Acquisition of an Acoustic Template Leads to Refinement of Song Motor Gestures

Jorge M. Méndez,1 Analía G. Dall’Asén,1 Brenton G. Cooper,2 and Franz Goller1
1Department of Biology, University of Utah, Salt Lake City, Utah; and 2Department of Psychology, Texas Christian University, Fort Worth, Texas

Submitted 23 November 2009; accepted in final form 10 June 2010

Méndez JM, Dall’Asén AG, Cooper BG, Goller F. Acquisition of an acoustic template leads to refinement of song motor gestures. J Neurophysiol 104: 984–993, 2010. First published June 16, 2010; doi:10.1152/jn.01031.2009. Vocal learning, a key behavior in human speech development, occurs only in a small number of animal taxa. Ontogeny of vocal behavior in humans and songbirds involves acquisition of an acoustic model, which guides the development of self-generated vocalizations (sensorimotor period). How vocal development proceeds in the absence of an acoustic model is largely unknown and cannot be studied directly in humans. Here we explored the effects of an acoustic model on song motor control by comparing peripheral motor gestures (respiration and syringeal muscles) of tutored birds with those of birds raised in acoustic isolation. Although the overall use of syringeal muscles during song was similar in both groups, tutored birds displayed enhanced temporal patterns of activation in respiratory and syringeal motor gestures. Muscle activation was more uniformly distributed throughout the song of tutored birds than that of untutored birds. Similarly, the respiratory effort was similar for both groups, but the expiratory pulses of song contained more modulations and temporal complexity in tutored birds. These results indicate that the acquisition of an acoustic template guides a refinement of experience-independent motor gestures by increasing temporal fine structure, but there is no difference in bilateral activation patterns for a given sound between the two groups. Nevertheless, these subtle temporal changes in muscle activation give rise to pronounced acoustic differences between the songs of the tutored and untutored birds. Experience with song during ontogeny therefore guides a more refined use of experience-independent motor programs.

INTRODUCTION

Vocal learning is an imprinting-like process in which acoustic models of species-typical vocalizations are acquired during a sensory period (Adret 2008; Brainard and Doupe 2002). These stored representations are thought to serve as a template for vocal development (Konishi 2004; Nick and Konishi 2005). Regardless of the presence of an acoustic model, vocal development proceeds through a sensorimotor practice phase in which vocal production is gradually perfected through acoustic feedback (Brainard and Doupe 2002). The timing of song ontogeny is species specific. For example, zebra finches acquire song during a sensory phase that typically lasts from posthatching day (phd) 25–60, whereas the sensorimotor phase starts at phd 30 and lasts until phd 90, at which time crystallization occurs. It is thought that, in the absence of a learned template, vocal ontogeny is shaped by an experience-independent representation of genetically predetermined vocalizations (Forstmeier et al. 2009).

Humans and songbirds share these basic features of vocal learning (Doupe and Kuhl 2008). However, in humans, it is not possible to study the role of an acquired acoustic template in modifying innate motor instructions for speech. Deaf children are unable to perceive the acoustic models during the sensory phase but are also not able to hear their own vocalizations during the sensorimotor practice phase, likely resulting in much greater speech deficits. Although a few historic “Kaspar Hauser” cases (children deprived of social contact and therefore potentially of acoustic models for speech) have been reported, it is unclear to what degree they have been exposed to speech at an early age (Fromkin 1974; Lane 1976). In songbirds, it is possible to test the influence of an acquired acoustic model on the motor instructions for learned vocal behavior by rearing birds in isolation, preventing exposure to song early in development, and comparing song motor patterns between group-reared and isolate-reared birds.

Song production involves the coordinated activity of motor systems for respiration, vocal organ (syrinx), and articulatory systems of the upper vocal tract (Goller and Cooper 2004). The respiratory system generates sequences of expiratory and inspiratory pulses, which determine the coarse temporal pattern of the song (Goller and Cooper 2004; Suthers and Zollinger 2004; Wild et al. 1998). The activity of the syringeal muscles is thought to provide fine control of airflow and thus temporal features as well as the acoustic structure of vocalizations (Goller and Suthers 1996a,b). In brown thrashers (Toxostoma rufum) for example, the activity of the ventral syringeal muscle (vS) is correlated with the fundamental frequency of the vocalization, and the dorsal and ventral tracheobronchial muscles regulate the syringeal valve. Although the respiratory dynamics of zebra finches during song were studied (Goller and Daley 2001; Franz and Goller 2003), the syringeal dynamics for this species are insufficiently understood, and the specific role of individual muscles is likely to deviate from the respective roles described in thrashers (Goller and Cooper 2004; Vicario 1991).

In the zebra finch, song is a stereotyped sequence of four to eight different syllables (motif), which is repeated a variable number of times in a song bout. Males exposed to a tutor during the critical period acquire song elements from its song and refine their own vocalizations by practice (Price 1979). Males raised without exposure to a tutor model (untutored) develop song that is somewhat similar to that of tutored birds but is characterized by atypical frequency structure of syllables (syllables are defined as uninterrupted traces on a spectrogram) and the presence of notes (acoustic units within a syllable) with atypically long duration (Fehér et al. 2009).
The coordinated activity of respiratory, syringeal, and upper vocal tract motor systems emerges from a well-characterized neural song production circuit (Brenowitz et al. 1997). The detailed nature of the motor program for song in the zebra finch is still largely unknown, despite recent advances in chronic recordings from neurons in premotor areas of this circuit (Hahnloser et al. 2002; Leonardo and Fee 2005). Nevertheless, recent behavioral and neural studies have unraveled exciting details about song learning in the zebra finch. Exposure of the juvenile bird to the song of a tutor causes rapid emergence of some features of the song model in the juvenile bird, such as syllable duration, frequency modulation (FM), and syllable sequence (Tchernichovski et al. 2004). Hearing tutor song also causes a rapid change in the activity of neurons in a premotor nucleus [robust nucleus of the arcopallium (RA)] of the song circuit (Shank and Margoliash 2009). In tutored birds, individual RA neurons fire short bursts during specific segments of song (Yu and Margoliash 1996), such that the combined output of many neurons may constitute the full sequence of premotor information (Leonardo and Fee 2005). It is unknown whether an acquired acoustic model of song leads to changes in the firing pattern of individual neurons or a change in the population output in RA.

Here we explored how the presence of an acoustic template during development affected the motor instructions for song by studying respiratory and syringeal motor gestures (i.e., the activity of the motor systems related to the execution of an action). By comparing the motor gestures after song development is completed in tutored and untutored birds, we are able to infer how the acoustic experience affects motor control of song. The main specific questions about the role of acoustic experience are as follows: 1) does acoustic experience qualitatively change the motor control of song production and 2) how do differences in motor gestures control the acoustic differences between experience-dependent and experience-independent song?

**METHODS**

**Animals and isolation protocol**

We used live tutored (tutored; n = 10) zebra finches and birds with no exposure to tutor song during development (untutored; n = 14) in this study. The tutored birds were bled and raised in a flight aviary, whereas the untutored birds were bled with parents and siblings in acoustic isolation chambers until posthatching day 7 (phd 7). At this point, the father was removed from the breeding cage, and at 35 phd, the bird was put in an acoustic isolation chamber and kept in a room with no adult male zebra finches until the bird’s song was recorded (after 120 phd).

**Surgical procedure and recording**

All experimental procedures were approved by the IACUC of the University of Utah. For the analysis of air sac pressure, data were collected from 10 tutored birds and 14 untutored birds. Data from five tutored birds and seven untutored birds were used for the analysis of the EMG activity of the syringeal muscles. For all the analyses, we used data from directed song.

The surgical procedures and recording methods were previously described in detail (Goller and Suthers 1996a,b). In summary, an elastic belt with a Velcro connector was fitted around the bird’s thorax to which a wire was attached for tethering the bird to a counterbalanced tether arm. After the birds resumed singing, the surgical procedure was carried out. The bird was deprived of food and water for 1 h before the surgery. Chloropent anesthesia was used for implanting the EMG electrodes, and isoflurane anesthesia was used for implanting a cannula into an air sac.

**Air sac pressure recording**

A small hole was made in the abdominal wall, and a flexible cannula (Silastic tubing, Dow Corning, Midland, MI) was inserted into an anterior thoracic air sac. The cannula was sutured to the ribs, and the free end of the cannula was connected to a pressure transducer (FPM-02PG, Fujikura) mounted on the backpack.

**Syringeal EMG recordings**

The syrinx was accessed through the interclavicular air sac, and bipolar electrodes were inserted into the syringeal muscles and secured into place with tissue adhesive. Two pairs of electrodes were implanted in left and right VS or dorsal tracheobronchial (dTB) muscles. The electrodes were routed subcutaneously to the backpack. Both syringeal muscle pairs were recorded in some individuals by performing successive surgeries on each bird with 1–3 mo between experiments.

**Data analysis**

All the analysis methods for EMG and air sac pressure data described in the following were programmed in Matlab (version 7.7.0, Mathworks, Natick, MA).

**ACTIVE TIME RATIO FOR THE AIR SAC PRESSURE.** The pressure pattern was adjusted and normalized to make the pressure amplitudes of the motif fit into the interval [0, 1]. The probability distribution of the amplitude of the pressure patterns were calculated for both tutored and untutored groups. Based on the results of a previous study measuring air sac pressures in zebra finches (Goller and Daley 2001), a threshold value separating expiratory from inspiratory values was established (threshold for normalized pressure = 0.3). The areas beneath the probability distributions between the threshold and 1 were computed. The sets of values for tutored and untutored groups were statistically compared.

**CV OF THE PRESSURE PATTERN.** The CV (CVx) was calculated for each bird as the ratio between SD and the mean of the pressure pattern during the motif.

**PROBABILITY DISTRIBUTION OF THE PRESSURE PATTERN DERIVATIVE (Pp).** The value of the derivative of the pressure pattern was calculated at each time point evaluating the derivative of the polynomial (degree 5) that provides the best fit for the pressure pattern in an interval of 10 ms around that time point. The probability distribution of these values was built (Pp). From the probability distributions, the following cumulative distributions C(x) were computed

\[ C(x) = \int_{-\infty}^{x} P_p(y)dy \]

For each group, the values x so that C(x) = 0.9 were calculated, and a Kolmogorov-Smirnov test was run to assess the significance of the differences between groups.

**Quantification of the modulation of pressure patterns**

To quantify the modulation characteristics of air sac pressure for individual pulses of the song motif, we fitted a fourth-order polynomial (using the least square method) to the expiratory pulse. This procedure reproduces the main shape of the pressure pulse but without the superimposed fine modulations of pressure. We calculated the summed error between the actual pressure signal during the expiratory pulse and the fitted curve per unit time. This pressure modulation...
index (PMD) provides a quantification of the modulation complexity that is independent of pulse duration.

PERMUTATION ENTROPY \((h(n))\). The method proposed by Bandt and Pompe (2002) to compute the permutation entropy of a time series was used. The permutation entropy is a widely used quantifier for the degree of complexity of a time series. Briefly, the method consists of the following algorithm. The time series \(x_i\), \(i = 1, \ldots, N\) is first embedded in a phase space that gives the order \((n)\) of the entropy with a time-delay \(T\) (given in number of points). Hence, the set of vectors \((x_i, x_{i+T}, \ldots, x_{i+(n-1)T})\) is the result of the embedding. A permutation \((p)\) of the elements of each vector puts the elements of the rearranged vector in ascending order. The relative frequency \([P(p)]\) of each permutation in the set of vectors is calculated. The permutation entropy \(h(n)\) for \(n > 2\) is defined as

\[
h(n) = -\frac{\sum P(p) \log[P(p)]}{\log(n!)}
\]

In the sum, all the possible permutations of \(n\) elements are added.

For the various analyses of EMG recordings, the raw EMG signals were rectified, integrated with a 2.5-ms sliding time window, and normalized. We named this transformed signal “the activation pattern” \((\text{VP}_{\text{EMG}})\).

MUSCLE USE \((MU)\). To assess the absolute activity of a muscle per unit time during the song, the following quantifier was calculated

\[
MU = \frac{T_{\text{quiet}}}{T_{\text{motif}}} \int_{\text{resp cycle}} |\text{VP}_{\text{EMG}}(s)| ds
\]

The upper line means average calculated over several respiratory cycles, \(T_{\text{quiet}}\) is the period of the quiet respiratory cycle, and \(T_{\text{motif}}\) is the duration of the motif. To calculate the average, usually five quiet respiratory cycles were used.

ACTIVE TIME RATIO OF THE EMG (\(ATR_{\text{EMG}}\)). We computed the ratio of time during while the muscle is active during the motif. The active state was defined whenever the muscle was activated at an amplitude 1.5 times greater than the maximal value during normal breathing.

INTEGRAL VARIATION OF THE ACTIVATION PATTERN FROM ITS MEAN \((\text{IV})\). The sum of the absolute values of the difference between \(\text{VP}_{\text{EMG}}\) and its mean during the motif was calculated.

CUMULATIVE VARIATION OF THE ACTIVATION PATTERN \((VC)\). If the motif goes from \(t = 0\) to \(t = T\), \(VC\) is calculated as

\[
E(t) = \int_{0}^{t} |\text{VP}_{\text{EMG}}(s)| ds, \quad D(t) = \frac{t}{T} \int_{0}^{T} |\text{VP}_{\text{EMG}}(s)| ds
\]

\[
VC(t) = E(t) - D(t)
\]

The ratio of time during the motif while \(VC\) is negative is used to quantify how long the activation amplitude is lower than what it would be if the activation was uniformly distributed across the motif. A statistical comparison between these sets of values for tutored and untutored birds was calculated.

CORRELATIONS. The autocorrelation function was computed for each EMG time series of birds in both groups. The lag was converted to time values for plotting. Also, the Pearson’s correlation coefficient was computed for each bird for the time series corresponding to the EMG activity of right and left muscles.

QUANTIFICATION OF THE DEGREE OF SYNCHRONY. To quantify the degree of synchrony of left and right muscles at a precise moment (corresponding to the element \(i\) of the time series), the instantaneous cross-correlation coefficient \([C_D(i)]\) at that time was calculated as

\[
C_D(i) = \frac{(x(i) - \bar{x})(y(i) - \bar{y})}{\sigma_x \sigma_y}
\]

HIGH-FREQUENCY MODULATIONS IN THE EMG SIGNAL. Power spectral density of the EMG activation patterns was calculated based on Thomson’s multitaper method (Thomson 1982). For frequencies between 10 and 300 Hz, we tested whether power spectral densities were different between tutored and untutored birds. To do so, we divided the frequency range into 20-Hz bins, calculated the signal power in that range, and performed a Kolmogorov-Smirnov test on the pooled data from all individuals in each group for each bin.

LATERALIZATION
The difference in activation of the left and right muscles per unit time is used for quantifying the degree of lateralization

\[
L = \frac{1}{T} \int_{0}^{T} \left| \text{VP}_{\text{EMG}}(i) - |\text{VP}_{\text{EMG}}(i)| \right| dt
\]

where the interval \([0,T]\) refers to the whole duration of the motif \((T_{\text{motif}})\) or the duration of each syllable \((L_{\text{syl}})\).

SOUND ANALYSIS. The analysis of the songs of every individual in both experimental groups was performed using SOUND ANALYSIS PRO (Tchernichovski et al. 2000; http://offer.sci.ccny.cuny.edu/2001/html/sound_analysis.html).

RESULTS
Examples for simultaneously recorded respiratory pressure and EMG activity of syringeal muscles during a song motif are shown in Fig. 1 for an untutored and a tutored bird. Because songs varied substantially between individuals in each group, quantitative analyses are used to explore systematic differences between the two groups.

Activation level of peripheral motor systems
The first question to address is whether exposure to a tutor song changes the overall activation level of peripheral motor systems for song. We compared the ratio of expiratory and inspiratory durations for the entire motif [active time ratio for the air sac pressure (ATRP)] and the CV \((C_V)\) of the amplitude of air sac pressure patterns between both tutored and untutored birds and found no significant differences [ATRP: Kolmogorov-Smirnov test, \(D = 0.134\) and \(C_V\) (untutored) = 0.47 ± 0.01, \(C_V\) (tutored) = 0.46 ± 0.02, Kolmogorov-Smirnov test, \(D = 0.24\)].

Next, we asked whether overall activation of syringeal muscles differs between both groups. The mean muscle use \((MU)\) for the motif did not differ between tutored and untutored groups for the left and right vS or for dTB muscles (Kolmogorov-Smirnov test, \(D = 0.27\) for vS and \(D = 0.9\) for dTB). Additionally, there was no significant difference in the time ratio of muscle activity during the motif \((ATR_{\text{EMG}})\) for vS and for dTB (Kolmogorov-Smirnov test, \(D = 0.87\) for vS and \(D = 0.67\) for dTB). These results indicate that tutoring does not change the overall use of respiratory or syringeal muscles for song production.
Temporal structure of respiratory and syringeal patterns

One important question to address is whether the acquired acoustic template leads to a change in the degree of modulation and complexity of the respective contributions to song by different motor systems. In a comparison of different syllables for tutored and untutored groups, songs in the tutored group contained a higher degree of respiratory modulation (Fig. 2). The increased modulation of respiratory pressure during song syllables in tutored birds produced more variable acoustic behavior than was found in the untutored group. Furthermore, to understand the complexity at the behavioral level, it is necessary to record the respiratory and syringeal motor patterns simultaneously. For example, in the syllables presented in Fig. 2 for the untutored birds, the pressure patterns do not reflect the frequency and AM displayed in the acoustic output, suggesting that these features were generated by syringeal control.

Based on these qualitative observations of increased modulation in expiratory pressure for the tutored birds, we compared the temporal structure of the motor gestures quantitatively between tutored and untutored birds. The dynamical complexity of the time series of pressure patterns was assessed by computing the permutation entropy (Bandt and Pompe 2002). The comparison between both groups yielded that tutored birds displayed more complex pressure patterns (Kolmogorov-Smirnov test, $D < 0.05$ for time delays between 3 and 9 ms and order 5; see METHODS; Fig. 3A). Furthermore, the shape of the probability distribution of the derivative of the air sac pressure patterns for the tutored birds is broader (see METHODS; Kolmogorov-Smirnov test, $D = 0.01$ for a cumulative value of 0.9), showing more temporal modulation of air sac pressure (Fig. 3B). This indicates that respiratory pressure patterns of tutored birds present more fluctuations along a broad time scale, consistent with the existence of more variable values of the pressure derivative. All these results point to a more frequent occurrence of high-frequency fluctuations in the respiratory motor gestures of tutored birds.

FIG. 1. Representative recordings of song and physiological parameters for tutored (right) and untutored (left) zebra finches. A: spectrogram of the song. B: air sac pressure patterns. C: activation pattern of the left ventral syringeal muscle (vS) muscles during the motif. D: activation pattern of the right vS muscles during the motif. Air sac pressure and EMG traces are shown as relative voltages. EMG activity was rectified, integrated, and normalized so that the maximal value during the motif is 1. The time series of the EMG patterns were not shifted to account for the delay between electrical activation and movement of syringeal effectors. Generally, the syllables are produced during expiratory air sac pressure pulses, although sometimes zebra finches produce sound during deep inspirations as the one shown for the tutored bird (2nd to last syllable accompanied by deep inspiratory pressure).
The relationship between the mean pressure and sound frequency differs between tutored and untutored birds. Whereas high-frequency sounds in tutored birds are generated with higher mean air sac pressure than low-frequency sounds, this difference is less distinct in untutored birds (Fig. 4A). Air sac pressure values for high-frequency sounds in untutored birds overlap widely with those for low-frequency sounds.

The relationship between modulation index of the expiratory air sac pressure and the activation of syringeal muscles also indicates a difference in motor control between tutored and untutored birds (Fig. 4B). The values for average EMG activity during syllables span a wider range, particularly for high values, in untutored birds. Pressure modulation indices extend to higher values in tutored birds. This difference suggests that tutored birds use a combination of respiratory and syringeal control to achieve sound modulation characteristics, whereas untutored birds predominantly use syringeal control.

We inspected the high-frequency sound elements that were generated with relatively low air sac pressure in untutored birds more closely to test whether these constituted examples of sounds that are generated with different motor gestures by tutored and untutored birds. This was not the case, because the sound elements in this pressure range are atypical sound elements that are rarely produced by tutored birds. The motor gestures for producing similar sounds were therefore similar in both groups, and the differences in the respective ranges of air sac pressure and EMG activity of syringeal muscles correspond to acoustic differences. This is further shown by similarities in motor gestures for other shared acoustic characteristics. Tutored and untutored birds produced frequency modulations in syllables by modulating the activity of the syringeal muscles.

In both groups, harmonic stacks were generated with low EMG activation of vS muscles (Fig. 5, A and B).

The main difference in song between untutored and tutored birds is how various notes and their corresponding motor gestures are used to compose the song motif. For example, untutored birds include atypically long notes that are modulated in frequency and amplitude at a high rate. This modulation is accompanied by burst-like, high-amplitude activation of the syringeal muscles (Fig. 5A). Tutored birds also produced such burst-like activity of their vS muscles, but these activation patterns and corresponding sound segments were much shorter than in untutored birds.

An analysis of the spectral content of vS muscle activation for tutored and untutored birds shows a difference in their respective use (Fig. 6A). In tutored birds, signal power between 10 and 300 Hz is more evenly spread across the full spectral range. Untutored birds had significantly lower power in the range of 40 – 100 Hz (Kolmogorov-Smirnov test, \( D = 0.05 \)) and a distinct peak in the range of 140 – 180 Hz (Fig. 5A). This difference in the spectral distribution of power is most likely explained by the fact that untutored birds sang uncharacteristically long syllables with high-frequency amplitude and frequency modulations, which are accompanied by EMG bursts that match the modulation frequency (Fig. 5A).

How is muscle activation distributed across the motif? Although per unit time, the integrated variation of the muscular activation patterns does not differ significantly between the songs of the two groups (Kolmogorov-Smirnov test, \( D = 0.05 \) for vS and \( D = 0.32 \) for dB), the distribution of vS activation across the motif is significantly different (see methods; cumulative analysis of variations; Kolmogorov-Smirnov test, \( D = 5 \times 10^{-2} \); Fig. 6B). Activity in untutored birds is typically concentrated toward the

---

**FIG. 2.** In tutored birds, air sac pressure of expiratory pulses is modulated more frequently. A and B: examples showing differences in modulation of pressure between tutored and untutored birds for long duration syllables. Untutored birds produce long syllables with simple, unmodulated pressure patterns despite high-frequency modulations of the acoustic output, which are produced by activity of the syringeal muscles. Air sac pressure patterns are shown as relative voltages and are offset to allow comparison of tutored and untutored birds. Each of the expiratory pulses shown corresponds to a syllable. The spectrograms of these syllables are displayed on the right side and identified by arrows.
end of the motif, whereas in tutored birds, it is more evenly distributed across the motif.

The differences in the temporal structure of muscle activation between tutored and untutored birds go beyond this coarse level. The analysis of the autocorrelation function shows a clear difference in the temporal organization of muscle activation bursts. In tutored birds, autocorrelation functions are less broad, which is consistent with more frequent switching between high and low activation and thus a more variable structure of the EMG pattern (mean of the full width at half-maximum (FWHM) for tutored birds = 9 ms and for untutored birds = 28 ms; Kolmogorov-Smirnov test, $D < 5 \times 10^{-4}$; Fig. 6C). Untutored birds maintain a certain activation level (high or low) for a longer time.

**Bilateral syringeal activation**

The quantification of lateralization of EMG activity ($I_{\text{mot}}$) for the whole motif or individual syllables did not yield significant differences between tutored and untutored birds (Kolmogorov-Smirnov tests, $D = 0.44$ for the motifs and $D = 0.7$ for the syllables). Also, the Pearson’s correlation coefficients for left-right comparisons were not significantly different between both groups (Kolmogorov-Smirnov test, $D = 0.53$). Both analyses show therefore that lateralization of sV activity is not different between tutored and untutored birds.

The beginning of an expiratory pulse of the song motif is typically accompanied by a unilateral or bilateral activation of the syringeal muscles. For these syllable onset events, activation of the left and right sV muscles was more synchronous in tutored birds than in untutored birds, which resulted in a bigger instantaneous cross-correlation at the beginning of the pressure pulse (Kolmogorov-Smirnov test, $D = 0.002$). Figure 7 shows one example of a synchronous burst occurring in the left and right sV muscles. Another example (Fig. 5) also shows this difference; the syllable from an untutored bird (Fig. 5A) shows no synchronous burst, whereas the one from a tutored bird (Fig. 5B) shows again the clear occurrence of this burst in both muscles.

**Speed of syringeal muscles**

To test whether the differences in the sensorimotor learning process impact the biomechanics of vocal production,
we compared the maximal EMG burst rates of the syringeal muscles between tutored and untutored birds. The maximal EMG burst rates for vS muscles were 237.0 ± 14.2 and 217.4 ± 10.6 Hz for tutored and untutored birds, respectively; the difference was not statistically significant (Kolmogorov-Smirnov test, $D = 0.2$). In addition, the relationships between the mean vS EMG activity and fundamental frequency of sound did not differ between tutored and untutored birds. These results suggest that biomechanics of vocal control are not strikingly different between tutored and untutored birds.

In summary, song motor gestures of tutored birds contain modulations more evenly distributed across time and across frequencies. This is consistent with shorter periods of increased activation of the muscles and also shorter periods of time displaying the maximal EMG burst rate. Furthermore, coordination between the left and right sides of the syrinx is different between songs of tutored and untutored birds. How are these differences in motor gestures expressed in the acoustic signals they generate?

**Corresponding acoustic differences**

As was found previously (Fehéř et al. 2009), note duration is significantly longer in untutored birds. The maximal EMG burst rates for vS muscles were 237.0 ± 14.2 and 217.4 ± 10.6 Hz for tutored and untutored birds, respectively; the difference was not statistically significant (Kolmogorov-Smirnov test, $D = 0.2$). In addition, the relationships between the mean vS EMG activity and fundamental frequency of sound did not differ between tutored and untutored birds. These results suggest that biomechanics of vocal control are not strikingly different between tutored and untutored birds.

A second striking feature of untutored songs is the more frequent occurrence of simultaneous, independent fundamental frequencies in harmonic stack elements. Whereas such elements were only present in 2.6% of the harmonic stack segments in tutored birds, they occurred in 28.6% of the elements in untutored birds. The two independent frequencies most likely originate from the two sound generators (Goller and Cooper 2004; Jensen et al. 2007) and may therefore indicate different coordination between the two sides in untutored birds. This difference, however, was not reflected in the integrated analyses of lateralization, in which data from all syllable types were pooled.

**DISCUSSION**

This study provides insight into how the acquisition of an acoustic model early in life shapes the motor control for song production at the level of the respiratory and syringeal motor gestures. Understanding the changes in peripheral motor gestures is a first step toward unraveling the effects of early acoustic experience on the central motor program for song. The results indicate that untutored and tutored birds make similar use of the motor systems over the course of a motif and use
similar respiratory and syringeal motor gestures to generate similar sounds and that there are no functional differences in the achieved activation rate of syringeal muscles. The main differences in motor gestures are found in the temporal features, which are a direct result of the temporal arrangement of acoustic elements within syllables and the motif. We first discuss the implications of these findings on the mapping of motor gestures onto acoustic output and then draw some implications about the effects of an acoustic model on the central control of song production.

Motor control in both groups uses qualitatively similar motor gestures

The exposure to an auditory tutor model results in a refinement of the experience-independent motor gestures and does not lead to qualitatively different motor control patterns. Such a difference could be expected if more complex acoustic output was achieved by an overall increase in muscle use throughout the motif. This expectation is consistent with syringeal biomechanics. Rapid gating of airflow and high-frequency sounds have been associated with high EMG activity in different syringeal muscles (Goller and Suthers 1996a,b). If sophisticated frequency control and rapid modulation of airflow are the physiological control mechanisms for generation of diverse acoustic features, more muscle use would translate into greater acoustic complexity. However, our data suggest that tutored zebra finches do not show increased utilization of muscles.

Another possibility how tutoring might affect motor control is through direct influences on the peripheral musculature. During development, muscle activation patterns could influence the fiber type composition of muscles. Innervation patterns of muscles are known to affect the development of motor units in other systems (Nacarrete and Vrbová 1993). Whereas this study did not directly investigate the organization of syringeal muscles, our data on the burst-like activation patterns suggest that both untutored and tutored birds used equally high burst rates of syringeal muscle activation to generate amplitude and FM patterns in the song. Similar burst rates indicate that syringeal muscles can perform at high contraction speeds in both groups (Elemans et al. 2008) and suggest that very fast motor units are present in both groups.

Exposure to model song causes changes in temporal distribution of qualitatively similar motor gestures

Differences in the temporal activation patterns of the peripheral muscles is an important difference between tutored birds and those without exposure to adult song. This difference is manifested as more frequent modulation of both the respiratory and the syringeal motor gestures over the course of the motif in the songs of tutored birds. It is particularly striking that the respiratory contributions to fine modulation of air pressure and

---

**FIG. 6.** Analysis of the activation patterns of vS syringeal muscles. A: averaged power spectral density of the tutored (black) and untutored (gray) birds. The black horizontal line indicates the region where the signal power of the tutored group is significantly greater than that of the untutored group (40–100 Hz), and the gray line indicates significantly greater power in the untutored group (140–180 Hz). B: cumulative analysis of variations for 2 representative examples of the activation pattern during the motif for the tutored (bottom) and untutored birds. The time is normalized to the duration of the motif. The quantifier $D(t)$ represents the ratio of the cumulative activity of the muscle up to time $t$ if the EMG activity were uniformly distributed over the motif, and $E(t)$ corresponds to the actual ratio of cumulative activity of the muscle up to time $t$. This allows a quantification of how much the distribution of the muscle activity deviates from a uniform distribution across the motif. C: representative examples of the autocorrelation function of the vS muscles for tutored (black) and untutored (gray) birds. Tutored birds switch more frequently between high and low levels of activation.
gestures onto acoustic output. This poses the question of whether additional nonlinearity exists at higher levels in the central control of the song motor program. This translation of motor activity into behavior could represent a more general mechanism of how motor skills are added without increasing the overall muscle activation or energy expenditure. Behavioral change may be effected more generally by the use of nonlinearities in mapping neural instructions onto movement.

**Implications for how experience shapes vocal motor programs**

Although single unit recordings in RA in tutored birds suggest that the output signal to the syringeal and respiratory areas is comprised of the combined activation of individual neurons, we still lack a thorough understanding of the central motor program of song production. More detailed data on functionally connected populations of neurons and their combined output to syringeal and respiratory areas may be necessary to elucidate the motor program, but such information is currently not available. However, syringeal motor gestures reflect motor activity that is directly controlled by premotor activity in RA and thus reflect the population output signal from RA. This motor information parallels insights gained from studying peripheral activation patterns in the process of learning other motor skills (Shadmehr and Holcomb 1997; Shadmehr and Thoroughman 2000; Thoroughman and Shadmehr 1999). However, the song learning system is unique in that it does not only represent a motor learning process but motor learning as a consequence of the acquisition of an acoustic template for song.

Experience-dependent modification of the song motor program seems to operate within an experience-independent framework of motor gestures. These basic gestures may be determined by the neural architecture and emerging motor control mechanisms of the song control network. A restriction of experience-dependent modification to temporal redistribution of a basic set of motor commands may be a general mechanism of vocal motor learning and may therefore reflect aspects of speech development in human infants as well. If these specific modifications to the motor program for birdsong are paralleled by how the motor program for human speech changes with acoustic and social experience (Goldstein et al. 2003), the notion of an experience-independent basic speech program is strongly supported (Chomsky 1975). The effects of skill learning on motor control in other systems are consistent with this interpretation. For example, the effect of learning a reaching task by rodents is manifested in a selection and tuning of an already present motor program (Kargo and Nitz 2003).

The two experimental groups differed in their experience with an acoustic model early in life, whereas both underwent the second learning component of vocal ontogeny, the sensorimotor learning phase. However, interactions between the acquisition of an acoustic model and the subsequent perfection of self-generated vocalizations are likely because the development of the neural architecture of the motor circuitry is directly affected by the acquisition of an acoustic model (Wilbrecht et al. 2006) and therefore must influence the sensorimotor learning process. This presumed interactive process may provide a mechanistic explanation for the observation that social experience, and not specific features of the acoustic model, may...
cause enhanced temporal structure in the songs of birds tutored by males singing untutored song (Fehér et al. 2009). Because the tutored birds used in this study also had social experience during vocal ontogeny, this interaction cannot be excluded. Irrespective of which aspect of experience led to the observed changes in motor control and acoustic output, the interactive aspects of experience and development of neural circuitry and control patterns most likely play an important role. It is therefore important that the changes to the motor program have been studied at the endpoint of vocal ontogeny, because experimental manipulation during development may affect this interaction.

ACKNOWLEDGMENTS

We thank the reviewers for the thoughtful and constructive comments.

GRANTS

This work was supported by National Institute on Deafness and Other Communication Disorders Grants DC-04390 and DC-06876.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

REFERENCES


