# 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
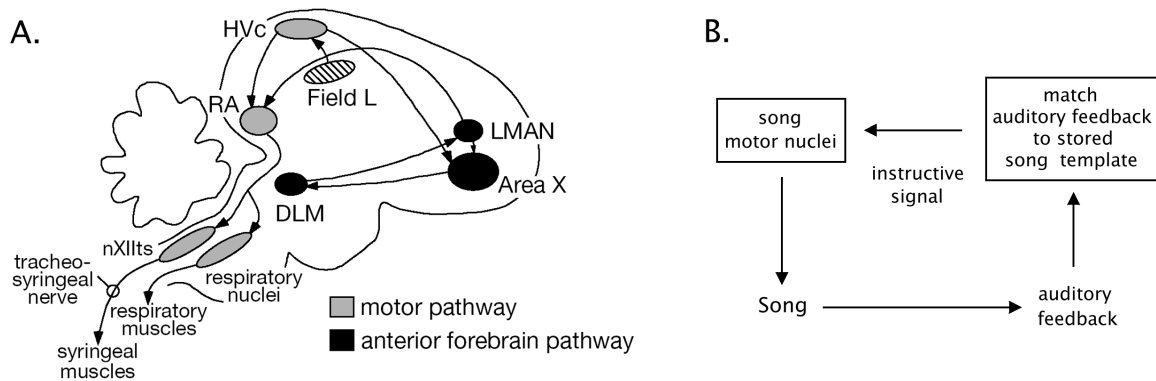
Figure 4.1. Zebra finch song control system and error-correction model. A) Schematic diagram of the zebra finch brain. The motor control pathway, from HVc to the syrinx, is shown in gray. The anterior forebrain pathway is shown in black. B) Error correction model of birdsong. The bird listens to his 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).
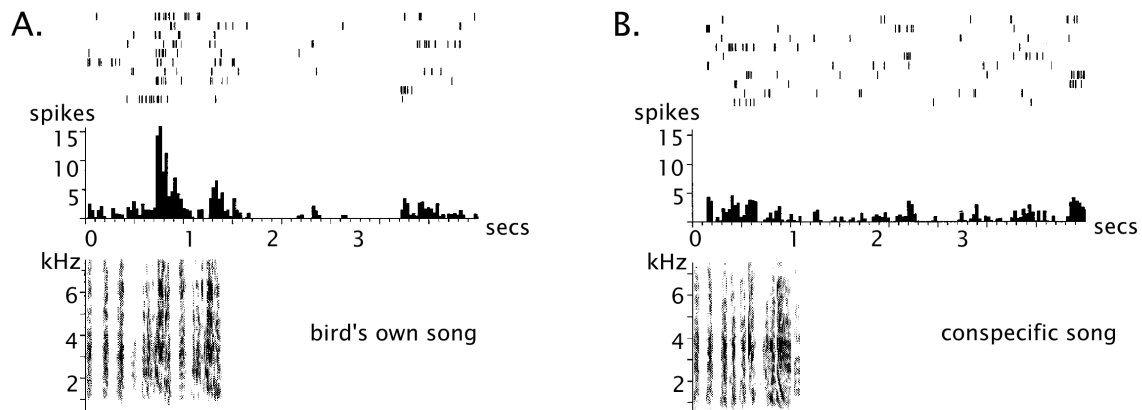
Figure 4.2. Song selectivity of anesthetized LMAN neurons. A) Extracellular response of a single LMAN neuron in a urethane anesthetized bird to playback of the bird's own song (BOS). Shown, from top to bottom, is the spike raster for each trial of sound playback, the average firing rate of the neuron, 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.
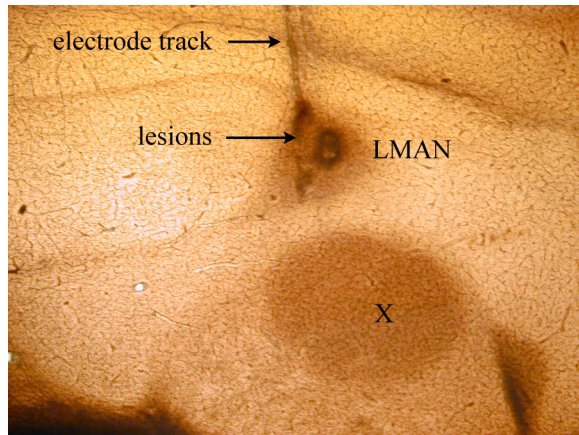
Figure 4.3. Histological reconstruction of motorized microdrive recording sites in LMAN. Two electrolytic lesions, and the electrode track, can be seen within nucleus LMAN. To the left of nucleus X and below are the fiber tracts running from X to the thalamic nucleus DLM, and from DLM to LMAN. The diameter 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
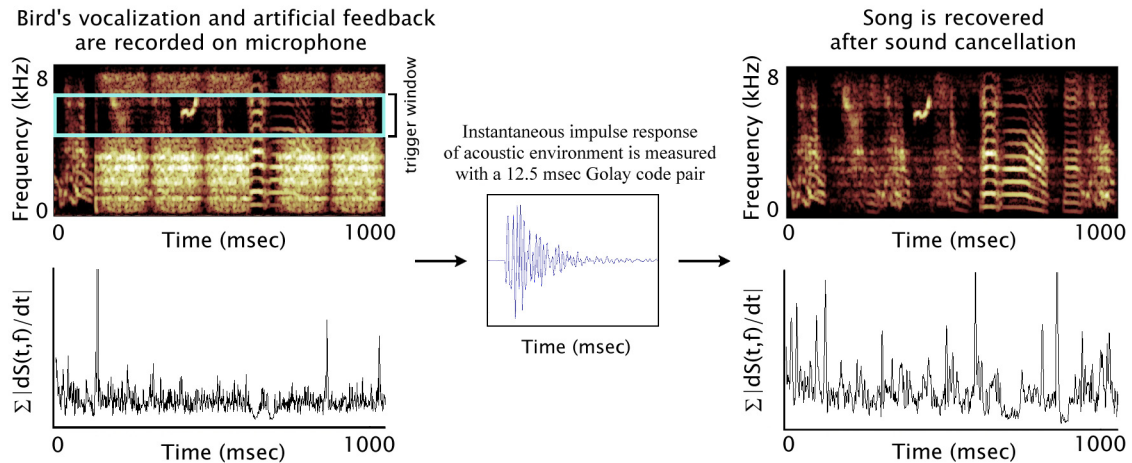
Figure 4.4. Auditory feedback perturbation system and active sound cancellation system. The computer continuously triggers on a narrow frequency band of sound with 50 msec temporal resolution. White noise modulated with a 7 Hz envelope is played to the bird for as long as song is detected. The microphone 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
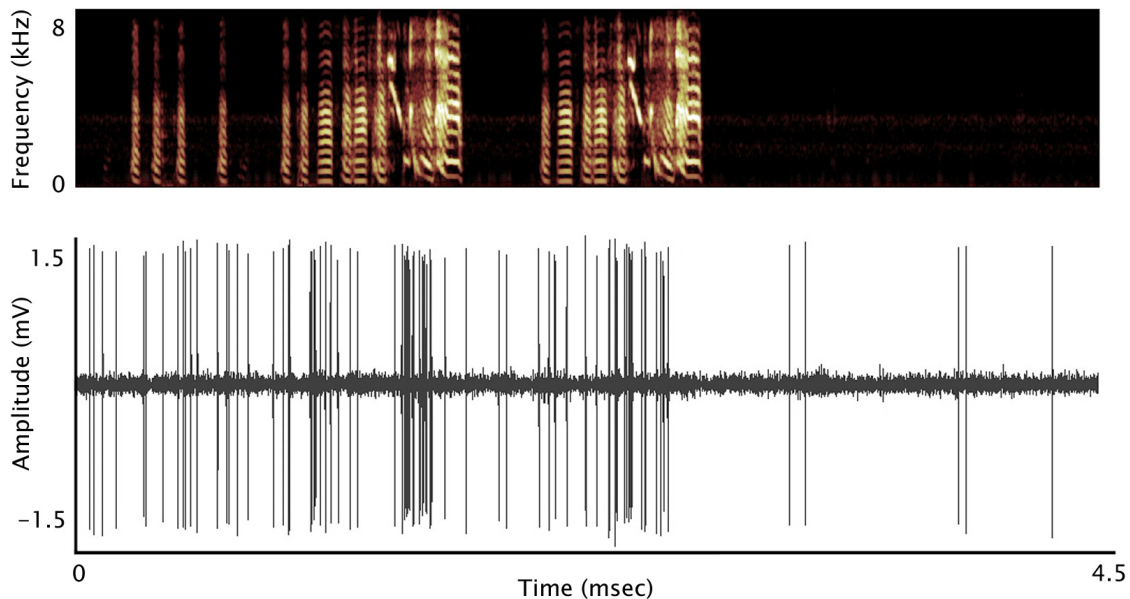
Figure 4.5. Recording from a single LMAN neuron during two motifs of singing.
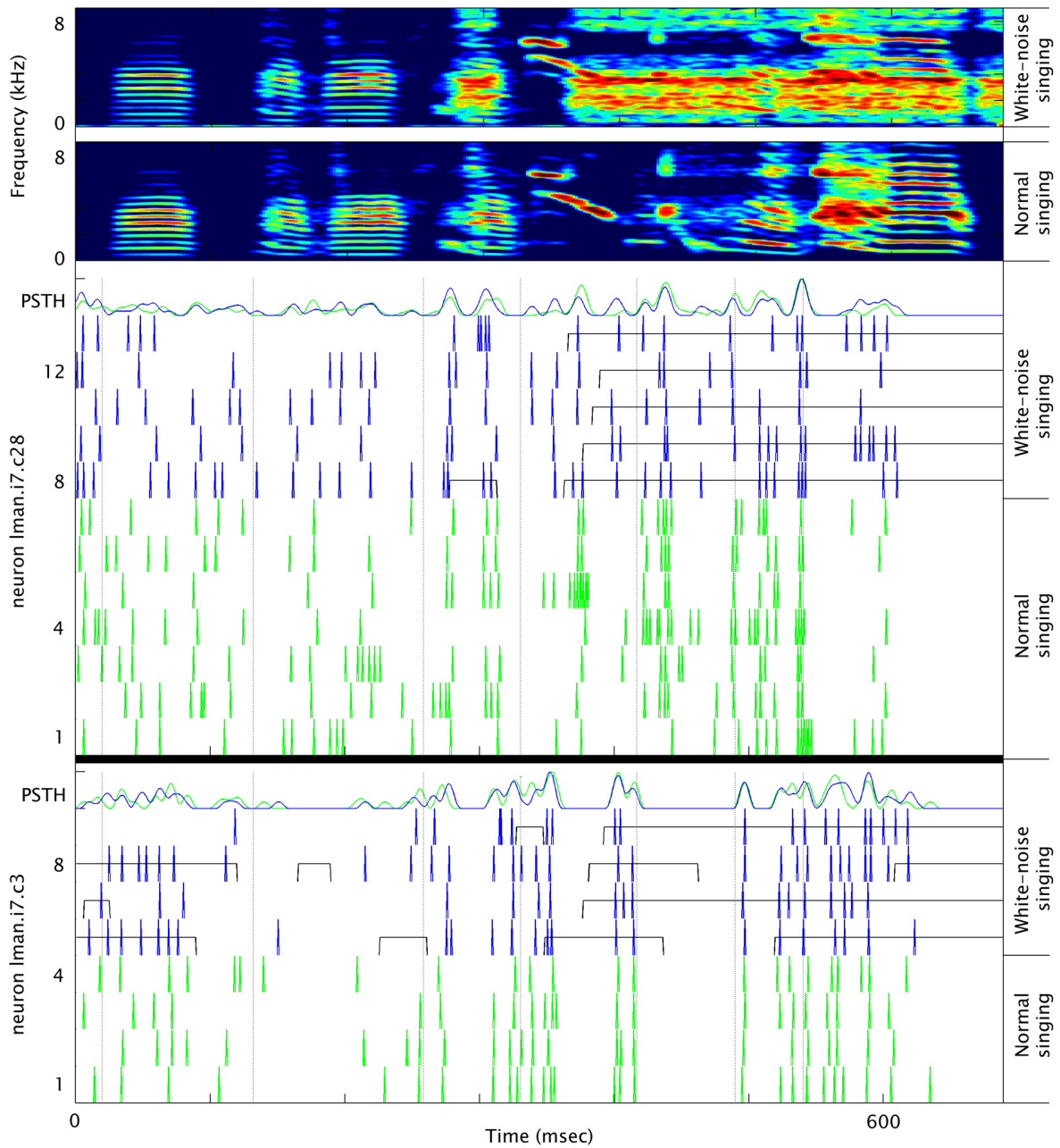
Figure 4.6. Song and spike train alignment for two LMAN neurons. Spikes produced during normal singing are show in green, and spikes produced during white noise feedback singing are shown in blue. The black square-wave lines running through the feedback trials represent the exact locations of the perturbing 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.
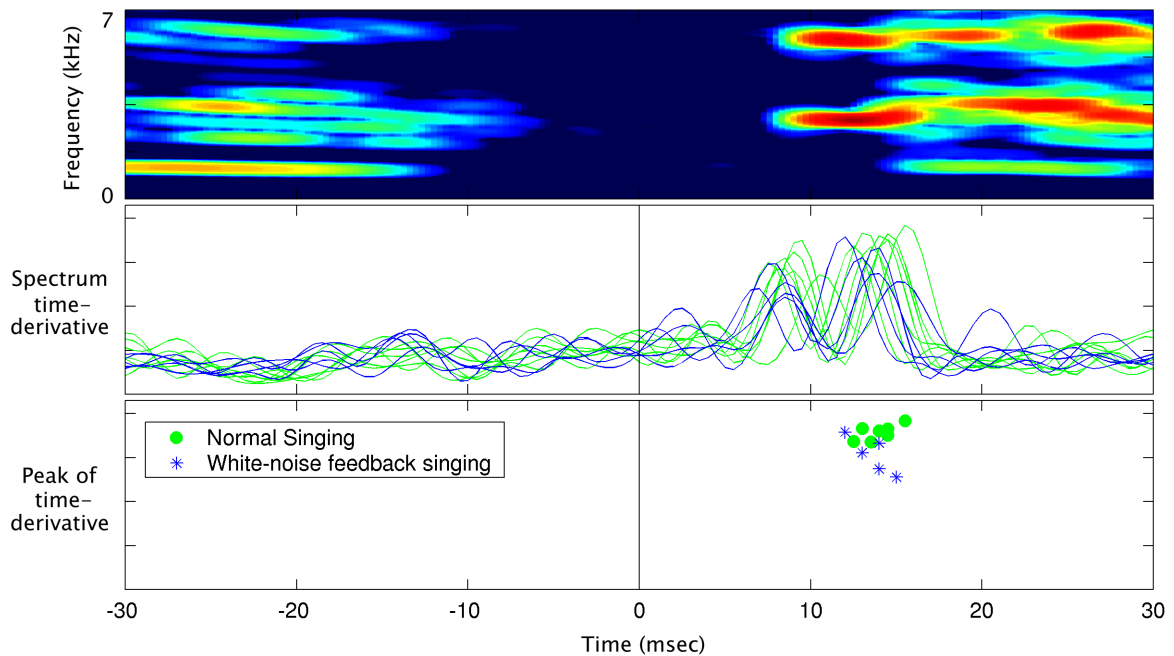
Figure 4.7. LMAN neurons fire spikes precisely timed to the bird's song. Top trace: time-frequency spectrogram of a 80 msec sound from the bird's song. Middle trace: for each motif of song recorded for neuron i7_c28 (Figure 4.6, top panel), we find the time of the first spike occurring in a window from 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
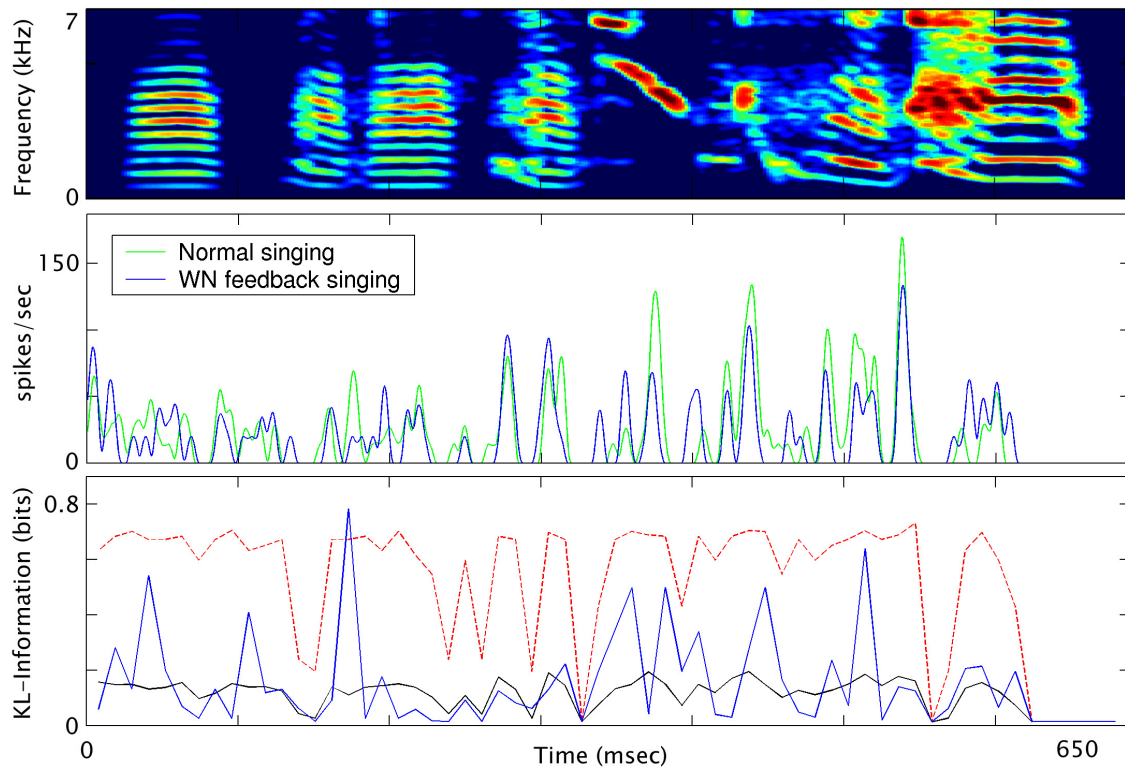
Figure 4.8. Time-varying Kullback-Leibler information for an LMAN neuron. Top: song spectrogram. Middle: average firing rate during normal (green) and perturbed auditory feedback (blue) singing for neuron i7.c28 (see Figure 4.6). Bottom: Kullback-Leibler information (blue), computed in a sliding 15 msec 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
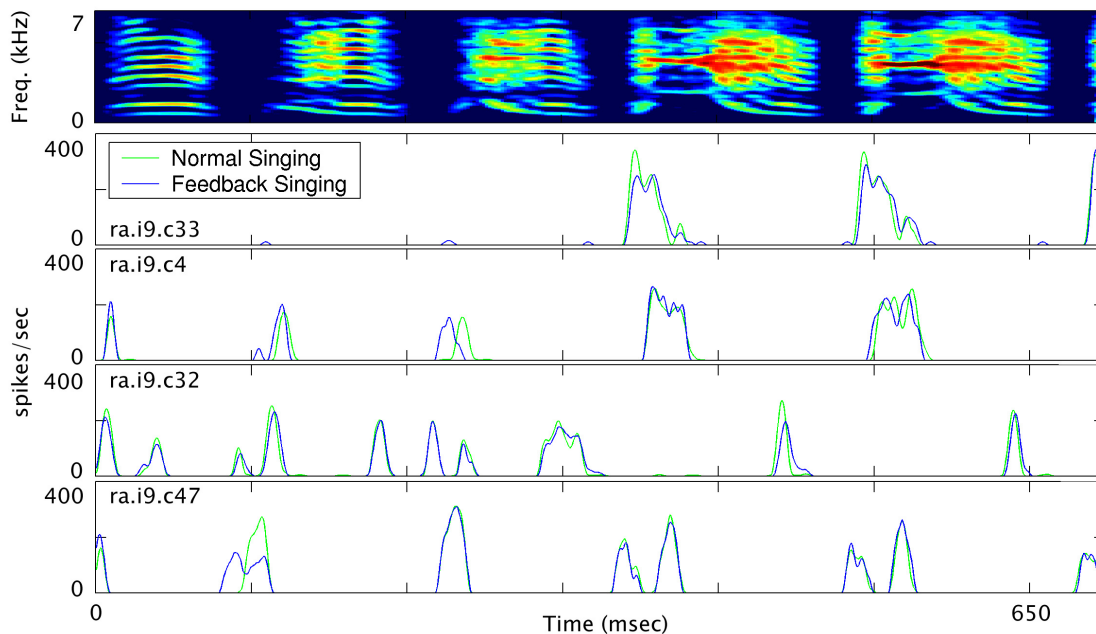

Figure 4.9. Average firing rate of four RA neurons during normal and perturbed auditory feedback singing. No significant feedback-induced changes in spike rate or pattern were found for any of the neurons.

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
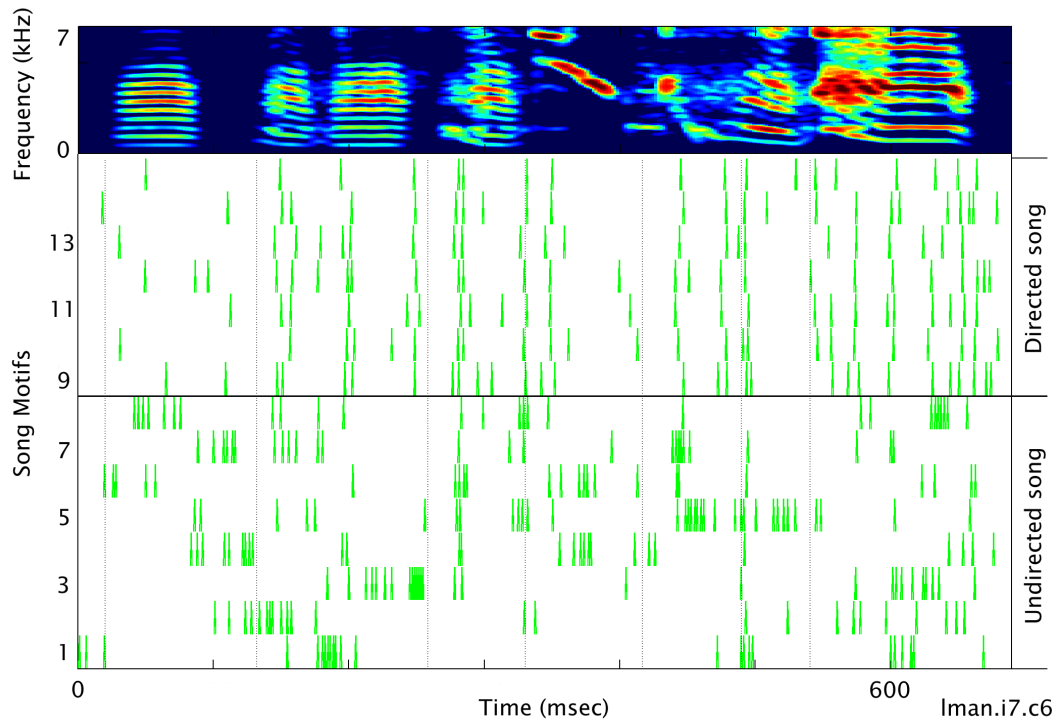
Figure 4.10. LMAN neural response during directed and undirected song. The first 8 spike trains show the spike activity of a single LMAN neuron during undirected song. Approximately one minute after the cessation of the 8th motif, a female bird was presented to the male and he immediately sang the following 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 R_i, 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 R_i 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.