experiment was described as a Markov chain which evolved in time. In this framework, the likelihood of singing a particular syllable depended only on the occurrence of the last syllable produced, and not on any prior syllables. We have found that many zebra finch songs can be modeled in this manner using only first-order transition probabilities between syllables (i.e., there is no significant difference between the raw empirical estimate of the probability for some string of syllables and the Markov estimate using only the first-order transition probabilities). This demonstrates that the temporal patterning in the bird's song has a fairly simple statistical structure. In addition, the conditional entropy of the Markov chain can be used as an estimate of the variability of the song on a particular day. Estimating the variability of a particular animals vocalizations has in the past been a problematic issue for the birdsong community. The conditional entropy provides a nice solution to this problem. The conditional entropy of the Markov chain is defined as

$$H(Y \mid X) = -\sum_{i} p(x_i)\, p(y_i \mid x_i) \log(p(x_i \mid y_i)$$

where the summation is taken over all y and x (i.e., the entire set of syllable transition probabilities). P(x) is the probability of seeing syllable x at any point in the song. H(Y|X) measures the uncertainty in observing syllable Y, given that syllable X was just produced. If each syllable was followed only by a single other syllable (i.e., a single type of syllable sequence is produced), then the conditional entropy will be low, whereas if any syllable was followed any other syllable (many different sequences are produced), the conditional entropy will be high.

Chapter 4 : An efference copy may be used to maintain the stability of adult birdsong

Zebra finches use auditory feedback to both learn and maintain their songs (Konishi, 1965; Leonardo and Konishi, 1999). Nucleus LMAN of the anterior neostriatum is believed to be crucial for these processes, and is thought to convey an error-correction signal to the motor control system based on the degree of match between the bird's vocalizations and a memorized song template (Brainard and Doupe, 2000a). We measured the activity of individual LMAN neurons while simultaneously manipulating the auditory feedback that birds heard during singing, thus controlling the level of error they could detect in their songs. LMAN neurons were found to produce spikes locked with millisecond precision to specific acoustic features in individual song syllables. This timing precision is comparable to that seen in the motor control neurons that generate the song itself (Chi and Margoliash, 2001). Furthermore, perturbation of the auditory feedback heard by singing birds had no effect on LMAN spike patterns, suggesting that rather than auditory feedback, nucleus LMAN processes an efference copy of the bird's motor commands. These findings cast a new light on the role LMAN plays in the learning and maintenance of the bird's song.

4.1 Introduction

Error-correction plays a critical role in many biological and man-made circuits. Feedback-based error-correction is often important for learning (Hertz e. al., 1991), as well as for robustness and stability against the degradation and perturbation of control signals by noise. A neurobiological implementation of these processes is contained in the zebra finch song control system. Juvenile birds memorize the song of a tutor and then gradually match their own vocalizations to this template (Konishi, 1963). Systematic correction of errors in the bird's own vocalizations, driven by auditory feedback, is essential to learn this complex behavior (Konishi, 1965). Auditory feedback remains critical for song maintenance in adult

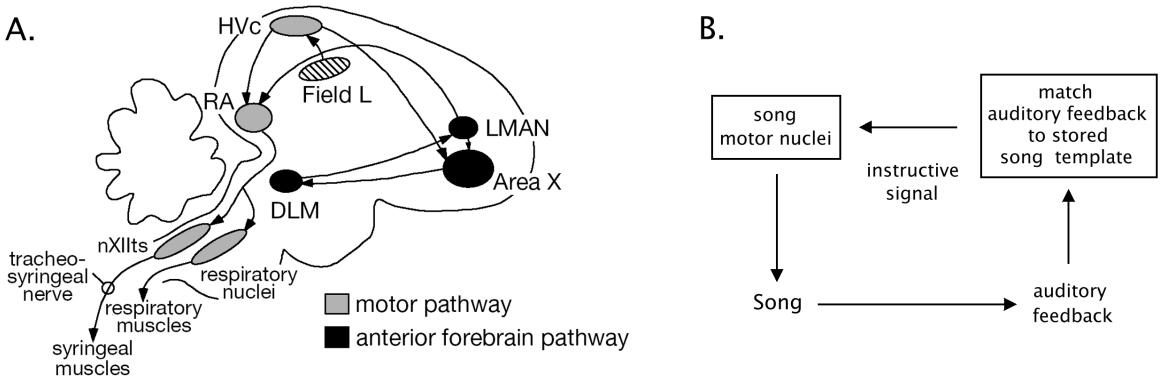

Figure 4.1. Zebra finch song control system and error-correction model. A) Schematic diagram of the zebra finch brain. The motor control pathway, from HVc to the syrinx, is shown in gray. The anterior forebrain pathway is shown in black. B) Error correction model of birdsong. The bird listens to his own vocalizations, compares them to an internal model, and uses this error signal to instruct the motor control program. *Figure 4.1a kindly provided by Allison Doupe, UCSF.*

birds (Nordeen and Nordeen, 1992; Leonardo and Konishi, 1999). The concerted activity of set of discrete brain nuclei, known collectively as the song system, generates this behavior (Figure 4.1a), and can be divided into two basic circuits: a pathway from HVc (the High Vocal Center; Margoliash et al., 1994) to RA (robust nucleus of the archistriatum; Nottebohm et al., 1976), that controls the instantaneous temporal and spectral structure in the song (Yu and Margoliash, 1996; Leonardo and Fee, 2002), and a pathway from HVc through the anterior forebrain, that is involved in the slower processes of song learning and song maintenance (Scharff and Nottebohm, 1991). Nucleus LMAN (lateral magnocellular nucleus of the anterior neostriatum, Arnold et al., 1976; Bottjer et al., 1989) generates the output of the anterior forebrain and projects directly back to the motor control system (RA). LMAN is thus ideally situated to process auditory information and relay it back to the motor system, and is thought to be intimately involved in the error-correction circuit used in song learning (Bottjer et al., 1984; Brainard and Doupe, 2000b)

The error-correction model of birdsong is based on auditory feedback: the bird sings, hears his own vocalizations, computes an error signal based on the match between the auditory feedback and an internal song model, and then uses this error signal to update the motor program (Konishi, 1965; Brainard and Doupe, 2000b; Figure 4.1b). Behavioral and

lesions studies strongly suggest that LMAN is the link between the auditory and motor system. Although LMAN is not required for song production in adult birds (Bottjer et al., 1984), lesioning LMAN prevents successful song learning in juvenile birds(Scharff and Nottebohm, 1991), and prevents the regeneration of song in adult birds that learn seasonally (e.g., white-crowned sparrows; Benton et al., 1998). LMAN is thus important for song stability throughout the bird's life. The completion of song learning is called crystallization; after this stage, song normally shows little variation in its spectral or temporal properties (Marler, 1970). However, if auditory feedback is removed by deafening during adulthood, song slowly deteriorates (Nordeen and Nordeen, 1992). If LMAN is lesioned when the birds are deafened, their songs remain relatively stable (Brainard and Doupe, 2000). There is thus a direct correlation with the presence or absence of LMAN and song plasticity. We have shown in previous work (Leonardo and Konishi, 1999) that the stable songs of adult zebra finches can be disrupted by controlling in real-time the auditory feedback that they hear while singing. Restoration of normal auditory feedback to these birds enables the recovery of their original songs, indicating that the auditory feedback-driven plasticity we observe in the song system is the result of an active control process and is not simply passive drift of an internal representation. This chapter investigates the role of nucleus LMAN in the real-time processing of auditory feedback during singing.

Despite the clear connections between LMAN lesions, auditory feedback, and song plasticity, attempts to measure the specific information that LMAN relays to the motor control system have been ambiguous. Studies carried out on anesthetized birds have shown that LMAN neurons are highly tuned to the bird's own song (BOS; Figure 4.2), and respond maximally to this stimulus and little to other sounds (Doupe and Konishi, 1991). The presence of BOS tuning and the auditory sensitivity of LMAN neurons have led to the hypothesis that LMAN allows auditory feedback to influence RA activity and thereby modifies the bird's vocalizations to more closely match an internal song model (Doupe, 1997). It should be noted that although LMAN has become the focus of research investigating the specific role of auditory feedback in the song control system, song-selective auditory neurons are found in all of the song control nuclei, including HVc (Margoliash, 1983), RA (Vicario and Yohay, 1993) and the motorneurons that drive the syrinx (Williams and Nottebohm, 1985).

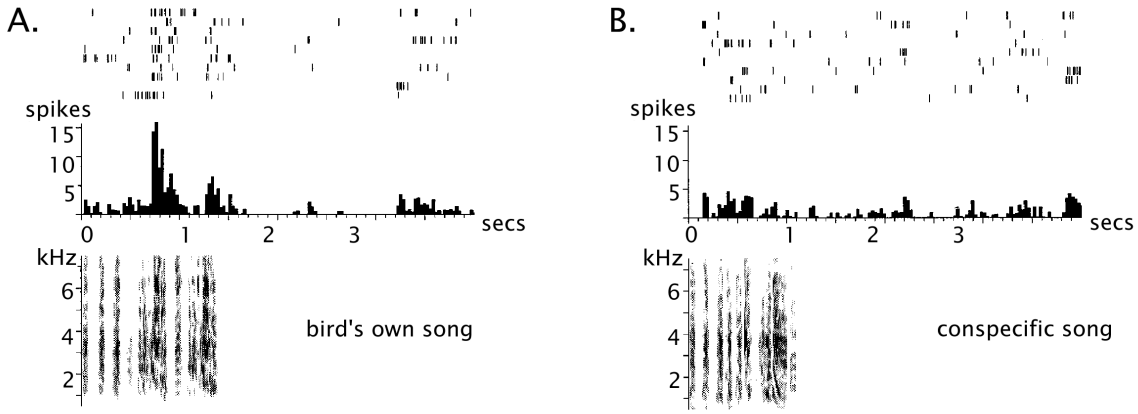

Figure 4.2. Song selectivity of anesthetized LMAN neurons. A) Extracellular response of a single LMAN neuron in a urethane anesthetized bird to playback of the bird's own song (BOS). Shown, from top to bottom, is the spike raster for each trial of sound playback, the average firing rate of the neuron, and the spectrogram of the BOS. B) Response of the same anesthetized LMAN neuron to the conspecific song of another zebra finch (CON). Note the large and reliable response which occurs during the playback of the BOS, and the relatively small response to playback of the CON, despite the similarity between the two songs. *Figure kindly provided by Allison Doupe, UCSF.*

Work by Hessler and Doupe (1999b) in the awake singing bird suggests that some component of LMAN responses could be motor-driven, based on the observations that neural signals increase in amplitude before the bird begins singing, and deafening the bird does not appear to alter LMAN activity. While compelling, these data are primarily multi-unit signals and thus lacking in temporal resolution. This makes it difficult to determine if the increase in LMAN activity pre-song is motor-based or simply reflects a general increase in LMAN excitability before the song begins. Further, changes in the timing of individual neurons after deafening are difficult to detect in multi-unit recordings. It is generally agreed that in order to understand the function of LMAN, it is necessary to record from single LMAN neurons while simultaneously manipulating the auditory feedback that the bird hears during singing (Brainard and Doupe, 2000b; Margoliash, 1997; Carr, 2000). If LMAN is sensitive to auditory feedback, single neuron spike patterns are expected to change during singing when auditory feedback is perturbed.

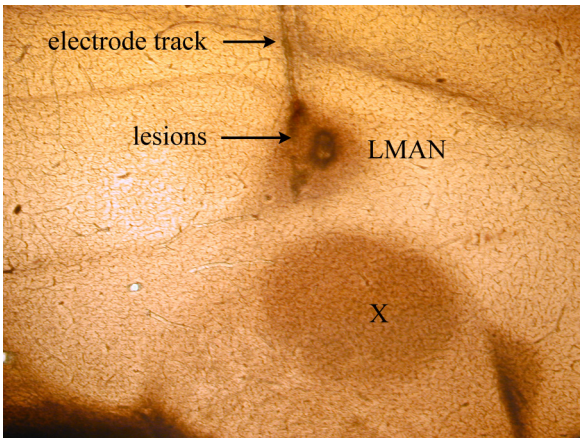

Figure 4.3. Histological reconstruction of motorized microdrive recording sites in LMAN. Two electrolytic lesions, and the electrode track, can be seen within nucleus LMAN. To the left of nucleus X and below are the fiber tracts running from X to the thalamic nucleus DLM, and from DLM to LMAN. The diameter of nucleus LMAN is approximately 500 um.

We developed a computer-controlled system to perturb the auditory feedback heard by singing birds. In combination with this, a miniature motorized microdrive (Fee and Leonardo, 2001) was used to measure the activity of LMAN neurons in the freely moving zebra finch. Upon isolation of a single LMAN neuron (Figure 4.3), each zebra finch was induced to sing by repeated presentation of one or more female birds (so called directed song). A computer monitored the song in real-time, and either generated artificial auditory feedback or allowed the birds to sing normally. Auditory feedback was produced from a wall-mounted speaker in the bird's cage, causing the bird to hear a superposition of his own vocalizations and the artificial feedback. Two different types of feedback were used - white noise and delayed song syllables - and were played to the bird continuously during the song-triggered feedback trials (Figure 4.4). As no difference in results was found between these two feedback conditions, we treat them as a single group in the analysis that follows. Both of these types of feedback are known to cause song deterioration if they are played every time the bird sings for 1-4 weeks (Leonardo and Konishi, 1999; Leonardo, unpublished observations on white noise feedback). However, the random interleaving of normal and feedback perturbation trials contained sufficient normal song that the bird's vocalizations were not expected to change over the duration of the one week experiment. Stability of the motor system was critical to the design of our experiment, which assumed that auditory feedback was the only variable being manipulated and all of the other mechanisms of song control were functioning normally. We ensured that the bird’s songs did not deteriorate by implementing an active sound cancellation system that allowed us to remove the artificial feedback from the recorded microphone signals and examine in detail the spectral signatures

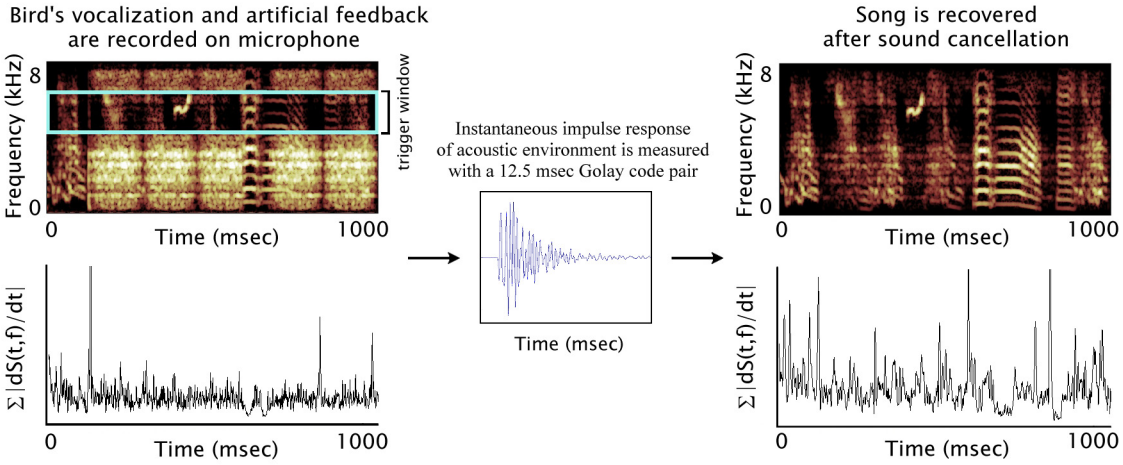

Figure 4.4. Auditory feedback perturbation system and active sound cancellation system. The computer continuously triggers on a narrow frequency band of sound with 50 msec temporal resolution. White noise modulated with a 7 Hz envelope is played to the bird for as long as song is detected. The microphone records the superposition of the bird's vocalization and the artificial auditory feedback, a signal roughly analogous to what the bird hears. For each motif of song, the instantaneous impulse response of the acoustic environment (speaker, cage, moving bird, and microphone) is measured and then used to predict and cancel the frequency structure of the artificial feedback on the microphone, allowing recovery of the bird's original vocalization. Shown in the figure is the spectrogram of the microphone signal (top left and right) and its time-derivative (bottom left and right), before and after sound cancellation. The time-derivative of the recovered vocalization is used both to confirm song stability and to align the simultaneously recorded LMAN spike trains.

of the songs produced by the birds during singing with altered auditory feedback (Figure 4.4; see Methods). Comparison of songs produced during feedback trials, baseline trials, and baseline trials obtained before implantation of the microdrive confirmed the stability of the songs of all the birds used in the experiment. With this protocol, we were thus able to measure the activities of individual LMAN neurons when the bird sang normally, and when he sang with altered auditory feedback. Any changes in LMAN activity that occurred during the presentation of altered auditory feedback would be due solely to real-time sensitivity to auditory feedback and not to instantaneous feedback-induced changes in the song (i.e., changes in the motor program), which have been observed to occur in human speech during the presentation of altered auditory feedback (Houde and Jordan, 1998).

4.2 Results

We recorded from 31 LMAN neurons in three zebra finches, and analyzed their activity during the simultaneously recorded songs (Figure 4.5). The individual sound elements of zebra finch song are called syllables and contain distinct spectral features. Syllables are produced in a fixed sequence known as a motif, and each time the bird sings the motif is repeated a variable number of times (Sossinka and Bohner, 1980). In order to analyze the song-dependent activity of each LMAN neuron, we must average its spike activity produced during different song motifs. However, as is the case in human speech, the lengths of the individual syllables in the song vary independently of each other, and are randomly stretched and compressed by approximately 5%. Using the methods of Leonardo and Fee (2002), we estimated the magnitude of this acoustic time-warping and compensated for it appropriately in each syllable produced by the bird. The simultaneously recorded neurons were then projected onto this aligned acoustic time-axis, so that structure observed in the aligned spike trains was due solely to their correlation with structure in the song (and not to an artificial imposition of structure by the alignment algorithm). Previous reports on LMAN neural responses in the singing bird have concentrated on the analysis of multi-unit data (Hessler

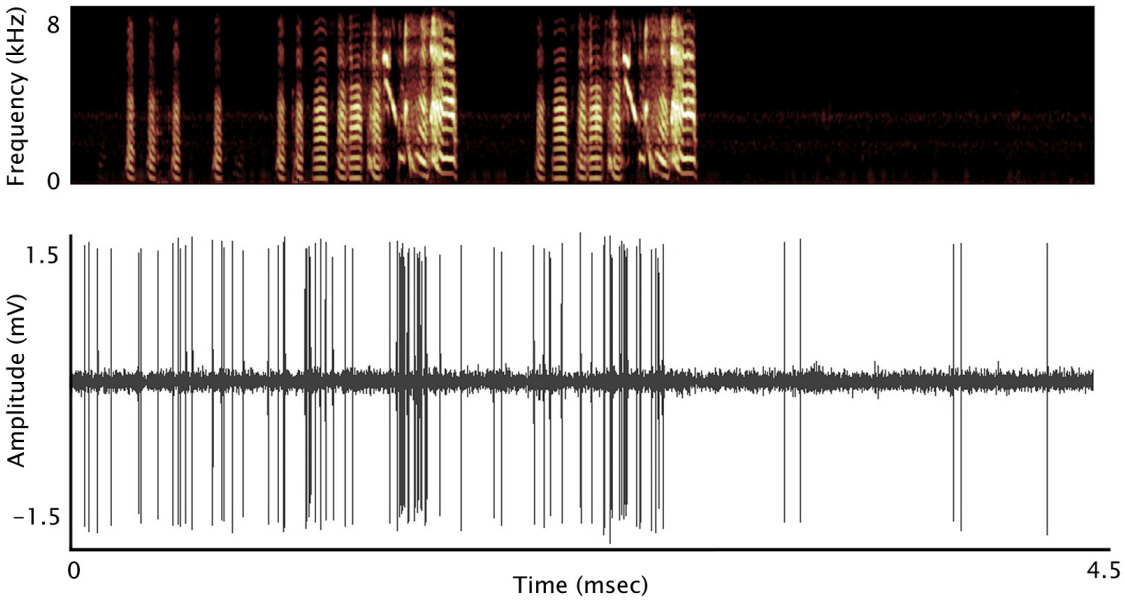

Figure 4.5. Recording from a single LMAN neuron during two motifs of singing.

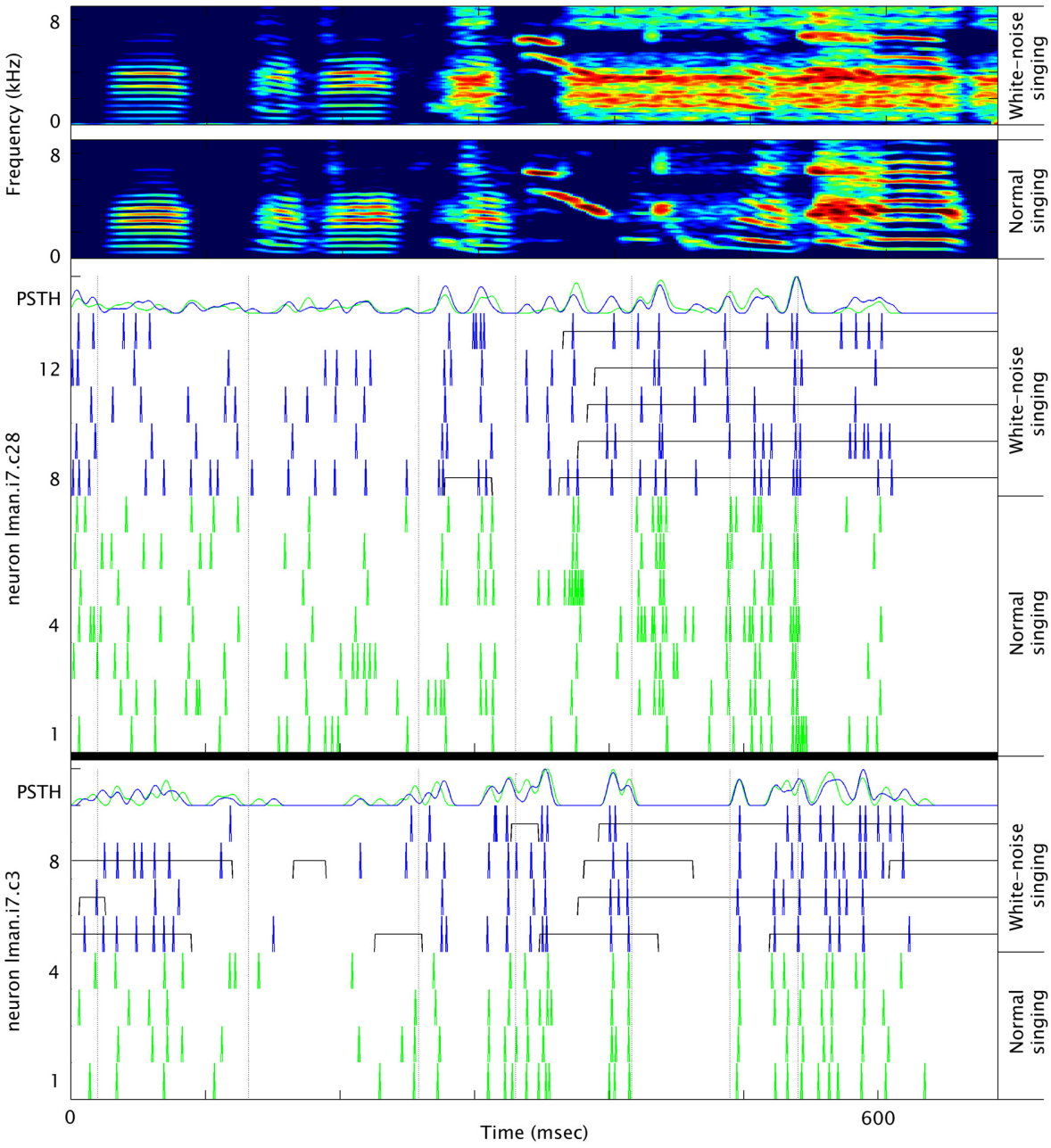

Figure 4.6. Song and spike train alignment for two LMAN neurons. Spikes produced during normal singing are show in green, and spikes produced during white noise feedback singing are shown in blue. The black square-wave lines running through the feedback trials represent the exact locations of the perturbing auditory feedback produced during those songs. Black dashed lines running vertically through each raster represent the location of the syllable alignment points. Note, for each neuron, certain spikes appear reliably locked to particular time points in the bird's song.

and Doupe, 1999), and have relied on a single syllable onset as a reference point for time-aligning the songs and neural signals. This results in considerable variability of the aligned neural signals at time-points far from the reference point, due to the syllable time-warping discussed above. For these reasons, it has been difficult to estimate how precisely LMAN activity is locked to acoustic features in the bird's song, and the general impression in the literature has been that LMAN neurons are considerably more variable than neurons in other nuclei in the song control system (Hessler and Doupe, 1999b; McCasland, 1987). However, as we describe in the next section, alignment of the songs based on multiple of acoustic features reveals considerable song-locked structure in the spike trains of most LMAN neurons.

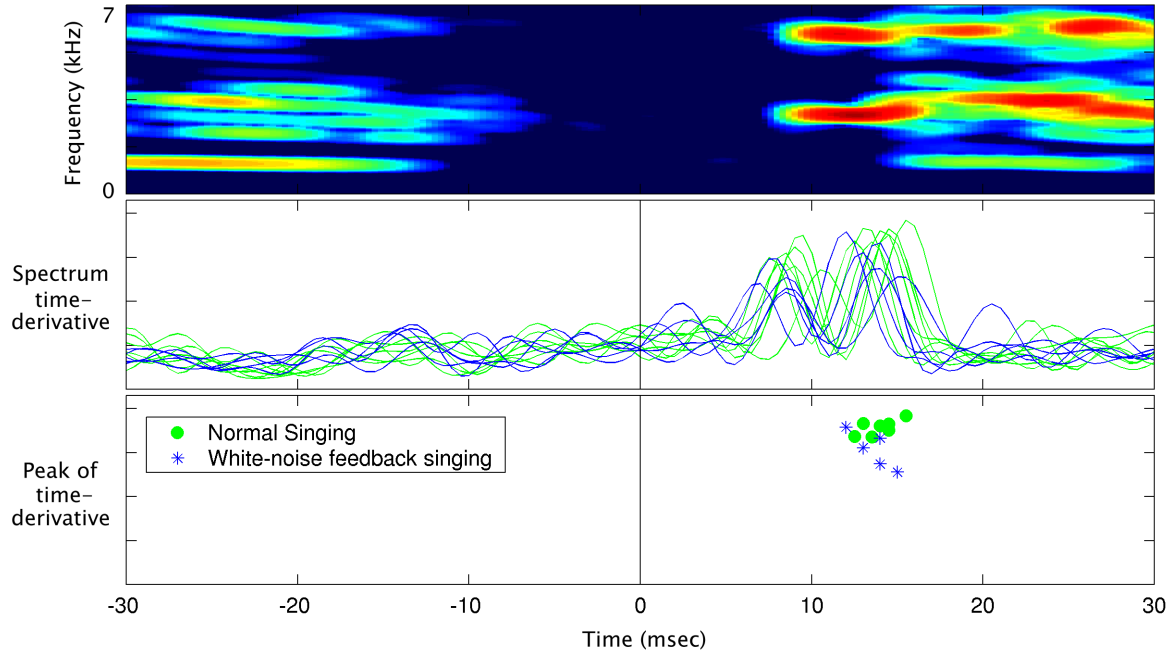

Figure 4.7. LMAN neurons fire spikes precisely timed to the bird's song. Top trace: time-frequency spectrogram of a 80 msec sound from the bird's song. Middle trace: for each motif of song recorded for neuron i7_c28 (Figure 4.6, top panel), we find the time of the first spike occurring in a window from 540-580 msec. We then extract an 60 msec window of song around this spike and estimate its time-derivative, which is plotted in green for trial with normal auditory feedback, and blue for trials with altered auditory feedback (the trigger spike occurs at t=0). Note the bimodal peak in the time-derivative around 10 msec, indicating the onset and offset of the narrowband burst of sound seen in the spectrogram above. Bottom trace: each point marks the location of the second peak of the time-derivative for one motif of singing. The standard deviation in position of the normal song peaks (green) is 1.02 msec and is not significantly different from that of the feedback song peaks (blue, 1.14 msec).

After song-alignment, we found that 26 of 31 LMAN neurons produced spikes that were reliably and precisely locked to time points in the bird's song (Figure 4.6). These neurons exhibited sharp peaks in firing rate, as can been seen in the average firing rate profile of the spike train rasters. The 5 remaining neurons showed a general increase in firing rate during song, but lacked any modulation clearly tuned to particular time points in the song. In order to quantify how precisely LMAN neurons could be locked to structure in the bird's song, we measured the jitter in the alignment of an acoustic feature, triggered on the spiking of an LMAN neuron at a specific point in song. The acoustic feature we used for this was the time derivative of the song spectrogram, summed across frequencies (Tchernichovski et al., 2000). The time-derivative has sharp, reproducible peaks at locations in the song where sound frequencies change rapidly (e.g., syllable onsets, or changes in syllable structure). When an LMAN neurons reliably spiked at some location in the song, we were able to line up these spikes (instead of lining up the syllables) and measure the associated jitter in a nearby peak of the song's time-derivative. Our goal here was not to quantify the precision of every LMAN spike, but rather to measure how accurately the most precise LMAN neurons were locked to the features in the bird's song. Surprisingly, we found that many LMAN neurons fired spikes that were locked to the song with approximately 1 msec of variance (Figure 4.7). This is comparable to the precision with which pre-motor neurons in RA are locked to the acoustic features they generate (Chi and Margoliash, 2001). Furthermore, many of these LMAN spikes occurred at the onset of a motif, before singing began, or at the onset of a syllable. Given that the latency for auditory feedback to reach LMAN is on the order to 25 msec, this implies that the auditory feedback that drove a syllable-onset spike would have occurred during the silent inter-syllable-interval. The observation that LMAN neurons can be locked to the song as accurately as RA pre-motor neurons raises the possibility that motor-input, and not auditory feedback, is responsible for generating some portion of the LMAN firing patterns.

By examining how precisely LMAN spikes were locked to acoustic features during singing with perturbed auditory feedback, we can determine if these spikes are generated by auditory or motor input. Neurons whose spike patterns were driven by auditory feedback should show increased variability in spike timing when the bird sang with perturbed auditory feedback, as the acoustic features driving these spike would be distorted by the artificial

feedback. Figure 4.7 shows the acoustic feature locking of a single LMAN neuron during the normal and perturbed auditory feedback singing conditions. No significant difference in spike feature locking was found between these two conditions across the set of precisely timed spikes we analyzed ($n=6$, $p < 0.01$, two-sample t-test). Spike-triggered acoustic feature locking continued to have 1 msec of jitter, despite the fact that the acoustic features were entirely obscured by white-noise feedback. Thus, at the very least, some LMAN spikes are not driven by auditory feedback, and are likely to have a motor basis.

We now rigorously quantify the difference between the entire spike trains produced by LMAN neurons during normal and perturbed auditory feedback singing, using methods from information theory. The spike-triggered acoustic feature analysis we described above is a convenient way of demonstrating qualitatively the degree of sensitivity of some spikes in the song to the presence of auditory feedback. However, many LMAN spikes were not produced with sufficient precision to allow such an analysis. We used two related methods to compare the entire spike trains of LMAN neurons during normal and perturbed feedback singing. First, we compared the two conditions using the d' statistic used to compare the song-selectivity of different neurons in anesthetized LMAN recordings. Then, in the following section, we compute the time-varying Kullback-Leibler information between normal and feedback singing spike trains, as a function of various window sizes. The results of these analyses show, quantitatively, that auditory-feedback did not induce changes into any portion of the spike patterns generated by LMAN neurons in the singing zebra finch.

Recall that BOS tuning (Figure 4.2) was one of the pieces of experimental data that has led to the hypothesis that LMAN processes auditory feedback and outputs the result of comparing how well the bird's vocalization matches an internal template. Under this hypothesis, vocalizations that poorly match the bird's template song, such as those heavily contaminated with perturbed auditory feedback, should produce LMAN spike trains that are substantially changed from the spike trains produced by the same neuron when the bird sings normally. The standard method for measuring the response selectivity for the BOS over other acoustic stimuli is the d' statistic (Solis and Doupe, 2000). d' measures the difference between the means of two Gaussian distributions, normalized by their standard deviations. In the case of the anesthetized LMAN studies, the d' value is an estimate of the discriminability between the total spike count produced by a neuron during BOS motifs

versus the total spike count produced by the same neuron during non-BOS motifs. In our data set, the motifs where the bird sang normally represented the BOS condition, and the motifs where the bird sang with altered auditory feedback represented the non-BOS condition. Of the 17 LMAN neurons from which we obtained sufficient samples of normal and perturbed feedback singing to calculate d', we found no significant difference in motif spike count ($p < 0.01$; see Methods). Thus, using the same statistic used to establish song selectivity in the anesthetized bird, we see no effects of auditory feedback on the total spike count produced by LMAN neurons in the singing bird. However, the d' results do not rule out the possibility of auditory feedback modulating of LMAN activity. The computer controlled auditory feedback does not trigger on all parts of the motif with the same degree of accuracy; this may increase the response variability in certain regions of the song without changing the mean neural activity. Further, it is possible that individual LMAN neurons code for errors to localized parts of song syllables and not the entire motif. Effects such as these could easily be missed by only examining changes in the motif spike count.

Small localized changes in LMAN activity patterns can be detected by examining the differences between normal and feedback singing spike trains in a window that slides across the motif. This could be done using the d' statistic, but the use of a different but related method will allow us to examine changes in both the mean and variance of the LMAN spike patterns. d' is actually the Gaussian case of a more general information theoretic metric know as the Kullback-Leibler information (Cover and Thomas, 1991; Johnson et al., 1999). The KL information measures the difference between two probability distributions, without making any a priori assumptions about the statistics underlying those distributions. If R represents a set of possible neural responses, and $p_1(R_i)$ and $p_2(R_i)$ represent the probabilities of observing response R_i given that the bird is singing normally (p_1) or with perturbed auditory feedback (p_2), the Kullback-Leibler information is given by

$$D(p_1 \| p_2) = \int p_1(R) \log \frac{p_1(R)}{p_2(R)} dR$$

The KL information is zero when the two distributions are identical, and nonzero and increasingly positive as the two distributions become increasingly different. We calculated the time-varying Kullback-Leibler information for LMAN spike trains produced during normal and perturbed feedback singing, using 1000 msec (whole motif), 100 msec, and 15

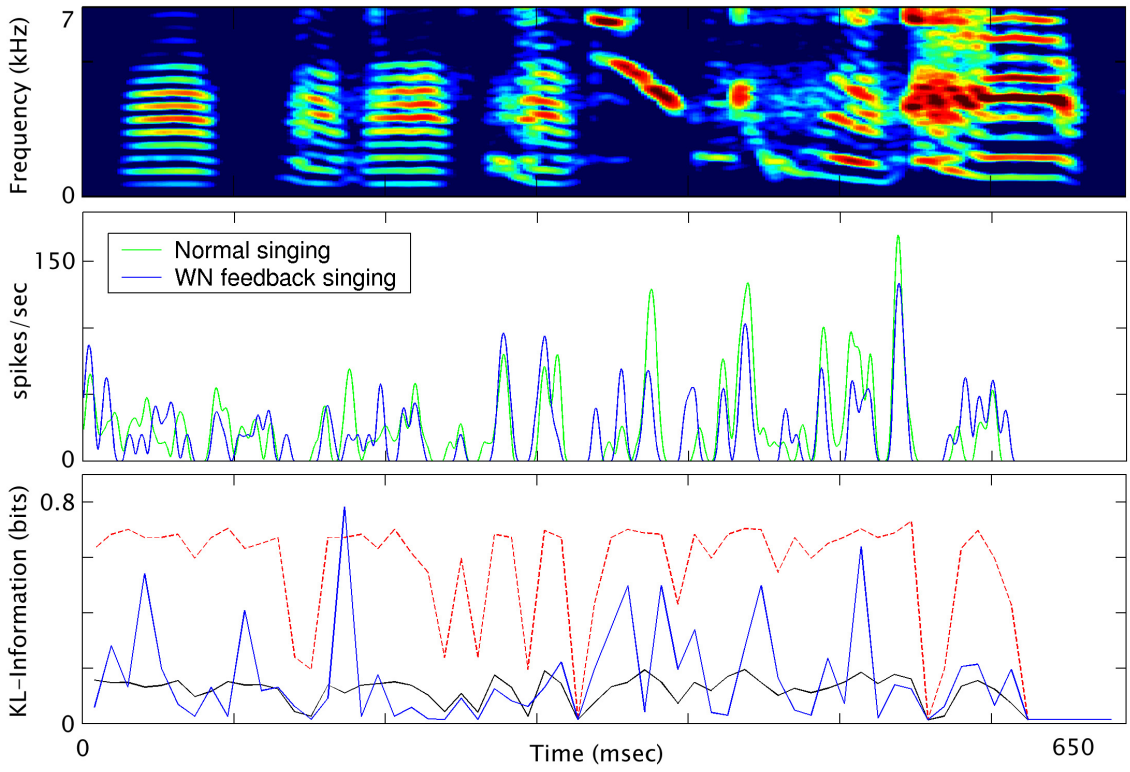

Figure 4.8. Time-varying Kullback-Leibler information for an LMAN neuron. Top: song spectrogram. Middle: average firing rate during normal (green) and perturbed auditory feedback (blue) singing for neuron i7.c28 (see Figure 4.6). Bottom: Kullback-Leibler information (blue), computed in a sliding 15 msec window, between normal and perturbed feedback singing conditions. The expected KL information, given that the neuron fires in the same way during normal and feedback singing (i.e., h0, the null hypothesis), is shown in black. Dashed red lines indicate the 99% confidence interval for the null hypothesis. One event exceeding the confidence level is expected by chance due to the multiple statistical comparisons made as the window steps through the motif. Such an event occurs at t=190 msec (note in Figure 4.6 that no feedback actually ever occurs at this time). As the total number of threshold events does not exceed chance levels, we conclude that the KL information is not significantly different from zero and that the LMAN neuron produces the same spike trains during normal and perturbed auditory feedback singing.

msec sliding windows. For each step of the window, we estimated the KL information between the probability distribution of spike counts during normal and perturbed singing. The use of these different windows enabled us to analyze the structure in LMAN spike trains over a wide range of time scales. Regardless of window size, we found that there was no sensitivity of LMAN neurons to auditory feedback during singing (n=17, $p < 0.01$, Figure 4.8; see Methods).

Consistent with our measurements in LMAN, single RA neurons (n=4) show no statistically significant changes in burst timing between normal and perturbed feedback singing based on the Kullback-Leibler information computed for a variety of window sizes ($p < 0.01$). Because all RA bursts are precisely timed to the bird's song, the lack of feedback sensitivity is more striking in RA than it is in LMAN (Figure 4.9). LMAN sends excitatory NMDA projections directly to RA pre-motor neurons (Rosen and Mooney, 2000), and any substantial changes in LMAN activity would be expected to alter RA firing patterns and thus alter vocal output. As we observe no real-time changes in vocal output when the bird is singing with perturbed auditory feedback, the LMAN and RA measurements are quite consistent with each other. This consistency between the lack of behavioral changes and the lack of neural changes is not trivial; although the behavioral effects of perturbed auditory feedback do not occur in real-time, auditory feedback must be registered in real-time somewhere in the song control system in order to be used offline to maintain the stability of the bird's song.

The measurements we have discussed so far were all made during directed song. Zebra finches sing two types of song: directed song, in which the bird is interacting with a

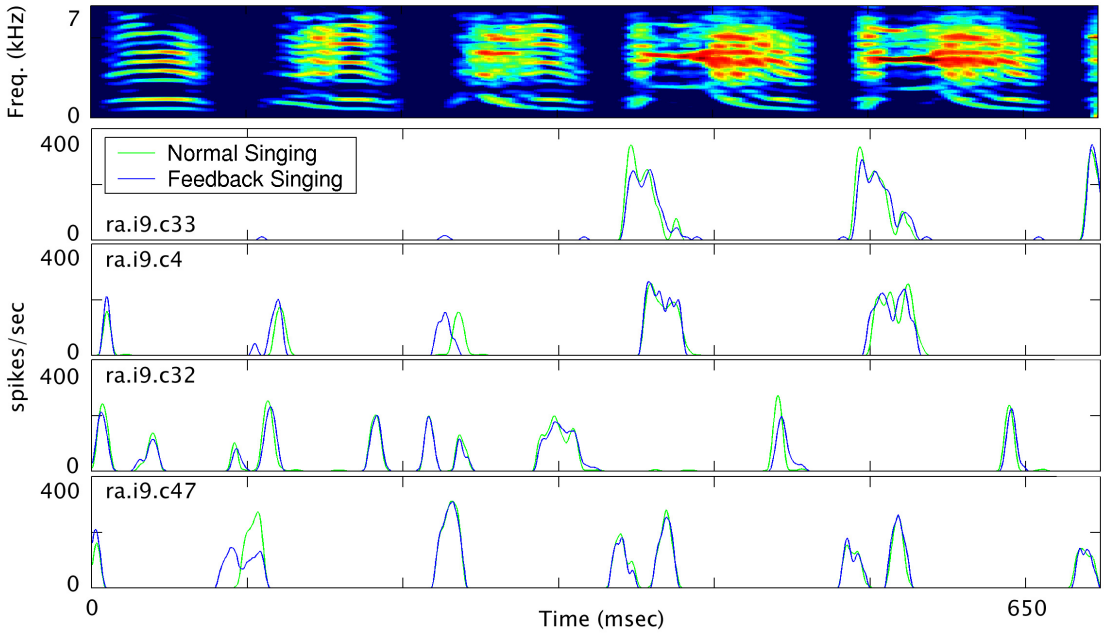

Figure 4.9. Average firing rate of four RA neurons during normal and perturbed auditory feedback singing. No significant feedback-induced changes in spike rate or pattern were found for any of the neurons.

female, and undirected song, in which the bird is singing to himself. Acoustically, directed and undirected songs are virtually identical (Sossinka and Bohner, 1980), and can only be distinguished vocally by the slightly faster delivery rate of directed song (~2%; Hessler and Doupe, 1999b). Physiologically, however, the two states are quite different. Previous reports have noted that LMAN multi-unit activity is substantially more variable during undirected song than directed song (Hessler and Doupe, 1999b), and different immediate-early genes are expressed in each state (Jarvis et al., 1998), leading to the hypothesis that directed and undirected songs serve two different, undetermined, behavioral purposes. The striking differences between these two states can be seen in our single neuron recordings. LMAN neurons frequently produced bursts of spikes during undirected song, whereas bursting rarely occurred during directed song. Furthermore, all of the LMAN neurons that we recorded during both directed and undirected song generated precise spike patterns during directed song and showed virtually no temporal structure during undirected song (n=3). An example of the transition in LMAN activity from directed to undirected singing is shown in Figure 4.10. Given that the song produced by the bird is essentially identical in directed and undirected song, this suggests that auditory feedback and motor efference should also be identical between the two states. How LMAN spike patterns can then be so precise during directed song, and so variable during undirected song, remains a mystery.

The experiments described here were not intended to measure LMAN activities during undirected song, but they do suggest a novel experiment through which the role of undirected song could be further explored. Undirected and directed songs are currently discriminated purely behaviorally, based on whether the bird is singing to another bird or to himself. This method is inaccurate, as it is difficult to visually distinguish when the bird is attending to another bird. Our LMAN recordings suggest that the two states can be distinguished physiologically, in real-time, by placing an electrode in LMAN and measuring the volume of bursting activity. As bursting in LMAN appears to be uniquely correlated with undirected song, this assay is likely to be highly accurate in distinguishing between the two song states. A detailed behavioral analysis of the songs generated in each state may then provide new insights into the purpose of undirected song. For example, the decrystallization protocol could triggered based on the presence or absence of bursting activity in LMAN. In this manner, perturbed auditory feedback could be delivered to one group of birds only

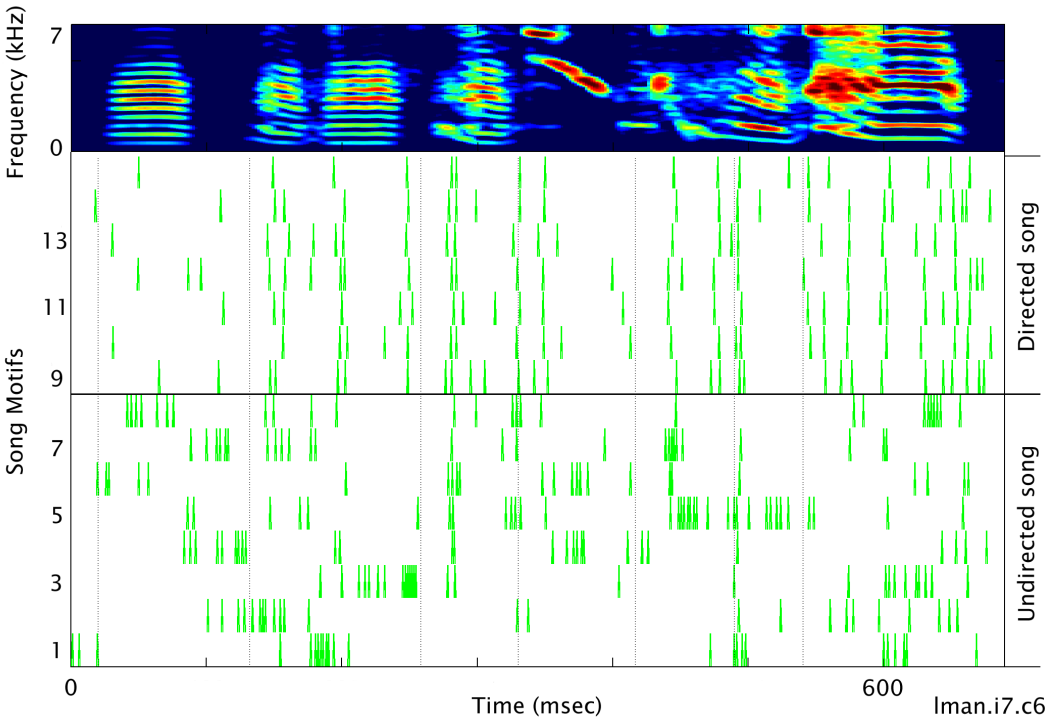

Figure 4.10. LMAN neural response during directed and undirected song. The first 8 spike trains show the spike activity of a single LMAN neuron during undirected song. Approximately one minute after the cessation of the 8th motif, a female bird was presented to the male and he immediately sang the following 7 motifs of directed song. The characteristic song-locked spike trains produced during directed song are almost entirely absent in the recordings from the same neuron during undirected song.

during directed singing, and to a second group of birds only during undirected singing. It would be interesting to determine if learning occurs during only one these two modes of singing.

Although our experiment primarily examined the activity of LMAN neurons during directed song, a small sample (n=2) of LMAN neurons was obtained during undirected song for singing with normal and perturbed auditory feedback. For these neurons, no significant different in spike rate or pattern was found between the two feedback conditions using the KL information test. Furthermore, previous work in LMAN by Hessler & Doupe (1999b) has shown that LMAN multi-unit activity during undirected song is unchanged by the complete removal of auditory feedback via deafening. Although physiologically some aspects of directed and undirected singing are quite different, LMAN neural activity shows no significant sensitivity to auditory feedback in either state.

4.3 Discussion

Based on the results we have described, we suggest that during directed singing, nucleus LMAN processes an efference copy of the motor commands for song. An efference copy is a record of the commands used to generate a motor output (Sperry, 1950; von Holst and Mittelstaedt, 1950; Bridgeman, 1995). Two observations form our argument that LMAN processes an efference copy. First, LMAN is not required for singing, but many LMAN neurons are locked to the song with the same millisecond precision seen in motor structures. Second, perturbation of the auditory feedback that the bird hears while singing will eventually destabilize the song, but has no real-time effect on the firing patterns of LMAN neurons. There are four possible signals that could produce the song-locked spike trains we observe in LMAN: an auditory feedback signal, a proprioceptive feedback signal from the vocal muscles, a motor command signal, or an efference copy. The Kullback-Leibler information analysis clearly shows that auditory feedback does not generate the spike patterns in LMAN. There is no known afference from the syringeal muscles back to the song control system (Bottjer and Arnold, 1984). LMAN can be lesioned without effecting the production of song (Bottjer et al., 1984), indicating that it does not generate a motor command. The only hypothesis that can explain our results is that LMAN activity is driven by an efference copy of the bird's song. No other signal could cause LMAN activity to have such a high degree of correlation to song structure while being immune to changes in auditory feedback. This conclusion represents a significant departure from the classical view of LMAN being primarily a processor of auditory feedback.

How could an efference copy be used to maintain the stability of the song control system? By itself, a copy of the motor commands used to generate the song contains no information about which sounds were produced correctly and which sounds were produced incorrectly. However, if the efference copy were compared to an internal model of the song, then a useful error-correction signal could be generated. Variations of an efference copy model of error-correction in the song control system have been suggested by others, most formally by Troyer and Doupe (2000; but see also Dave and Margoliash, 2000; Vates et al., 1997; Wild, 1993). However, the results described within this paper are the first experimental measurements to show the signature of an efference copy in LMAN spike

activity. Our results are consistent with the notion that LMAN represents the output (and possibly the computation) of an error-correction process based on an efference copy. Because the bird's song is not changing instantaneously in response to perturbed auditory feedback (Leonardo and Konishi, 1999), the efference copy does not change, and the LMAN spike patterns (and putative error-correction signal) are always the same when observed over a short time period. However, if the zebra finch were to sing a syllable incorrectly (i.e., send a different motor command to the syrinx), the efference copy would change, a large error signal would be generated, and this could be used to modulate the generation of spike patterns in RA so that the next time the bird sang the error would be reduced. Purely motor errors in song generation, detectable entirely independently of auditory feedback, are likely to occur. For example, when one is playing the piano, one often knows an incorrect key has been struck before the sound is heard, based only the information that the hand moved incorrectly. LMAN's role during singing might be to correct for motor errors such as these, by comparing the efference copy to the stored song template.

The manner in which auditory feedback is used to maintain the stability of the song control system remains an enigma. Changes in auditory feedback clearly drive plasticity in the motor control system and this plasticity is mediated by nucleus LMAN (Brainard and Doupe, 2000a). Our results are not in conflict with these observations, as they occur slowly whereas we were measuring the real-time effect of auditory feedback on LMAN spike patterns. However, a description of error-correction in the song control system which only depends on an efference copy is incomplete, as it leaves out the role of auditory feedback and thus cannot account for auditory-feedback induced changes of song. The original idea of an efference copy (Sperry, 1950) was specifically conceived of as a mechanism for canceling unwanted sensory reafference. For example, in mormyrid electric fish, an efference copy of the electric organ discharge (EOD) is used to generate a prediction of the incoming sensory reafference from the EOD. This prediction is then sent to the electrosensory lobe where it is used to cancel effects of the EOD from the incoming sensory signal, thereby allowing the fish to distinguish between its own reafferent input and input from external sensory sources (Bell, 1981). This feedback control is critical for active electrolocation. Could an efference copy be used in the song control system in a similar manner, generating a prediction of expected auditory feedback? One of the hallmark features of the efference copy system seen

in electric fish is that if the sensory feedback is artificially manipulated, the canceling effects of the efference copy rapidly adapt to cancel the new sensory feedback (Bell, 1982). In contrast, the spike patterns in LMAN show no sensitivity to perturbations of auditory feedback. To add a further twist to the puzzle, it is possible that auditory feedback never modulates the song control circuit when the bird is singing because it is gated during song to prevent the interaction of auditory and motor signals (McCasland and Konishi, 1981; Schmidt and Konishi, 1998). If this is the case, then the manner in which the efference copy and the sensory reafference interact must be fundamentally different than what has been seen in other systems.

In conclusion, we would suggest that the mechanisms through which auditory feedback maintains the stability of the song pattern generator are substantially more complex than that suggested by existing models. Auditory feedback has no real-time effect on neural activity in LMAN or RA, nor does it exert real-time effects on the vocal output generated by the bird. It may be the case that the auditory feedback is registered in the song control system (prior to LMAN) as the bird is singing, but it is only offline that it is used for song maintenance. If this is the case, LMAN could participate in two types of feedback control. When the bird is singing, the efference copy could be compared to the song template to correct motor control errors (i.e., activating the incorrect RA neurons). When the bird is not singing, stored auditory feedback could be regenerated and compared to the song template to correct for auditory control errors (i.e., expected RA neurons not generating the correct sound). Such a revised model of vocal control has considerable intuitive appeal; when one is speaking, one rarely modulates vocal output based on auditory feedback in real-time. Incorrectly pronounced words are noted, and the next time one speaks these errors are taken into consideration. The zebra finch may use an algorithm similar to this to maintain the stability of song.

4.4 Methods

Birds were housed in custom designed plexiglass cages, and had unlimited access to food and water. All birds used in the experiment were male zebra finches, approximately 120 days old (adults with crystallized songs). Female zebra finches were housed in similar, separate

plexiglass cages, and were presented to the males upon isolation of one or more LMAN neurons. Sound was played through a speaker mounted on one wall of the cage, and was recorded via an omnidirectional microphone mounted on a perpendicular cage wall. The amplitude of the artificial auditory feedback was calibrated to be approximately equal to the bird's own vocalizations upon reaching the bird's ear. The calibration was approximate because the bird was able to move freely in the cage. At the onset of each bout of song, the computer randomly chose, with equal probability, whether to allow the bird to sing normally or to produce the artificial auditory feedback.

Neurophysiology. Birds were anesthetized with 1-2% isoflurane and nucleus LMAN was identified with an extracellular targeting electrode based on stereotaxic coordinates and physiological activity. Upon identification of the center of LMAN, a three-electrode miniature motorized microdrive was cemented onto the skull using the procedure described in Fee and Leonardo (2001). The microdrive weighs 1.5 grams, and contains three independently controlled motors, allowing each electrode to be remotely positioned extracellularly with 0.5 um spatial resolution. The electrode tips were implanted ~700 um above LMAN. Electrodes were made from 80 um tungsten wire, insulated with paralyene, and had ~3 MOhm impedance (5-10 um tips; Microprobe, Inc.). Birds were allowed to recover for sufficient time so that they were singing reliably upon presentation of a female bird (~1-2 days). During each day of recording, a custom modified Sutter MP-285 microdrive controller was used to position the electrodes in nucleus LMAN and record their depths. Upon conclusion of the experiment, electrolytic lesions were made in LMAN by passing -3 uA of current through each electrode for 10 seconds (3 times). The bird was sacrificed with an overdose of isoflurane, and the brain were recovered and fixed overnight in 3% paraformaldehyde. The following day, the brain was sliced into 100 um sections on a vibratome, and the location of the lesions was verified to be in LMAN.

There are two general classes of neurons in LMAN: large spiny projection neurons, which synapse onto RA pre-motor neurons, as well as other LMAN neurons, and small aspiny local interneurons which only make synapses within LMAN (Boettinger and Doupe, 1998). In this experiment we had no physical mechanism to distinguish between these two cell types (this could be done using the antidromic stimulation methods of Hahnloser et al., 2002). However, intracellular experiments have demonstrated that LMAN projection

neurons are much more easily isolated than interneurons (Livingston and Mooney, 1997; Rosen and Mooney, 2000). The LMAN neurons we recorded from were very homogeneous in their responses, and had large action potentials consistent with large soma of projection neurons. It is thus highly likely that all of the neurons we recorded from were projection neurons.

Each electrode signal was passed through a single FET (Motorola part # MMBF5457LT1) configured as a unity-gain follower (mounted adjacent to the microdrive on the bird's head). The signals were then amplified 1000x and sampled at 40 kHz (TDT Digital bioamp DB4, National Instruments PCI-6052E data acquisition card). Single neuron recordings were verified with two methods. First, individual spike waveforms of 1.5 msec in length were extracted from each raw electrode signal using a 3x RMS threshold, and then interpolated by a factor of 10 to remove sampling jitter. This produced a matrix of aligned spike waveforms. We calculated the singular value decomposition (SVD) of this matrix. The eigenvectors associated with the largest two eigenvalues of the SVD represent a subspace of the matrix, which contains most of the variability of the spike waveforms. We projected the spike waveforms onto this two-dimensional subspace (a great reduction from the original 60 dimensional space) and looked for well-defined clusters of points. These clusters represent well-isolated neurons. As we are isolating single neurons, this clustering process was not used for spike-sorting, but rather as an reliable denoising mechanism to automatically remove the occasional contamination from a second neuron or electrical artifact. Single-units isolations were further verified by confirming the presence of a spike refractory period in the inter-spike-interval histogram. Because of the difficulty of getting sufficient numbers of normal and altered auditory feedback motifs of song, we did not record simultaneously from multiple single LMAN neurons in this experiment.

Automated Computer Control. A custom designed computer program (written in Labview, National Instruments) was used to control all aspects of the experiment including: simultaneous recordings from the electrodes, microphone, and MP-285 microdrive controller (for electrode depths); generation and triggering of the perturbing auditory feedback; and control of the active sound cancellation system.

Auditory Feedback Perturbation. Microphone data were sampled at 40 kHz, after being low-pass filtered (10 kHz cutoff, 7-pole anti-aliasing filter). The computer continuously

acquired 50 msec segments of sound from the microphone. Each of these bins of sound was passed through a software based infinite-impulse-response (IIR) filter. If the RMS amplitude of the filtered signal exceeded a threshold, artificial auditory feedback was triggered by the computer. Sound played back to the bird was passed through a second IIR filter with a notch in the location of the sound bandwidth used for song triggering. This decoupling of the playback and trigger bandwidths allowed us to continuously trigger on the bird's song, without risk of creating a positive feedback loop between the speaker and the microphone (i.e., triggering on artificial feedback played from the speaker). Previous implementations of our auditory feedback protocol (Leonardo and Konishi, 1999) involved alternating song triggering and sound playback, resulting in a large portions of the song not receiving any artificial feedback and considerable variability in the locations of the song that were exposed to the artificial feedback. This original protocol was not suitable for use with the neural recordings, in which it was essential to reliably cover most of the bird's vocalizations with the same type of auditory feedback in order to facilitate averaging of the neural response to different motifs of song.

Two types of feedback were used, white noise and delayed song syllables. The white noise was continuously regenerated from a fixed distribution and modulated with a 7 Hz envelope (the approximate periodicity of the song syllables forming a motif). We verified that the white-noise feedback causes song decrystallization similar in time course and magnitude to that of the original adaptive feedback protocol used in Leonardo and Konishi (1999). The delayed syllable feedback was a single fixed song syllable played back to the bird repeatedly while singing continued (this is analogous to the syllable-triggered feedback used in Leonardo and Konishi, 1999).

Active Sound Cancellation. Signals recorded on the microphone contained the superposition of the vocalizations produced by the bird and the artificial auditory feedback produced by the computer. Recovering the bird's original vocalizations from the microphone signal was essential for two reasons. First, we needed to verify that the acoustic structure of the songs produced by the bird was not changing in real-time due to the artificial auditory feedback, otherwise one could argue that changes in LMAN activity during singing with altered feedback were due to changes in the motor program itself and not to real-time sensitivity to auditory feedback. Second, we used the precise spectral features of each

syllable to align the song motifs produced at different times by the bird (Leonardo and Fee, 2002). This song alignment is impossible if the syllable features are covered by the artificial feedback.

To circumvent these problems, we implemented an active sound cancellation system. The basic idea behind this system was to use the impulse response of the acoustic environment (speaker, cage, bird, microphone) to predict what the feedback will look like on the microphone and then to subtract this prediction off of the microphone signal, leaving only the bird's vocalizations. However, the bird was constantly moving around in his cage, making the transfer function of the acoustic environment highly nonstationary. It was not possible to simply calibrate the system at the start of the experiment and then use this single fixed transfer function for feedback prediction. Our solution to this problem was to measure the impulse response of the acoustic environment in real-time, effectively taking a snapshot of it each time the bird sang. This instantaneous impulse response could then be used offline to cancel the feedback signal from the microphone signal. In order for this system to work, the transfer function of the acoustic environment must be measured rapidly, immediately after the bird stops singing but before he moves or begins singing again. We measure the transfer function of the acoustic environment using a 12.5 msec Golay code pair (Foster, 1986; Zhou et al., 1992; Braun, 1998). Golay codes are complementary series of binary numbers whose autocorrelation sidelobes are inverses of each other such that their sum is a delta function (Golay, 1961). This property enables the measurement of a transfer function with a substantially shorted probe sequence than that used in many other systems (e.g., sequences of pure impulses, or sequences of sinusoids of varying frequency). The cancellation of artificial auditory feedback from the microphone signal using the active sound cancellation system was approximately 30 dB.

Spectral Analysis. We calculated the time-frequency spectrogram for each song with an 8 msec window sliding in 0.5 msec steps, in which each time point consisted of the direct multitaper estimate of the power spectrum (with a time-bandwidth product NW = 2; Thomson, 1982). Spectral time-derivatives were estimated using the methods of P. P. Mitra (personal communication; see also Tchernichovski et al., 2000).

d' Statistics. Confidence intervals for the d' analysis were estimated using a bootstrap procedure (Bradley, 1993). Briefly, the spike trains for the neuron being analyzed were

randomized with respect to their recording condition (normal singing vs. perturbed auditory feedback singing). The d' statistic was then calculated for the randomized distribution. Repeating this procedure 5000 times generated the expected d' value and 0.01 confidence intervals for the null hypothesis that the neuron was insensitive to auditory feedback.

Kullback-Leibler Information. For each neuron and time window that we estimated the KL information, let N be the number of normal motifs of singing, and let F be the number of perturbed auditory feedback motifs of singing. This allows for a total of N+F possible different spike counts, and defines a set of N+F intervals R, where R ranged uniformly between the largest and smallest values of the spike counts observed in the analysis window across all the recorded motifs. We discretized the spike counts observed in the data into the interval of R in which it lied. This discretization of neural activity represents the best tradeoff between too few responses (causing all spike counts to look the same) and too many responses (causing all spike counts to look different), given the number of data samples we obtained (Panzeri and Treves, 1996; Panzeri et al., 1999). The Kullback-Leibler information was then estimated from the discretized spike trains. Sometimes the normal singing spike count distribution had a zero probability for some spike count Ri, while the feedback singing distribution did not. This could be caused to finite sampling limitations, in which we simply did not collect enough data to observe spike count Ri during normal singing. The effect of this was a division by zero, which forced the KL estimate to infinity. To compensate for this sampling bias, we employed a bias correction (Krichevsky and Trofimov, 1981), in which a small constant offset was added to all the spike count probability estimates so that none were be zero.

A bootstrap procedure was used to estimate confidence intervals for the KL information (Johnson et al., 2000). The null hypothesis for our statistical test was that there was no difference between normal and perturbed auditory feedback singing. By randomizing the experimental condition (normal vs. feedback singing) with respect to the recorded data, calculating the KL information of the randomized data set, and then repeating this process 5000 times, we obtained the expected KL information and associated confidence intervals for the null hypothesis. This procedure was repeated for each step of the 1000 msec, 100 msec, and 15 msec sliding windows. The effect of multiple statistical comparisons must be taken into account when evaluating the significance of the data. Given that we were analyzing the

responses of 17 neurons with normal and perturbed auditory feedback singing, and the length and step size of our sliding windows, we made a total of 1452 statistical comparisons across the entire data set Based on this, we expected 14 false positive events by chance at the 0.01 significance level, and 7 false positive events by chance at the 0.005 significance level. We observed 16 and 4 significant events, respectively, randomly distributed across the neurons and window sizes as expected by chance levels. Based on this, we accepted the null hypothesis that there is no significant difference between LMAN spike trains during normal and perturbed auditory feedback singing.

The KL information is non-negative; this asymmetry produces a positive bias in the KL information even when the two probability distributions are equal, due to random sampling fluctuations. This bias can be seen in the "DC" offset of the estimate shown in Figure 4.8. Note that a similar offset is seen in the randomized KL information estimate, indicating that it is not significant. The shuffling procedure we use is not the best possible estimate of this bias (Panzeri and Treves, 1996). However, our goal was only the computation of the statistical significance of the data.

We required at least 4 motifs of normal singing and 3 motifs of altered feedback singing for analysis using the d' statistic and the KL information metric. Of the 17 neurons in this data set, 7 neurons had at least 7 motifs of song per condition, and none of the 17 showed significant auditory-feedback induced modulation of spike patterns. These sample sizes are not as large as one would prefer, but were unavoidable given the length of time we could hold LMAN neurons (~15-60 minutes) and the average volume of songs produced by birds with the implanted microdrive. The effect of these small sample sizes was to limit the sensitivity of our statistical tests by reducing the possible number of neural responses that could be observed in an analysis window. This does not appear to be a fundamental problem as we acquired sufficient data to establish for nearly every cell an average firing rate pattern that had strong characteristic peaks at various part of the song, indicating that the activity patterns for those cells were well defined. However, it is possible that there are small effects of auditory feedback that occur at variable locations in the song and are thus invisible to the statistical tests we have used. Note that a 1 spike change that occurs reliably at a specific part of the song will be detected as a highly significant event by the KL information test – it is only changes that are both highly variable and small that would be difficult for us to find.

We conducted our analysis over a variety of time windows (1000 msec, 100 msec, 15 msec) in order to detect changes in the temporal pattern of spikes generated by LMAN neurons. Within this large range of time scales, we saw no effect of auditory feedback on LMAN spike patterns. Any changes that did occur would have to be subtle, occurring on time scales less than 15 msec long, and not involving changes in spike rate. This seems fairly unlikely as there are only a very small class of patterns that produce changes in spike pattern but not spike rate in windows of these short sizes.

Appendix 1 : Miniature motorized microdrive construction protocol

The motorized microdrive was designed primarily by Michale Fee. My contributions to the design of the instrument include work on the split connector that allows the motor assembly to be removed from the microdrive barrel, the lateral positioner that provides limited xy positioning of the electrodes, and some of the electrical modifications to the Sutter MP-285 that is used to remotely position the electrodes. I also made numerous contributions to the motordrive construction protocol described below. I built by hand each of the drives used in the LMAN project. The work described in this appendix is excerpted from: Fee, M.S. and Leonardo, A. Miniature motorized microdrive and commutator system for chronic neural recording in small animals. J. Neuroscience Methods. 112, 83-94 (2001).

The use of chronically implanted electrodes for neural recordings in small, freely behaving animals poses several unique technical challenges. Because of the need for an extremely lightweight apparatus, chronic recording technology has been limited to manually operated microdrives, despite the advantage of motorized manipulators for positioning electrodes. Here we describe a motorized, miniature chronically implantable microdrive for independently positioning three electrodes in the brain. The electrodes are controlled remotely, avoiding the need to disturb the animal during electrode positioning. The microdrive is approximately 6 mm in diameter, 17 mm high and weighs only 1.5 grams, including the headstage preamplifier. Use of the motorized microdrive has produced a tenfold increase in our data yield compared to those experiments done using a manually operated drive. In addition, we are able to record from multiple single neurons in the behaving animal with signal quality comparable to that seen in a head-fixed anesthetized animal.

The motorized microdrive described in this appendix was used to record from neurons in nucleus RA and nucleus LMAN in the song control system of the zebra finch. These small birds tolerated the microdrive very well, exhibited all of their normal activities

in the aviary, and were able to fly freely with the microdrive implanted. Singing behavior was unaffected by the microdrive, and the temporal and spectral pattern of syllables retained the hallmark stability seen in normal adult zebra finches. As we discuss in more detail below, the number of useful neural recordings obtained concurrently with singing behavior was dramatically larger in birds implanted with the motorized microdrive rather than with a traditional manually operated microdrive.

Background

Prior to the development of the motorized microdrive, recordings were attempted in four birds with a non-motorized version of this microdrive (similar to that described in (Venkatachalam et al., 1999)). In these four animals, a total of ten cells were recorded during singing behavior, four of these were in one bird and six in another bird. In only one instance were two neurons recorded simultaneously during singing, and only two song motifs were recorded while holding this pair. This is in sharp contrast to the 87 single units, 40 pairs and 5 triplets recorded with the motorized drive in nucleus RA (chapter 2).

Two central difficulties were encountered in the experiments with the manual microdrive, both of which were a result of the need to catch, restrain, and then release the animal in order to adjust the electrodes. Although it was often straightforward to isolate single neurons while manually manipulating the electrodes, it was extremely difficult to release the animal without losing the cell. Furthermore, with the manually operated drive it was not possible to 'tweak' a slightly degraded signal from an electrode, since attempts to catch the bird inevitably resulted in losing the cell. Efforts to record two cells simultaneously on two electrodes were confounded by the same problem. Once a cell was successfully recorded during singing, attempts to catch the bird and isolate another cell on a different electrode inevitably failed.

The second fundamental difficulty was that handling the birds to manipulate the electrodes had the effect of suppressing singing behavior. The zebra finches used in these experiments could normally be reliably induced to sing by placing a caged female zebra finch nearby. The effectiveness of this stimulus was greatly reduced, often for several hours, by the process of capturing and restraining the bird. On many occasions, with a manually

operated microdrive, single-unit signals were obtained and held for tens of minutes, during which the bird could not be induced to sing.

Using the motorized microdrive largely eliminated these difficulties. The greater controllability afforded by the motorized control made the process of getting high-quality signals much simpler that with the manual microdrive. Isolating single neurons was done in the same manner as in acute recording experiments, simply advancing and retracting the electrode to find the optimum signal. Simultaneous recordings could often be isolated by dialing in a single-unit on one electrode and then dialing in a single-unit on another electrode. In addition, electrodes whose signal quality had degraded after some time could usually be improved by adjusting the electrode position. Furthermore, the singing behavior seemed unaffected by the manipulation of the electrodes; the birds showed no response to the operation of the motors.

Neural recordings obtained in singing zebra finches with the motorized microdrive were of the same quality as those seen in head-fixed anesthetized animals. Because of the thread size used to drive the shuttles, and the computer-control of the MP-285 (described below), we were able to move the electrodes with a positional resolution of less than one micron. This allowed us to obtain signal-to-noise ratios that are particularly high compared to those normally seen in chronic neural recordings with microwires. Typical peak-to-peak amplitudes for cells across the entire population were 1-4 mV; excellent signals could exceed 10 mV in amplitude. Good single-unit isolations with the motorized microdrive were often found to be stable for over an hour, and degradation in signal quality could be compensated for by readjustment of the electrode position. A quantitative measure of the effectiveness of the motorized microdrive system is that the yield of cells recorded during the singing behavior was roughly ten times higher than without the motorized microdrive. This dramatic increase in data yield is mirrored by a corresponding increase in data quality; the acquisition of many simultaneous pairs and triplets of cells is virtually unobtainable with a manually operated microdrive in singing zebra finches.

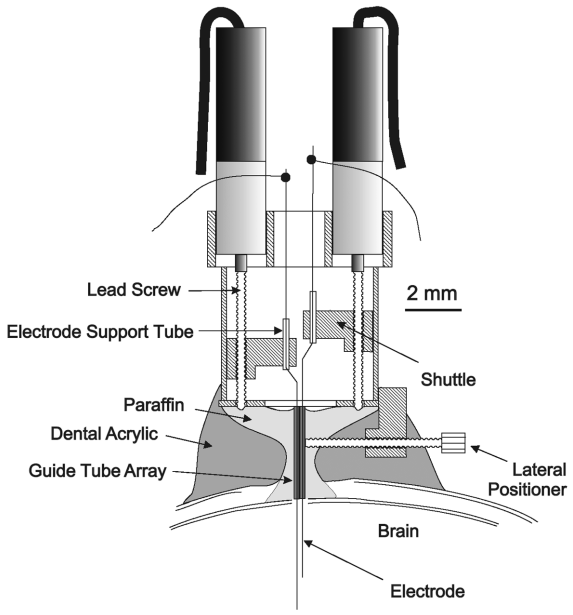

Figure A.1. Overview of motorized microdrive. Each electrode is held by a moveable shuttle that can be advanced and retracted by rotating a threaded lead screw. The shuttle moves in a cylindrical channel within the microdrive body. The lead screw is rotated by a small brushless DC motor that weighs ~100 mg. The device described here has three motors and shuttles arranged in a circle. Limited lateral positioning of the electrodes is accomplished with a threaded rod placed against the electrode bundle.

Methods

The basic design of the microdrive is derived from one described previously (Venkatachalam et al., 1999). Electrodes are held by threaded shuttles that travel along small threaded rods (Figure A.1). The shuttles for multiple electrodes are arranged concentrically to permit a compact arrangement of bundles of electrodes. In contrast to previous designs in which the threaded rods are rotated manually, each threaded rod is mounted to the output shaft of miniature synchronous motor. The motors are 1.9 mm in diameter and weigh approximately 100 mg. The motorized microdrive consists of three main subassemblies. The microdrive/connector assembly is constructed first, followed by the motor assembly, and finally, the electrode assembly. (See Figure A.2 b, c and d). In the following subsections we describe the construction of each of these components in detail.

Microdrive/Connector assembly

The bottom plate is attached to the microdrive body. A double loop of 0.016” diameter solid hook-up wire is wrapped around the base of the microdrive to make the ground connections between the two connectors and to the microdrive body, and also serves to hold the connectors in place prior to gluing. The ground lead of the main connector (Omnetics Connector Co., #A7255-001 and A7732-001) is soldered to the hook-up wire and the

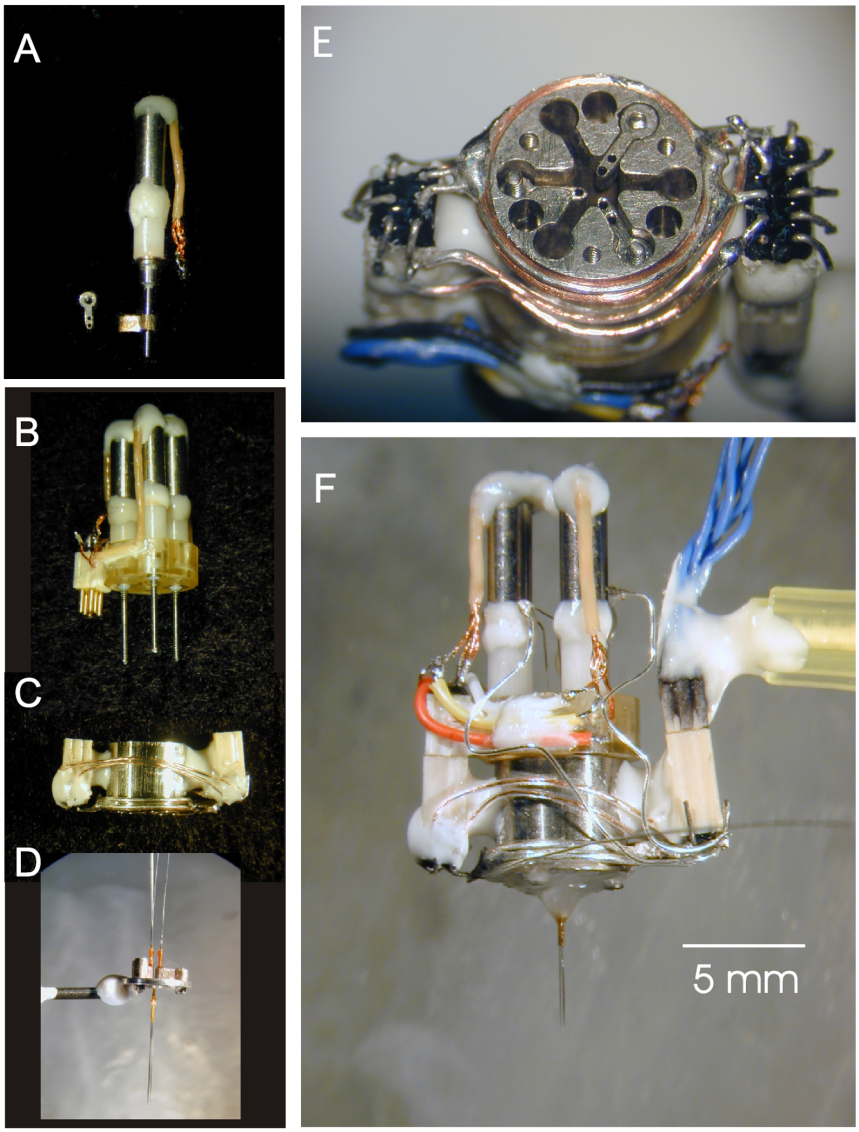

Figure A.2. Photographs of microdrive subassemblies and completed microdrive. A) A single micromotor with attached lead screw and shuttle. Rotations of the motor shaft produce a translation of the shuttle. To the left is a top view of an individual shuttle. B) The motor assembly: Individual motors are pressed into the motor mounting plate. Electrical connections are made from the motors to the connector glued to the side of the motor mounting plate. C) The microdrive body showing the attached miniature electrical connectors (Omnetics, Inc.). On the right is the main connector through which all signals are routed to and from the commutator. The motor control signals are further routed to the motor connector on the left. When the motor assembly is attached to the microdrive body, the motor connectors are mated. D) The electrode assembly can be constructed by temporarily screwing the shuttles to a ring in the proper orientation. The electrodes are inserted and polyamide guide tubes are placed over the electrodes and arranged into an array. E) Bottom view of the microdrive (bottom plate removed) showing the shuttles threaded onto the lead screws. F) Side view of the completed microdrive, loaded with electrodes and ready for implantation. A thin coating of paraffin has been applied to the bottom of the drive to keep the electrodes and the interior of the microdrive free of the acrylic used to attach the microdrive onto the cranium.

connector is positioned as desired. The ground lead of the motor connector is soldered to the hook-up wire on the opposite side and positioned parallel to the microdrive body. Both connectors are then glued in place with Torr Seal (Varian Vacuum Products, Inc.). The motor control signals from the main connector are soldered to the appropriate pins on the motor connector (Cooner Wire, Inc., #CZ-1187, teflon-coated copper stranded wire).

Motor assembly

The motors were purchased from Micro Mo Electronics (www.micromo.com, Part #0206A0.5B+02/1 47:1). The attachment of the motor to the planetary gearhead is extremely fragile and must be reinforced (Figure 3A). A small drop of cyanoacrylate glue is applied to the junction of the motor and planetary gearhead. Once this has hardened, a ring of Torr Seal epoxy (Varian Vacuum Products, Inc.) is applied around the junction and slightly heated with a heat gun (100 C) to facilitate flowing, and to speed the hardening of the epoxy.

The threaded rod is attached to the output shaft of the planetary gearhead. The output shaft is cut with small diagonal cutters to a length of 0.5 mm (± 0.1 mm). The shaft is brittle and can be trimmed with further clipping. A 5.5 mm length of #0000-160 threaded rod is cut from stock with diagonal cutters and the ends cleaned up with a small sharpening stone to remove burrs. The motor is mounted vertically in a holder under a dissecting microscope with the output shaft facing upward. The output shaft (and rotating plate) is covered with a small amount of Torr Seal and the shaft coupler is placed over the output shaft. One end of the threaded rod is heated with the heat gun and the tip is dipped in Torr Seal. The threaded rod is inserted into the shaft coupler. The motor is electrically connected to the motor controller (SC-1900, Minimotor Inc.) and started at a slow speed (~60 RPM). The threaded rod is centered with forceps so that no wobble is visible at the top or bottom as it rotates. Care must be taken since excess epoxy between the coupler and the threaded rod can come in contact with the microdrive body and the resulting friction will impede shuttle movement.

The motor is connected to the controller (SC-1900) to test for proper functioning. Then the connector is displaced to make connection with only a single pair of contacts at a time. Motor rotation is observed for each pair of contacts. There is usually one pair of contacts for which there is the least amount of motor rotation. This pair is chosen for the motor ground and motor common connections (this usually corresponds to the green and

unmarked bundles in the motor cable). The motor cable is bent back down the side of the motor by heating (with a soldering iron) the plastic cable covering at the back end of the motor. With the cable secured along the side of the motor, the cable connection to the motor is reinforced with a small application of Torr Seal.

Although there are only three connections to each motor, there are fifteen extremely fragile copper wires inside the motor cable. The end of the cable covering is removed by melting a ring in the plastic cover with the soldering iron and pulling the end off with forceps. The fifteen wires are bundled in three groups of five. One bundle of wires is labeled with a red enameled wire, one with a green wire and the other with no colored marking. The insulation at the end of the wires is removed by applying a small drop of methylene chloride based paint remover gel (Zip-Strip) to soften the enamel, which is then scraped off gently with sharp forceps. This process may require two applications of Zip-Strip. Each group of five wires is then twisted together and soldered to the connector. The plain bundle is connected to motor ground; the green bundle is connected to motor common; and the red bundle is connected to the individual motor drive signal.

The motors are pressed into the motor mounting plate and the shuttles are screwed all the way onto the lead screws. Construction of the motor subassembly takes place in two stages. In the first stage, the most crucial aspect of the construction process is that the alignment of the motors and lead screws with the shuttle channels in the microdrive body be as precise as possible. Although the output torque of the motors is greatly improved by the 47:1 planetary gear system, the torque (300 uNm) is just sufficient to drive the shuttles. Because the fit between the body of the microdrive and the electrode shuttles is precise, even a small misalignment of the motor shaft can produce sufficient friction to prevent movement. Careful alignment of the motor mounting plate on the microdrive body is required (see Figure A.2 b and c). The motor mounting plate is attached (with #0000-160 screws) to the microdrive body with the shuttles positioned in the shuttle channels. The shuttles are tested one at a time over their full travel range. If there is any binding, the motor mounting plate is loosened, repositioned and reattached. When all three shuttles are free to travel over the full range of motion, then the motor mounting plate and the microdrive body have been properly aligned.

The second stage of the construction of the motor assembly involves gluing the motor connector to the motor mounting plate and making the electrical connections from the motors to the motor connector. The mating part of the motor connector is inserted into the motor connector on the microdrive/connector assembly. This connector is then glued to the motor mounting plate (at the left in Figure A.2 b). It is wise to double-check for proper shuttle freedom of movement before the glue fully hardens since the motor connector strongly constrains the alignment of the motor mounting plate. At this point, the electrical connections from the motors to the motor connector are made, as described in the notes. Once this process is complete, the tops of each motor are linked to each other with a small bridge of epoxy (this more firmly secures the positions of the motors in the motor mounting plate; refer to Figure A.2 b and f).

Construction of the electrode array

The electrode bundle may be assembled directly onto the motor subassembly, or it may be constructed outside of the microdrive body by attaching the shuttles to a temporary mounting ring (identical to the bottom plate; see Figure A2 d). In either case, the bottom plate needs to be removed from the microdrive body to allow the shuttles to be inserted into the bottom of the microdrive and retracted upwards. Short lengths (1.0 mm) of 0.008" ID polyimide tubing (AM Systems, Inc.) are glued into the electrode holes in the shuttles to provide mechanical support for the electrodes. Three electrodes (~3 MOhm tungsten electrodes insulated with parylene; Microprobe, Inc. part #WE300312H3) are cut to the correct length and crimped (see Fig. 1) to provide the proper spacing of the tips. The end of each electrode shank is stripped of 1 mm of insulation. The electrodes are then inserted backwards into their shuttles, and secured in place with a drop of epoxy. The exact orientation of each electrode may now be fine tuned by carefully manipulating the crimp angle with a pair of forceps. A polyimide guide tube (0.004" ID , 2.5 mm) is placed over each electrode and the guide tubes grouped so that the electrode tips form a bundle with 100-200 um spacing. This spacing is constrained primarily by the diameter of the polyimide tubing. The guide tubes are tied with gold wire (0.003" OD) and secured with epoxy.

Once the electrode array is assembled, the motor controller is activated and the shuttles are moved to the top of the microdrive body. The electrode shanks are bent outward

between the motors, and the electrical connections are made from the electrodes to the main connector on the microdrive body using 0.005" teflon coated silver wire (AM Systems, Inc.) and silver epoxy (Epoxy Technology, Inc.). The bottom plate is reattached to the base of the microdrive, and a length of 0.005" bare silver wire is soldered onto the microdrive (to the hook-up wire) for animal ground. A differential ground wire (0.001" teflon coated platinum-iridium wire) is attached to the electrode bundle and aligned so it protrudes ~700 um into the brain near the implanted electrodes. The differential ground is essential to cancel out signal artifacts induced by bird movement. Finally, the electrode guide tube array and bottom plate of the microdrive body are coated with a thin layer of paraffin to protect the moving parts of the drive from the dental acrylic.

The design of the drive permits approximately 3.5 mm of movement in electrode depth through the brain. A second dimension of movement control can be added by the use of a lateral positioner coupled to the electrode bundle (see Figure A.1). A modified shuttle is epoxyed to the side of the microdrive body, perpendicular to the electrode bundle and oriented with the long axis of the target brain nucleus. A #0000-160 threaded rod is advanced through the shuttle until it is in contact with the electrode guide tube array. The threaded rod and the associated shuttle are coated with mineral oil to prevent binding to the acrylic used to cement the microdrive onto the skull. An eighth-turn of the threaded rod will shift the position of the electrode bundle by ~ 20 um. This manually operated positioner can be used periodically during the course of experimentation when the column of tissue associated with the current lateral position of the electrodes has become exhausted of isolatable cells. This typically occurs after a few days of motorized recordings throughout the moveable depth of the electrodes. At this point, the electrodes are retracted to their top position, the bird is restrained, and the lateral positioner is advanced by a small amount (20-50 um) suitable to move the electrodes into fresh tissue.

Motor control electronics

The brushless DC motors used in the microdrive have three windings, and are normally driven with three sinusoidal voltage inputs that have a 120° phase difference. Using this approach, a total of 9 wires are required to control three independent motors. An alternate technique was developed that requires only four motor control wires to be used in

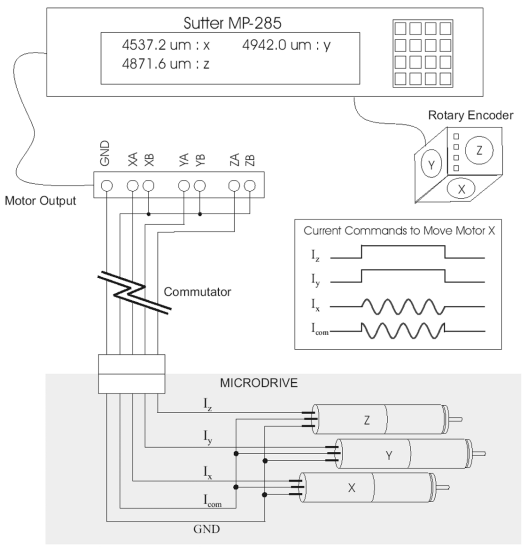

Figure A.3. Schematic of motor connections to the modified Sutter MP-285 controller. Each motor has three connections. These connections are normally driven by three sinusoidal voltage inputs with 120° phase shift. To reduce the number of connections required to drive the three motors, all motors were provided with one common ground and a common sinusoidal input current (I_{com}). The third input was controlled independently for each motor. The Sutter manipulator controller was modified to apply a sinusoidal input current (with a 90° phase shift) only to the motor being driven (e.g., I_x), and a constant bias current to the other inputs (e.g., I_y and I_z) to prevent uncommanded movement of the other motors.

addition to those required for recording neural signals: analog ground, + Vcc for the headstage preamplifier, and the four neural signals (three electrodes and a differential ground reference). In the alternate approach, the motors are driven by two sinusoidal current inputs, one at 0° and one at 90°, as a stepper motor is usually driven. The third connection on each motor is connected to analog ground. The wiring is reduced because all motors share one of these sinusoidal signals (e.g., the 0° signal), referred to as I_{com}, so that when any one motor is 'on', the 'off' motors also have one winding energized (Figure A.3). The second (i.e., 90°) current input is applied only to the motor selected to be 'on'. The 'off' motors are unlikely to turn with only one energized winding, but to prevent any possible spurious rotation, a constant DC current is applied to non-energized winding of the two 'off' motors, locking them in place.

The motor control was implemented using a modified commercial manipulator controller (Sutter Instruments, MP-285). As originally designed, the MP-285 is used to control a stepper-motor-driven, three-axis manipulator. Manipulator movements are computer-controlled in response to commands from a cluster of three rotary-encoded wheels. The embedded computer also keeps track of the current position of the manipulator axes and turns off power to the motors after some delay (t_i) during periods of inactivity. A serial port output allows the depth of the electrodes to be logged by custom-designed computer control

software used to record the neural data. These features make the MP-285 well suited to the control of the three-motor microdrive, with each axis controlling one electrode and motor.

Several modifications of the MP-285 were required. Modifications of the controller firmware were kindly provided by Sutter Instruments (Joe Immel, personal communication). One modification was a re-calibration of position display to reflect the electrode displacement per motor cycle of the motorized microdrive (which is different from the original manipulator). Another modification allows t_i to be user programmed. This delay is set short (<0.5 sec) so that high values of drive current may be used for transient motor movements without thermally overloading the motor.

A simple circuit was added internally to the MP-285 controller to detect command input from the rotary encoder and then perform two functions. First, since analog ground is used also for motor ground, the circuit connects the analog ground line to the controller power supply ground when any motor is activated. When the motors are not in use, the ground is automatically disconnected from the controller power supply to eliminate noise on the electrode signals. Second, the circuit applies the bias 'locking' current to the motors that are not in use. For example, if command input to the x-axis is detected, bias current is applied to the y- and z-axis motors. This design results in the constraint that only one electrode may be moved at a time. Circuit details are available from Sutter Instruments.

Microdrive reproducibility was limited on occasions in which the motors would briefly stall; this problem was minimized by careful attention to the construction process. However, it should be noted that since the microdrive control via the MP-285 is open-loop, any stalling of the motors will produce errors in the estimate of the electrodes depth. The motorized microdrive was found to be quite robust. Across all implants done to date, none of the motors appeared to suffer damage inflicted by the bird. In addition, because the motor subassembly detaches from the microdrive body, the process of reusing the microdrive after an experiment is straightforward. The motor subassembly is detached, the microdrive body is cleaned of acrylic, the motor unit is reattached and the drive is reloaded with electrodes to prepare for the next experiment. The motor or gearboxes occasionally fail for unknown reasons and must be replaced during the reconstruction process.

References

Aamodt, SM., Nordeen, EJ. and Nordeen KW. Blockade of NMDA receptors during song model exposure impairs song development in juvenile zebra finches. *Neurobiology of Learning and Memory*. 65 (1), 91-98 (1996).

Arnold, AP, Nottebohm, F. and Pfaff, DW. Hormone concentrating cells in vocal control and other areas of the brain of the zebra finch. *J. Comp. Neurology.* 301, 245-272 (1976).

Basham, ME., Nordeen, EJ. and Nordeen, KW. Blockade of NMDA receptors in the anterior forebrain impairs sensory acquisition in the zebra finch (Poephila guttata). *Neurobiology of Learning and Memory.* 66 (3), 295-304 (1996).

Benton, S., Nelson, DA., Marler, P. and DeVoogd, TJ. Anterior forebrain pathway is needed for stable song expression in adult white-crowned sparrows. *Behavioral Brain Research*. 96, 135-150 (1998).

Bell, CC. An efference copy which is modified by reafferent input. *Science*. 214, 450-453 (1981).

Bell, CC. Properties of a modifiable efference copy in an electric fish. *J. Neurophysiology*. 47 (6), 1043-1056 (1982).

Bialek W., Rieke F., vanSteveninck RRD. and Warland, D. Reading a neural code. *Science*. 252 (5014), 1854-1857 (1991).

Boettiger, CA. and Doupe, AJ. Intrinsic and thalamic excitatory inputs onto songbird LMAN neurons differ in their pharmacological and temporal properties. *J. Neurophysiology*. 76, 2615-2628 (1998).

Bottjer, SW. and Arnold, AP. The role of feedback from the vocal organ.1. Maintenance of stereotypical vocalizations by adult zebra finches. *J. Neuroscience.* 4 (9), 2387-2396 (1984).

Bottjer, SW., Halserna, KA., Brown, SA. and Miesner, EA. Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. J. Comp. Neurology. 297, 312-326 (1989).

Bottjer, SW., Miesner, EA. and Arnold, AP. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224 (4651), 901-903 (1984).

Bradley, E. *An Introduction to the Bootstrap* (Chapman and Hall, New York, 1993).

Brainard, MS. and Doupe, AJ. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature*. 404 (6779), 762-766 (2000a).

Brainard, MS. and Doupe, AJ. Auditory feedback in learning and maintenance of vocal behavior. *Nature Reviews Neuroscience*. 1 (1), 31-40 (2000b).

Braun, V. Golay sequences for identification of linear systems with weak nonlinear distortion. *IEE Proc.-Sci. Meas. Technol*. 145 (3), 123-128 (1998).

Bridgeman, B. A review of the role of efference copy in sensory and oculomotor control systems. *Annals of Biomedical. Engineering.* 23, 409-422 (1995).

Carr, C. Locating an error correction signal for adult birdsong. *Nature Neuroscience.* 3 (5), 419-421 (2000).

Casacuberta, F., Vidal, E. and Rulot, H. On the metric properties of dynamic time warping. *IEEE Trans. Acoust. Speech.* 35 (11), 1631-1633 (1987).

Chi, ZY. and Margoliash, D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron*. 32 (5), 899-910 (2001).

Cover, TM. and Thomas, JA. *Elements of Information Theory* (John Wiley and Sons, New York, 1991).

Dave, AS. and Margoliash, D. Song replay during sleep and the computational rules for sensorimotor vocal learning. *Science*. 290, 812-816 (2000).

Doupe, AJ. Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J. Neuroscience*. 17 (3), 1147-1167 (1997).

Doupe, AJ. and Konishi, M. Song-selective auditory circuits in the vocal control-system of the zebra finch. *PNAS*. 88 (24), 11339-11343 (1991).

Fee, MS. and Leonardo, A. Miniature motorized microdrive and commutator system for chronic neural recording in small animals. *J. Neuroscience Methods.* 112, 83-94 (2001).

Fee, MS., Shraiman, B., Pesaran, B. and Mitra, PP. The role of nonlinear dynamics of the syrinx in birdsong production. *Nature*. 395, 67-71 (1998).

Feller, W. *An Introduction to Probability Theory and its Applications* (John Wiley and Sons, New York, 1968)

Foster, S. Impulse response measurement using Golay codes. *IEEE international conference on acoustics, speech and signal processing*. 2, 929-932 (1986).

Georgopoulos, AP., Schwartz, AB. and Kettner, RE. Neuronal population coding of movement direction. *Science* 233 (4771), 1416-1419 (1986).

Georgopoulos, AP., Pellizzer, G., Poliakov, AV. and Schieber, MH. Neural coding of finger and wrist movements. *J. Computational Neuroscience.* 6 (3), 279-288 (1999).

Golay, MJE. Complementary series. *IRE Trans. Information Theory.* II-7, 82-87 (1961).

Goller, F. and Suthers, RA. Implications for lateralization of bird song from unilateral gating of bilateral motor patterns. *Nature.* 373 (6509), 63-66 (1995).

Goller, F. and Suthers, RA. Role of syringeal muscles in controlling the phonology of bird song. *J. Neurophysiology.* 76 (1), 287-300 (1996).

Hahnloser, RHR., Kozhevnikov, AA. and Fee, MS. An explicit representation of time underlies the generation of neural sequences in a songbird. *Submitted*. (2002).

Hertz, JA., Krough, AS. and Palmer, RG. *Introduction to the Theory of Neural Computation.* Addison-Wesley, Redwood City (1991).

Hessler, NA. and Doupe, AJ. Social context modulates singing-related neural activity in the songbird forebrain. *Nature Neuroscience.* I 2 (3), 209-211 (1999a).

Hessler, NA. and Doupe, AJ. Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. *J. Neuroscience.* 19 (23), 10461-10481 (1999b).

Ho, CE., Pesaran, B., Fee, MS. and Mitra, PP. Characterization of the structure and variability of zebra finch song elements. *5th Annual Joint Symposium on Neural Computation*. (May 16, 1998).

Houde, JF. and Jordan, MI. Sensorimotor adaptation in speech production. *Science*. 279, 1213-1216 (1998).

Hubel, DH. and Wiesel, TN. Receptive fields of single neurons in the cat's striate cortex. *J. Physiol. London*. 148, 574-591 (1959).

Immelmann, K. Song development in the zebra finch and other estrildid finches. *Bird vocalizations (Hinde, R.A., editor).* Cambridge, England: Cambridge University Press (1969).

Jarvis, ED., Scharff, C., Grossman, MR., Ramos, JA. and Nottebohm F. For whom the bird sings: Context-dependent gene expression. *Neuron*. 21 (4), 775-788 (1998).

Johnson, DH., Gruner, CM., Baggerly, K. and Seshagiri, C. Information-theoretic analysis of neural coding. *J. Computational Neuroscience.* 10 (1), 47-69 (2001).

Kadambe, S. and Boudreauxbartels, GF. Application of the wavelet transform for pitch detection of speech signals. *IEEE Trans. Information Theory.* 38 (2), 917-924 (1992).

Kelley, DB. and Nottebohm, F. Projections of a telencephalic auditory nucleus in the canary. *J. Comp. Neurology.* 183, 455-470 (1979).

Konishi, M. PhD Thesis. University of California, Berkeley. (1963).

Konishi, M. Birdsong: from behavior to neuron. *Annual Review of Neuroscience.* 8, 125-170 (1985).

Konishi, M. The role of auditory feedback in the control of vocalization in the White-crowned sparrow. *Z. Tierpsychol*. 22, 770-783 (1965).

Krichevsky, RE. and Trofimov, VK. The performance of universal encoding. *IEEE Trans. Information Theory.* IT-27, 199-207 (1981).

Leonardo, A. and Konishi, M. Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature*. 399, 466-470 (1999). [Chapter 3]

Leonardo, A. and Konishi, M. An efference copy may be used to maintain the stability of adult birdsong. *In preparation.* (2002). [Chapter 4]

Leonardo, A. and Fee, MS. Different neuronal ensembles underlie similar vocal outputs in a songbird. *In preparation.* (2002). [Chapter 2]

Lettvin, JY., Maturana, HR., McCulloch, WS., and Pitts, WH. What the Frog's Eye Tells the Frog's Brain. *Proc. of the IRE.*, 47(11), 1940-51 (1959).

Lewis, JE. and Kristan, WB. A neuronal network for computing population vectors in the leech. *Nature.* 391 (6662),76-79 (1998).

Livingston, FS. and Mooney, R. Development of intrinsic and synaptic properties in a forebrain nucleus essential to avian song learning. *J. Neuroscience.* 17(23), 8997-9009 (1997).

Marler, P. A comparative approach to vocal learning: song development in white-crowned sparrows. *J. Comp. Physiol. Psychology.* 71, 1-25 (1970).

McCasland, JS. Neuronal control of bird song production. *J. Neuroscience.* 7 (1), 23-39 (1987).

McCasland, JS. and Konishi, M. Interaction between auditory and motor activities in an avian song control nucleus. *PNAS*. 78 (12), 7815-7819 (1981).

Margoliash, D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J Neuroscience*. 3 (5), 1039-1057 (1983).

Margoliash, D., Fortune, ES., Sutter, ML., Yu, AC. and Wren-Hardin, B.D. Distributed representation in the song system of oscines: evolutionary implications and functional consequences. *Brain, Behavior, and Evolution.* 44, 247-264 (1994).

Marler, P. Sensory templates in species-specific behavior. *Simpler networks and behavior, J. Fentress, Ed.* Sinauer, Sunderland, MA. 314-329 (1976).

Marler, P. and Peters, S. A sensitive period for song acquisition in the song sparrow, Melospize-melodia - a case of age limited learning. *Ethology*. 76, 89-100 (1987).

Marler, P. and Sherman, V. Inate differences in the singing behavior of sparrows reared in isolaton from adult conspecific song. *Animal Behavior.* 33, 57-71 (1985).

Myers, C., Rabiner, LR. and Rosenberg, AE. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. *IEEE Trans. Acoust. Speech.* 28 (6), 623-635 (1980).

Nordeen, K. and Nordeen, E. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav. and Neural Biology*. 57, 58-66 (1992).

Nordeen, KW. and Nordeen, EJ. Long-term maintenance of song in adult zebra finches is not affected by lesions of a forebrain region involved in song learning. *Behavioral and neural biology*. 59, 79-82 (1993).

Nottebohm, F. Auditory experience and song development in the chaffinch, Fringilla coelebs. *Ibis*. 110, 549-569 (1968).

Nottebohm, F., Stokes, TM. and Leonard, CM. Central control of song in the canary. *J. Comp. Neurology.* 165, 457-486 (1976).

Okanoya, K. and Yamaguchi, A. Adult bengalese finches require real-time auditory feedback to produce normal song syntax. *J. Neurobiology*. 33, 343-356 (1997).

Panzeri, S. and Treves, A. Analytical estimates of limited sampling biases in different information measures. *Network Computation and Neural Systems.* 7 (1), 87-107 FEB (1996)

Panzeri, S., Treves, A., Schultz, S. and Rolls, ET. On decoding the responses of a population of neurons from short time windows. *Neural Computation.* 11 (7), 1553-1577 (1999).

Price, PH. Developmental determinants of structure in zebra finch song. *J. Comp. Physiol. Psychology.* 93 (2), 260-277 (1979).

Rosen, MJ. and Mooney, R. Intrinsic and extrinsic contributions to auditory selectivity in a song nucleus critical for vocal plasticity. *J. Neuroscience.* 20 (14), 5437-5448 (2000).

Scharff, C. and 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. Neuroscience.* 11 (9), 2896-2913 (1991).

Schmidt, MF. and Konishi, M. Gating of auditory responses in the vocal control system of awake songbirds. *Nature Neuroscience.* 1 (6), 513-518 (1998).

Selim, SZ. and Ismail, MA. K-means-type algorithms: a generalized convergence theorem and characterization of local optimality. *IEEE. Trans. on Pattern Analysis and Mach. Intell.* 1, 81-87 (1984).

Sossinka, R. and Bohner, J. Song types in the zebra finch. *Z. Tierpsychol*. 53, 123-132 (1980).

Solis, MM. and Doupe, AJ. Compromised neural selectivity for song in birds with impaired sensorimotor learning. *Neuron*. 25 (1), 109-121 (2000).

Sperry, R. Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psychology.* 43, 482-489 (1950).

Spiro, JE., Dalva, MB. and Mooney, R. Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *J. Neurophysiology.* 81 (6), 3007-3020 (1999).

Suthers, R. Contributions to birdsong from the left and right sides of the intact syrinx. *Nature*. 347, 473-477 (1990).

Tchernichovski, O., Mitra, PP., Lints, T. and Nottebohm, F. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science*. 291 (5513), 2564-2569 (2001).

Tchernichovski, O., Nottebohm, F., Ho, CE., Pesaran, B. and Mitra, PP. A procedure for an automated measurement of song similarity. *Animal Behavior.* 59: 1167-1176 (2000).

Thomson, DJ. Spectrum estimation and harmonic analysis. *Proc. IEEE.* 70, 1055-1096 (1982).

Thorpe, WH. The learning of song patterns by birds, with especial references to the song of the chaffinch. *Ibis*, 100, 535-570 (1958).

Todt, D. and Hultsch, H. Zum Einfluss des vokalen Lernen auf die Ausbildung gesanglicher Repertoires bei Drosselvoegeln. *Proc. Int. Symp. Verhaltensbiol., ed. G. Tembrock, R. Sigmund, W. Nichelmann.* (1985).

Troyer, TW. and Doupe, AJ. An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. *J Neurophysiology* 84 (3), 1204-1223 (2000).

Vates, GE. and Nottebohm, F. Feedback circuitry within a song-learning pathway. *PNAS.* 92 (11), 5139-5143 (1995).

Vates, GE., Vicario, DS. and Nottebohm, F. Reafferent thalamo-"cortical" loops in the song system of oscine songbirds. *J. Comp. Neurology.* 380 (2), 275-290 (1997).

Venkatachalam, S., Fee, MS., and Kleinfeld, D. Ultra-miniature headstage with 6-channel drive and vacuum-assisted micro-wire implantation for chronic recording from the neocortex. *J. Neurosci. Meth.*, 90, 37-46. (1999).

Vicario, DS. and Nottebohm, F. Organization of the zebra finch song control system: I. Representation of syringeal muscles in the hypoglossal nucleus. *J. Comp. Neurology*. 271, 346-354 (1988).

Vicario, DS. and Nottebohm, F. Organization of the zebra finch song control system: II. Function organization of outputs from nucleus robustus archistriatalis. *J. Comp. Neurology*. 309, 486-494 (1991).

Vicario, DS. and Simpson, HB. Electrical stimulation in forebrain nuclei elicits learned vocal patterns in songbirds. *J. Neurophysiology*. 73(6), 2602-2607 (1995).

Vicario, DS. and Yohay, KH. Song-selective auditory input to a forebrain vocal control nucleus in the zebra finch. *J. Neurobiology.* 24 (4), 488-505 (1993).

von Holst, E. and Mittelstaedt, H. Das Reafferenzprinzip. *Naturwissenschaften*. 37, 464-476 (1950).

Vu, ET., Mazurek, ME. and Kuo, YC. Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neuroscience.* 14 (11), 6924-6934 (1994).

Wild, JM. Descending projections of the songbird nucleus robustus archistriatalis. J. Comp. Neurology. 338, 225-241 (1993).

Williams, H. and Mehta, N. Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. *J. Neurobiology.* 39 (1), 14-28 (1999).

Williams, H. and Nottebohm, F. Auditory responses in avian vocal motor neurons - a motor theory for song perception in birds. *Science*. 229 (4710), 279-282 (1985).

Wilson, MA. and McNaughton, BL. Dynamics of the hippocampal ensemble code for space. *Science*. 261 (5124), 1055-1058 (1993).

Wooley, S. and Rubel, E. Bengalese finches Lonchura-striata-domestica depend on auditory feedback for the maintenance of song. *J. Neuroscience.* 17, 6380-6390 (1997).

Yu, AC. and Margoliash, D. Temporal hierarchical control of singing in birds. *Science*. 273 (5283), 1871-1875 (1996).

Zhou, B., Green, DM. and Middlebrooks, JC. Characterization of external ear impulse responses using Golay codes. *J. of the Acoustic Society of America.* 92 (2), 1169-1171 (1992).