Journal of Comparative Physiology A
https://doi.org/10.1007/s00359-017-1231-3

ORIGINAL PAPER

From electromyographic activity to frequency modulation in zebra finch song

Juan F. Döppler[1] · Alan Bush[1] · Franz Goller[2] · Gabriel B. Mindlin[1]

Received: 11 September 2017 / Revised: 14 November 2017 / Accepted: 17 November 2017

Abstract
Behavior emerges from the interaction between the nervous system and peripheral devices. In the case of birdsong production, a delicate and fast control of several muscles is required to control the configuration of the syrinx (the avian vocal organ) and the respiratory system. In particular, the syringealis ventralis muscle is involved in the control of the tension of the vibrating labia and thus affects the frequency modulation of the sound. Nevertheless, the translation of the instructions (which are electrical in nature) into acoustical features is complex and involves nonlinear, dynamical processes. In this work, we present a model of the dynamics of the syringealis ventralis muscle and the labia, which allows calculating the frequency of the generated sound, using as input the electrical activity recorded in the muscle. In addition, the model provides a framework to interpret inter-syllabic activity and hints at the importance of the biomechanical dynamics in determining behavior.

Keywords Birdsong · Biomechanics · Nonlinear dynamics · Frequency modulation · Electromyography

Introduction

Behavior requires the central nervous system (CNS) to generate instructions, which ultimately drive a set of biomechanical apparatuses, which in turn interact with the environment. These instructions are transduced by muscles into time-dependent forces responsible for changing the configuration of the biomechanical systems. This translation is dynamical in nature. Birdsong is a great example of this process. Its production requires a delicate (and extremely fast) control of a peripheral biomechanical organ, the syrinx, which is achieved through the generation of several gestures by the CNS.

A seminal work (Goller and Suthers 1996a) revealed that the air sac pressure establishes the necessary conditions for phonation, generating adequate flow through the syringeal lumen. It also showed that gating is actively controlled by some syringeal muscles and that the frequency modulations during the production of a syllable are affected by the configuration of the syringealis ventralis (vS) muscle. This picture emerged from studies performed on brown thrashers (*Toxostoma rufum*). In particular, the study showed a clear relationship between the average electromyographic (EMG) activity across a syllable and its mean fundamental frequency.

Yet, it was soon realized that the translation of gestures into acoustical features is complex and species specific. The study of song production by the zebra finch (*Taeniopygia guttata*) shows, for example, a complex relationship between the syllabic fundamental frequency and the pattern of the vS muscle activation (Vicario 1991). Unlike in the brown thrasher, the relationship between vS EMG activity and fundamental frequency shows no distinct relationship for high-frequency syllables (Goller and Riede 2013). Zebra finches present distinctive bursts of activity in the vS muscles right before syllable onsets, and in many cases, bursts of activity right after syllable endings. These are not features with an obvious interpretation in terms of frequency control, which seemed to be the primary role in other species like the brown thrasher. The relationship between the syllabic mean activity at the vS muscle, and the mean fundamental frequency holds for many species, but as soon as the temporal modulation within a syllable is inspected, the relationship does not hold for every syllable type.

✉ Juan F. Döppler
jdoppler@df.uba.ar

[1] Department of Physics, FCEyN, University of Buenos Aires, and IFIBA, CONICET, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina

[2] Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112, USA

Published online: 23 November 2017

This seminal study opened a wider investigation on the roles played by the different muscles, in different contexts, for different species. This led to the realization, for example, that in some species, frequency modulation might be mostly controlled through pressure, like it was the case for the suboscine Great Kiskadee (*Pitangus sulphuratus*) (Amador et al. 2008). Even the simplest theoretical models of birdsong production require a synergistic interaction between the control parameters to determine the frequency modulations of the song (Alonso et al. 2014). In fact, the multi-functional and context-dependent nature of the frequency modulation in birdsong production was verified experimentally in studies performed on Bengalese finches (*Lonchura striata domestica*) (Srivastava et al. 2015). An additional difficulty arises from the fact that the electrical activity of the vS muscle measured by electromyography affects the vocalized fundamental frequency indirectly. Electrical activity generates muscle contractions that indirectly affect the tension of the vibrating labia via movement of cartilage components, which provide the main insertion sites for the muscle. Thus, we have to understand how to translate electrical activity into syringeal configuration changes to unveil its association with frequency modulation. This will be particularly important in syllables presenting fast frequency modulations, for which EMG envelopes might not be an immediate proxy for labial stretching.

In this work, we present a simple model for muscle and labial dynamics that allows transducing electrical activity in the syringeal muscles into the fundamental frequencies of the recorded songs. We present this model in the Materials and methods section, in which we also discuss the experimental methods for recording EMGs. In this section, we present the fitting procedures carried out to reconstruct the model´s parameters as well. The data for one bird are used to present these methods. In the Results section, we introduce and analyze data from four other birds, and in the Discussion section, we discuss our results.

Materials and methods

We plan to test the hypothesis that with a biomechanical model of muscle activation and labial dynamics, it is possible to transduce EMG recordings of the vS muscle into frequency modulations, consistent with the song uttered, while the EMG activity was recorded. Our approach requires then the development of a theoretical model capable of translating EMG into frequency, as well as the experimental methods for measuring the EMGs in muscles during birdsong production. For this reason, we organize this section in two parts: the theoretical methods are described first, followed by the experimental ones.

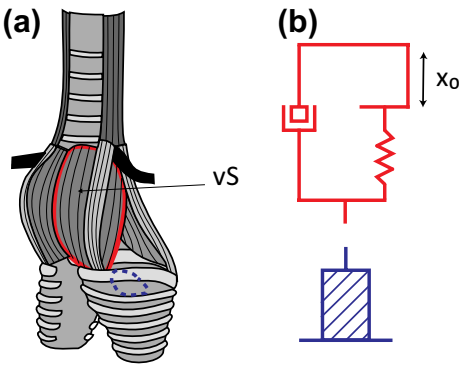

Fig. 1 **a** Oscine syrinx. Schematic illustration of the tracheobronchial syrinx, located at the junction of the trachea and the bronchi. The muscle syringealis ventralis (red) is inserted in the second bronchial ring, which is also connected to the vibrating labia (blue). **b** Schematic representation of the model. The muscle is modeled as a spring with variable slack and dissipation (red). The labium (blue) is modeled as a string subjected to the force generated by the muscle

Theoretical methods

In Fig. 1a, we display a sketch of the tracheobronchial syrinx. It is a bipartite structure at the juncture between the bronchi and the trachea. There are two pairs of labia, one in each bronchus. When airflow is established between these labia, they are set in motion, and sound is generated. The syringeal muscles can change the configuration of the syrinx, stretching or relaxing the oscillating tissue, which generates frequency modulation. For example, the ventral syringeal muscles (right and left vS muscles) connect to the second cartilaginous annulus. Therefore, the contraction of the vS muscle of either side results in the stretching of the labia located between the second and the third annuli of the respective bronchi. The schematics at the right (Fig. 1b) illustrate the muscle model (in red), which consists of a sliding slack (of length x_0) in parallel with a dissipative component. In the model, the slack´s length is controlled by the electrical activity in the muscle. The labium is modeled through a stretching viscoelastic element (in blue), subjected to the force exerted by the muscle (see Fig. 1b).

As the electrical activity shortens the slack, the difference between the muscle´s actual length (x_1) and the slack (x_0) is transduced into force (Shapiro and Kenyon 2000) through a nonlinear (exponential) function, which reflects the fact that muscles produce force efficiently on contraction ($x_0 < x_1$) but not on elongation ($x_1 < x_0$). In this way, their dynamics are ruled by the following equations:

$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_2}{dt} = -k_{vS}\left(x_1 - x_0(t)\right) - \beta_{vS}x_2,$$

where k_{vS} and β_{vS} represent the elastic and viscous properties of the muscle, and the relationship between the slack and the measured EMG is given by

$$x_0\,(t) = \Delta_0 - \delta g\,(\mathrm{EMG}(t)),$$

with the function $g\,(\mathrm{EMG}(t))$, the envelope of the rectified electromyogram. Once the dynamics of the muscle is computed (i.e., once x_1 is computed), the force can be estimated as

$$f(t) \equiv \alpha e^{B\left(x_1 - x_0(t)\right)-1},$$

with B a constant to be determined phenomenologically (Rokni and Sompolinsky 2012).

The second part of our model consists of a viscoelastic element representing the labium, which is stretched by the force exerted by the muscle. We can describe it in terms of its length x_3, whose dynamics will be ruled by

$$\frac{\mathrm{d}x_3}{\mathrm{d}t} = x_4$$

$$\frac{\mathrm{d}x_4}{\mathrm{d}t} = -k_l\,(x_3) - \beta_l x_4 + f,$$

where f stands for the muscle force, β_l represents dissipation in the labium, and k_l its elastic properties. The scaling of the equations of the model allows finding the minimal number of parameters (time constants) determining the system's behavior, which leads us to the final set of equations for our model:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = x_2$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \frac{1}{\tau_2}\left(-x_2 + \frac{1}{\tau_1}\,(-x_1 + x_0)\right)$$

$$f(t) \equiv A e^{\mathrm{B}\,\left(x_1 - x_0(t)\right)-1}$$

$$\frac{\mathrm{d}x_3}{\mathrm{d}t} = x_4$$

$$\frac{\mathrm{d}x_4}{\mathrm{d}t} = \frac{1}{\tau_4}\left(-x_4 + \frac{1}{\tau_3}\,(-x_3 + f)\right).$$

These scaling parameters are related to the physical ones by these relationships:

$\tau_1 = \beta_{vS}/k_{vS,}\ \tau_2 = 1/\beta_{vS},\ \tau_3 = \beta_l/k_l,\ \tau_4 = 1/\beta_l,\ A = \alpha/k_l$. The time constants are measured in seconds.

Experimental methods

To record the EMG, the syringeal muscle was exposed by an incision in the skin and the interclavicle air sac between the clavicles (Goller and Suthers 1996b). A pair of bipolar electrodes was inserted in either the left or right syringealis ventralis (vS) muscle, and secured with a drop of surgical tissue adhesive. For one bird, we inserted bipolar electrodes in both right and left vS muscles. In all the cases, the wires were led subcutaneously to the back, and connected to signal conditioning electronics, which amplified (× 225) and filtered (high pass filter + 150 Hz) the signals. After recovery from the surgery, the EMG activity and the song were simultaneously recorded.

The first step in the processing of the data consisted of computing the signal´s envelope. To do so, we computed the maxima of the rectified signal in 7 ms windows. In Fig. 2a, we display the rectified EMG, and the envelope in Fig. 2b.

As discussed in the section describing the theoretical methods, we propose for the slack that $x_0\,(t) = \Delta_0 - \delta g\,(\mathrm{EMG}\,(t))$ where the function $g\,(\mathrm{EMG}\,(t))$ stands for the EMG´s envelope. The values of the parameters were chosen as Δ_0, $\delta = (0.005, 0.004)$

Since our fitting aimed at reproducing the song´s fundamental frequency (F_0) modulation, the first step in the processing of the song was to compute its spectrogram. To do so, we used custom written software using Python´s "scipy" library. From these spectrograms, we extracted the fundamental frequency, by choosing at least four fundamental frequency values at times t_j, $j = 1, 2, 3, 4$ per syllable, and extrapolating the values for the rest of the sampling times. In this way, we generated a set of values $\omega_0\,(t)$, $i = 1, \ldots, N$, with N the number of sampling times needed to cover a syllable. To compare these reconstructed values with the output of our model, we assumed that the frequency is related to the labial stretching through

$$\varpi = a\sqrt{x_3} + b,$$

where the parameters (a, b) are found by linear regression. We ran the equations of the model for different sets of the parameters (τ_1, $\tau_{2,}$ $\tau_{3,}$ B), with an ODE integrator (Runge–Kutta of 4th order). Notice that τ_4 was not included among the parameters to be fitted. We found that as long as $\tau_4 \ll \tau_3$, the solutions would not differ significantly. For each set of parameter values, we computed

$$\varsigma = \sqrt{\sum_{i=1}^{N} \frac{1}{N}\left(\omega_0(t_i) - \left(a\sqrt{x_3(t_i)} + b\right)\right)^2}$$

We changed the parameters through a genetic algorithm, searching for the set that would minimize ς.

Each step of our genetic algorithm consisted of 30 numerical simulations of the model´s equations, each one using a different set of parameters τ_1, $\tau_{2,}$ τ_3, and B. For each simulation, the value of ς was computed, and was

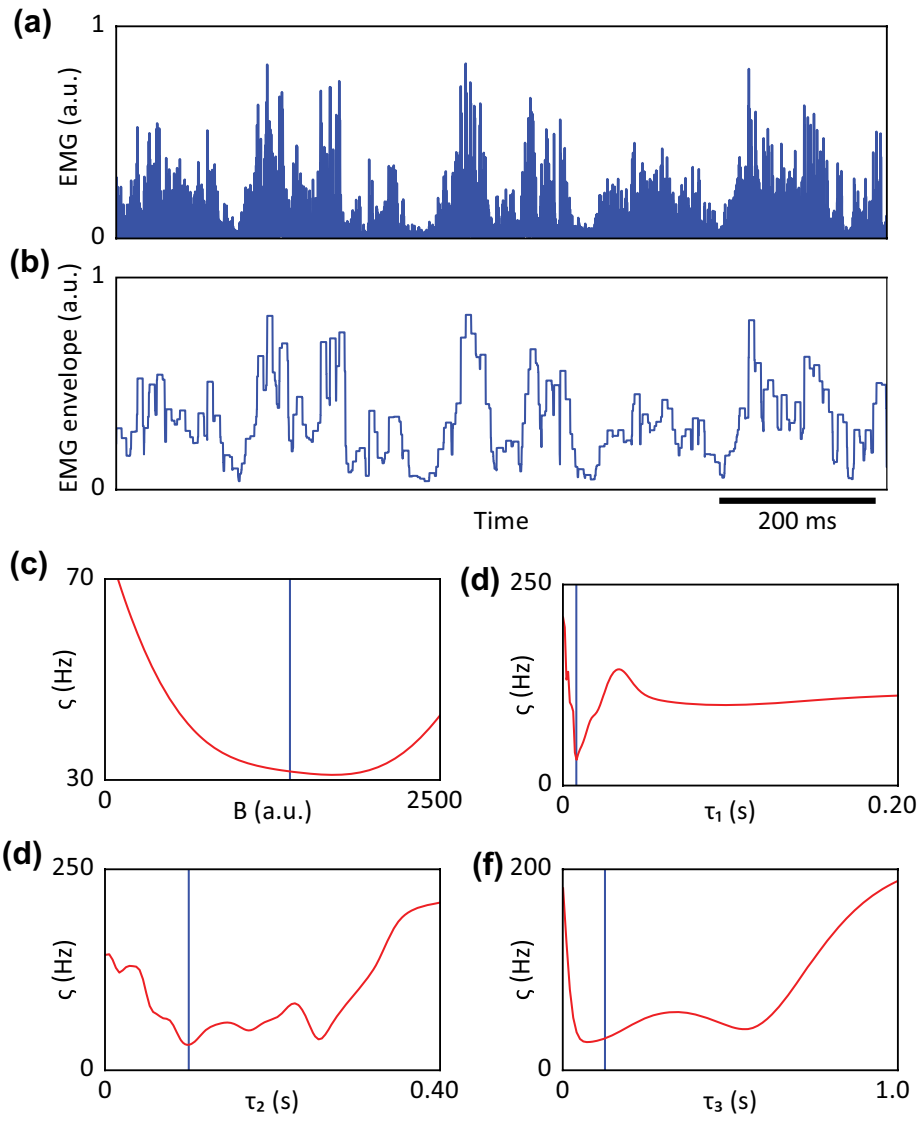

Fig. 2 **a** Rectified EMG signal collected from the vS muscle as described in the experimental methods. **b** Envelope of the rectified signal. **c–f** ς as a function of each of the parameters. The parameters were changed one at a time, around the values obtained by the genetic algorithm (marked in blue)

used to order the set of parameters. The seven sets giving the smallest ς values were retained for the next step of our procedure. The remaining 23 sets of parameters were constructed in the following way. For each set, a random number between 0 and 1 was generated. If this number exceeded 0.8, the set was also retained. Then, for each of the retained sets, each parameter was changed with a probability of 25%. The last step of the procedure consisted of completing the initial number of sets (30), through a procedure that is called "crossing"; the parameters of these new sets are defined as (${\tau_1}^i$, ${\tau_2}^i$, ${\tau_3}^j$, and B^j), where the superindexes (i, j) denote parameters sets already retained. This procedure was repeated for at least 200 steps. To test whether the final values corresponded to minima of ς, we explored the values obtained for parameters in the neighborhood of the ones obtained through the genetic algorithm. These explorations are displayed in Fig. 2c–f, for the simulations carried out for one of the birds. In these figures, a blue line indicates the actual values returned by the genetic algorithm. Notice that even in the cases for which the returned values did not correspond to minima, both in terms of ς and the actual parameter values, the differences were minimal.

For the values that were obtained using the genetic algorithm described above, we performed the simulations of the model and obtained the time traces displayed in Fig. 3. In Fig. 3a, we show the slack length x_0. In Fig. 3b, we show the difference between the muscle length and the slack value ($x_1 - x_0$), and the resulting force exerted by the muscle in this dynamical process is shown in Fig. 3c. Finally, the dynamics of the variable representing the labial stretching is shown in Fig. 3d.

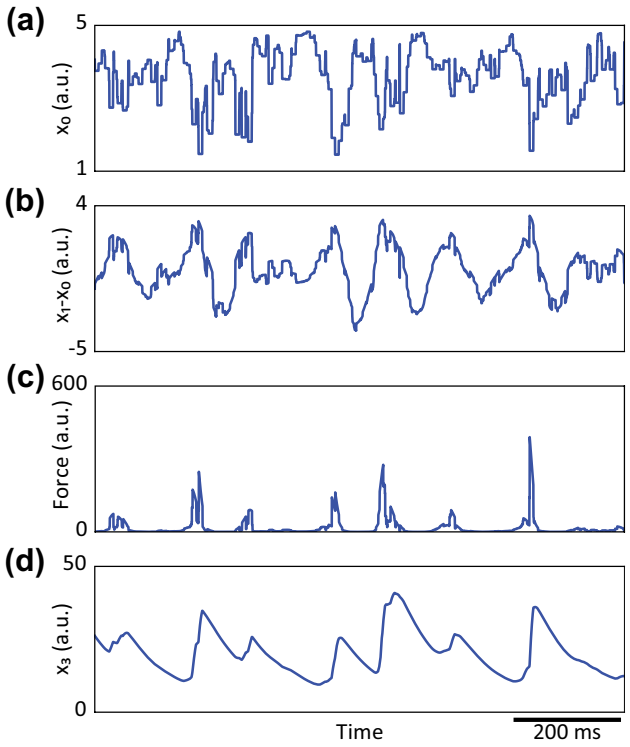

Fig. 3 Intermediate steps of the model. **a** Slack length of the muscle (x_0) as a function of time. **b** Difference between the actual length of the muscle (x_1) and the slack length. **c** Force exerted by the muscle on the labia as a result of this length difference. **d** Stretching of the labia (x_3) as a function of time

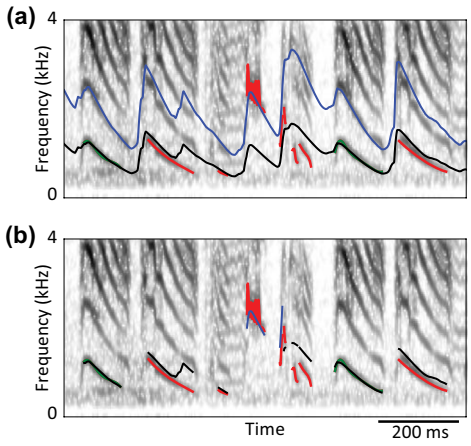

Fig. 4 Predicted frequency of the model for bird AB009. **a** Time trace of the predicted frequency of the model (black line) after its parameters were fitted using the fundamental frequency of only two syllables (green), which gives a good approximation for the rest of the low-frequency syllables. In blue, time trace of twice the computed frequency. Notice that it approximates the high-frequency syllables. **b** Model frequency for each syllable, selecting either the predicted frequency (black) or twice its value (blue)

Results

In Fig. 4a, we display the spectrogram of a zebra finch song. As it is typically the case in this species, we observe two types of syllables: those with low (i.e., < 1.3 kHz) and high (i.e., $\gtrsim 1.3$ kHz) mean fundamental frequencies. We selected two segments of the sound data, corresponding to the syllables, whose fundamental frequencies are painted in green (Fig. 4a). The fundamental frequency of the rest of the syllables (of either high or low fundamental frequency syllables) is painted in red. The two syllables selected for the fitting are low-frequency ones, and therefore, we will assume that they are generated with the two syringeal sound sources (Goller and Cooper 2004; Jensen et al. 2007). Using only those two segments, we fitted the parameters of the model of muscle and labium, and integrated the equations for the entire time lapse of the recording. The resulting frequency predicted by the model, using as input the EMG of the left vS muscle, is illustrated with a black continuous line. We quantified the goodness of the fit using the reduced Chi-squared:

$$\chi^2_{\text{reduced}} \equiv \frac{1}{\upsilon}\sum_{1}^{N}\left(\omega_0\left(t_i\right) - \varpi\left(t_i\right)\right)^2/\sigma_i^2,$$

with υ the number of degrees of freedom (number of points - number of fitted parameters $-$ 1), and σ_i the error of the $\omega_0\,(t_i)$. Since the temporal window Δt used in the computation of the fundamental frequencies was $\Delta t \leq 20$ ms, we used $\sigma_i = 1/\Delta t = 50$ Hz. For this bird, we obtained $\chi^2_{\text{reduced}} = 0.4$, for the fitted syllables.

Notice that despite not being used to fit the model parameters, the fundamental frequency of some syllables is well approximated by the black line. On the other hand, the fundamental frequency of the high-frequency syllables is well out of the range of the model´s prediction. Notably, if we compute twice the value of the prediction (blue line in Fig. 4), we can obtain a good approximation for those fundamental frequencies.

The song of the zebra finch is generated with the two sides of the syrinx. Some syllables present fundamental frequencies of less than about 1.3 kHz, and are generated using both sides of the syrinx in a 1:1 locked dynamical state. Other syllables of higher fundamental frequencies are generated with the right syringeal side only (Goller and Cooper 2004). In this way, to generate the high-frequency syllables, the left syringeal valve has to be closed. Since the speed of the air through the valves is much smaller than the speed of sound (i.e., Mach numbers much smaller than one), the air can be considered an incompressible gas. In this way, it is reasonable to assume that closing the left valve leads to a duplication of the flow through the right valve. It has been reported that the mucosal wave along the labia within the syringeal valves is proportional to the airflow through them (Doellinger and Berry 2006; Boessenecker et al. 2007).

Moreover, in the case of the zebra finch vocal organ, a linear relationship between the resulting fundamental frequency and the mucosal wave velocity has been measured (Elemans et al. 2015). In this way, closing one valve would lead to at least a duplication of the fundamental frequency of the sound generated by the valve´s operation, with respect to the fundamental frequency values that would be obtained with both sides open. Notice that within these hypotheses, if the valve that remains opened is also slightly contracted, the speed of the air through it can be even higher, leading to an even greater increase in frequency. This picture is consistent with the reported observation that low-frequency sounds are generated by pulse like labial oscillations, while high-frequency ones are more tonal (Jensen et al. 2007): when the flow is doubled, the labia are less likely to collide against each other in each cycle, leading to a smoother, more tonal modulation of the airflow.

Consistent with this conjecture, once we select the standard output of the model (black curves in Fig. 4a) for syllables of low fundamental frequency, and the frequency-doubled output (blue curves in Fig. 4a) for high fundamental frequency syllables, we obtain a good approximation for all the fundamental frequencies. This final result is illustrated in Fig. 4b.

In Fig. 5, we show an alternative strategy, for the same bird. We fitted the two segments in green (the ones that were fitted with the frequency-doubled output in the previous discussion), and integrated the equations of the muscle-labial model. The result of the computed fundamental frequency is displayed in black. Then, we halved the resulting fundamental frequency, and found that the predicted time trace was an excellent proxy for the low-frequency syllables (blue line). The selected segments are displayed in Fig. 5b.

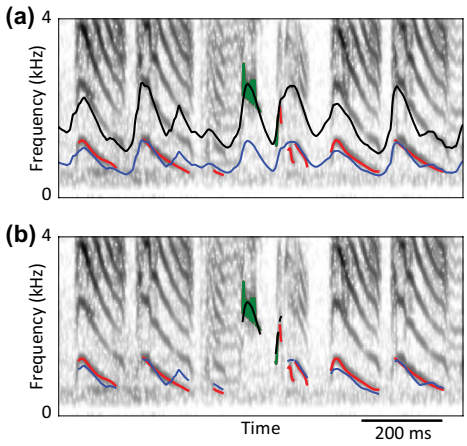

Fig. 5 Alternative strategy for fitting the model. **a** Predicted frequency (black) and its half (blue) when fitting the model using only high-frequency syllables (green). Notice that half the predicted frequency gives a good approximation to the frequency modulation of the low-frequency syllables. **b** Model frequency for each syllable, selecting either the predicted frequency (black) or half its value (blue)

We proceeded to repeat this analysis for four additional birds (apart from the one used to present the methods). The results are shown in Fig. 6. In all cases, the green time traces correspond to the fitted segments. In the cases illustrated in Fig. 6a and b, it is safe to assume that these segments correspond to low fundamental frequency syllables. In both cases, the black segments correspond to the predicted frequency, while the blue segments correspond to twice this frequency. The computations used EMG recorded from the left vS muscle.

In the case displayed in Fig. 6c, the selected segments presented fundamental frequencies starting around 1400 Hz, and therefore, we assumed that they were high-frequency syllables. With those values, we fitted our models, plotted the predicted frequencies with a black line, and computed a second time trace with the halved frequencies (blue line). As we show in Fig. 6c, the blue line approximates well the fundamental frequencies of the other syllables in the song.

For Fig. 6d, we used EMG recordings from both (right and left) vS muscles. The green lines are the fundamental frequencies of the segments that were fitted. The red segments represent the fundamental frequencies of the rest of the song. The black lines correspond to the fundamental frequency obtained from integrating the equations of the model using as input the EMG from the left side. The magenta lines correspond to the fundamental frequency obtained by fitting the same segments, but using as input the EMG of the right side. Finally, the blue lines are twice the fitted values that were computed using the EMG of the right side. It is worth remarking that in all the cases, the fitting segment was a small fraction of the total song. Moreover, syllables presenting a frequency modulation qualitatively different from those used to fit the model´s parameters could be well approximated by the model.

For all the cases analyzed, we report the values of the fitting parameters obtained through the genetic algorithm, and tested that these parameters were close to minima of ς. These results are displayed in Table 1. The averages and standard deviations are $\left(\langle\tau_1\rangle \pm \sigma_{\tau_1}, \langle\tau_2\rangle \pm \sigma_{\tau_2}, \langle\tau_3\rangle \pm \sigma_{\tau_3}\right) = (0.03 \pm 0.02, 0.06 \pm 0.04, 0.18 \pm 0.08)$.

A remarkable feature in the EMG of the vS muscle of zebra finch during song production (Vicario 1991; Goller and Cooper 2004) is the existence of activity bursts both at the beginning and end of many low-frequency syllables. It is tempting to conclude from that observation that these patterns of activity are associated with gating. However it is possible to interpret the patterns in terms of frequency modulation as well. According to our simulations of the muscle and labial model, the peaks of the forces at the beginning of the syllables are followed by a relaxation of

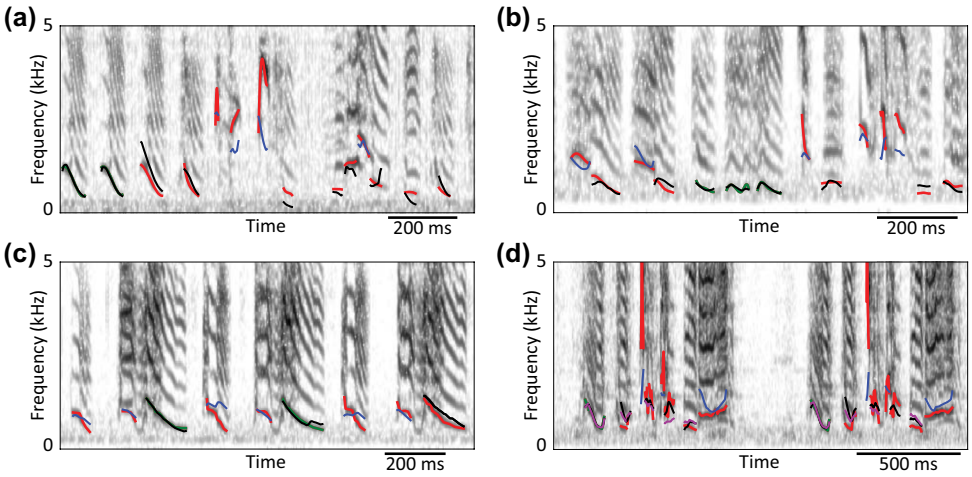

Fig. 6 Fitting of four other birds. In all cases, the green segments represent the fitted syllables, while the red ones are the fundamental frequency of the rest of the song. **a, b** Fittings (black) using low-frequency segments and its frequency-doubled segments (blue) allow us to reconstruct the song for birds ZFMCV and AB005, respectively. **c** Fitting (black) using two high-frequency segments and its frequency-halved segments (blue) for bird CH012. **d** In the case of bird 411, the two sides of the vS were measured. The black segments correspond to the frequency predicted using the left side EMG, while the magenta ones correspond to the right side EMG. The blue segments correspond to the frequency doubled of the predicted frequency of the right side. The correlation between the envelopes of the right and left EMGs is $C = 0.74 \pm 0.2$, in 50 ms windows. In each case **a**–**d**, the goodness of the fit was quantified by the reduced chi-squared, which for the cases in **a**–**d** were $\chi^2_{\text{a,reduced}} = 0.26$, $\chi^2_{\text{b, reduced}} = 0.39$, $\chi^2_{\text{c, reduced}} = 1.85$, $\chi^2_{\text{d right,reduced}} = 0.92$, $\chi^2_{\text{d left,reduced}} = 1.0$

Table 1 Fitted parameters of the model obtained for each of the birds (and for both syringeal sides in the case of bird 411). These were obtained using a genetic algorithm to minimize the difference between the predicted frequency and the song frequency

	B	τ_1	τ_2	τ_3
AB009	1380	0.008	0.1	0.126
AB005	557	0.006	0.051	0.209
CH012	448	0.085	0.13	0.169
ZFMCV	730	0.029	0.006	0.082
411-left side	86	0.022	0.031	0.329
411-right side	216	0.027	0.038	0.154

the labium. This explains the slowly decaying fundamental frequencies in the typical down-sweeps of zebra finch song. Most remarkably, the activity bursts ("kicks") at the end of the syllables allow labial stretching to reach values, such that the following activation (i.e., immediately before the next syllable) is sufficient to provide a labial tension compatible with the attained fundamental frequency. In Fig. 7, we display four down-sweeps, where forcing kicks exist at the beginning and end of the syllables (Fig. 7a), as well as the fitted force (Fig. 7b), variable x_3 (Fig. 7c) and frequency (Fig. 7d). Notice that without the kick received at the end of a syllable, the activation before the start of the following syllable would not place the labial stretching at a value required to obtain the observed fundamental frequency. This mechanism bares some resemblance with the co-articulation found in human vocalizations.

Other methods had been proposed in the past to link EMG measurements with the acoustic features of the sound. The most frequently used one involves the computation of the envelope of the rectified EMG data and assumes that muscle activation is correlated with labial tension. We looked how the fit of our model with the fundamental frequency of the sound compares to the results obtained using envelopes. To do so, we computed the envelope of the rectified EMG data for the five birds in our study using two different methods: the one described in the Methods section, which we call box-envelope, and the Hilbert transform method (Boari et al. 2015). Then, we computed the optimal delay between EMG activation and its putative effect on fundamental frequency as the delay with the highest correlation between the envelope and the square of the fundamental frequency of the syllables (optimal delay OD = − 8 ms). Then, we fitted the fundamental frequency, using the same data segments as we used to fit the parameters of our physical model. Finally, we calculated the reduced χ^2 using the shifted envelopes. The average ratio between the reduced χ^2 using our model ($\chi_{\text{model}}{}^2$) and the box-envelope method ($\chi_{\text{box}}{}^2$) is $<\chi_{\text{model}}{}^2/\chi_{\text{box}}{}^2> \sim 0.17$, while with the Hilbert method, we obtained $<\chi_{\text{model}}{}^2/\chi_{\text{Hilbert}}{}^2> \sim 0.13$

Discussion

We were successful in transducing EMG into frequency for a wide range of syllable types found in the songs of the zebra finch. We managed to do so through the use of a dynamical

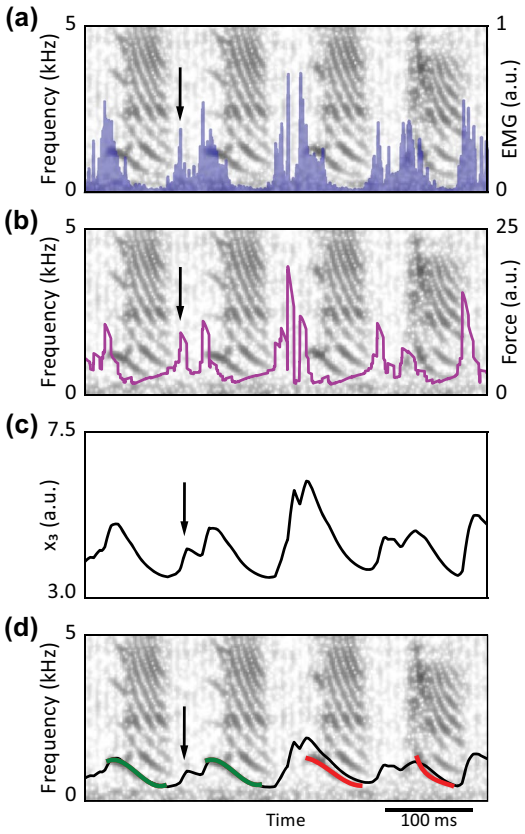

Fig. 7 Role of the inter-syllabic activity. **a** Rectified EMG is presented in blue, over the song sonogram. The arrow marks a burst of activity in a silent period. **b** Force generated by the muscle (magenta). Notice the force 'kick' after the syllable ends. **c** Labial stretching. The labium is quickly stretched after the first syllable, in preparation for the next one. **d** Predicted frequency (black) over the sonogram. The green segments are the fitted syllables, and the red ones, the fundamental frequency of the rest of the song

model for the joint dynamics of a system consisting of an elastic labium attached to a muscle. This allows advancing further from the relationship between average activity and fundamental frequency, which was reported in a seminal work on syringeal control (Goller 1996a).

The measurements of EMG from syringeal muscles in zebra finches posed some puzzles. The observation of activity bursts in the syringealis ventralis muscles right before syllable onset, as well as right after syllable offset, suggested a role in gating, somewhat independent of the frequency modulation. Surprisingly, these EMG bursts are associated, in the framework of our model, with the frequency modulation for most syllables in the zebra finch. In many cases, this modulation consists of a monotonic decrease in frequency value. According to the interpretation provided by our model, this modulation is, therefore, the passive decay of the stretching after a kick, and therefore, it emerges as a coarse instruction is transduced into the smooth temporal evolution of a biomechanical parameter.

In several cases (Fig. 7 displays an example), the EMG presents a peak of activity at the end of the syllable (indicated by the arrows in Fig. 7a). Interestingly, this allows the activity burst right before the onset of the next syllable to induce the right initial tension value. In Fig. 7b we display the peaks of the force exerted by the muscle, and in Fig. 7c the computed value of the labial stretching (x_3). Notice the rapid increases in labial stretching between the syllables due to the forces right after the syllable offsets. For completeness, the fitted frequency is displayed in Fig. 7d.

Surprisingly, in all the cases that we reported, fittings from small segments of the temporal trace allowed the reconstruction of the frequency modulation in the whole song. It is also remarkable that the high-frequency syllables could be approximated computing twice the value obtained fitting the low-frequency syllables. It is known that when the high-frequency sounds are generated, airflow is established through the right side only (Goller and Cooper 2004). We have shown the plausibility of accounting for this phenomenon if we assume incompressibility of the fluid, a hypothesis consistent with the Mach numbers involved. If the right and left vS instructions are highly correlated, the signals from either side can be used as a proxy for the other. This is not necessarily the case, since the degree of correlation between right and left side varies from bird to bird. In the case for which we measured both sides in this work, the correlation was high (see values in the caption of Fig. 6d).

In this work, we have built models for the syringeal muscles and labia to account for the modulation of the fundamental frequencies. It is known that air sac pressure does have some influence on the resulting fundamental frequency. Yet, its role in modulation is secondary. Nerve cut experiments in zebra finches (i.e., cuts of the tracheo-syringeal nerve) have shown that the without syringeal muscle activity, the syllables present almost no modulations (Goller and Cooper 2004; Riede et al. 2010). It is also known that other muscles are active during phonation. Vicario reported that during some zebra finch calls, the syringealis ventralis muscle is silent while the dorsalis muscle is active (Vicario 1991). Moreover, in a recent work (Düring et al. 2017) it has been shown that the musculus syringealis dorsalis medialis (MDS) attaches to cartilaginous pads embedded in the medial labium. It must, therefore, influence the tension of the labia directly and, therefore, modify the fundamental frequency during song production. Furthermore, the gating of airflow via movement of the labia by the ventral and dorsal tracheobronchial muscles also affects fundamental frequency [e.g., (Mencio et al. 2017) and (Laje et al. 2002)]. It is, therefore, likely that during the syllables, where in our examples the fitting is less successful, the synergistic contributions of the other syringeal muscles play a larger role in frequency control.

The models presented in this work are ultimately an integration of the rectified EMG. For slow modulations, the predictions of this model would not differ much from the rectified EMG´s envelope. In this way, we can put the results obtained from different species in the same framework, allowing us to explain the apparently discrepant results from brown thrasher (Goller 1996a) and zebra finch with its rapid frequency modulations. For the case of rapid modulations, it is crucial to account for the muscle´s dynamics: the net force is the result of the mismatch between muscle length and slack, which might be non trivial in the case of rapid EMG variations.

It has been argued that behavior emerges from the interaction between a central nervous system (CNS) and a peripheral device (Chiel and Beer 1997). The fact that the frequency modulations of many zebra finch syllables are the passive decay of an elastic labium after being rapidly "kicked" is a clear and dramatic example of this observation, and leads to a clear hint on how the CNS controls the periphery.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Funding This work describes research partially funded by National Council of Scientific and Technical Research (CONICET), National Agency of Science and Technology (ANPCyT), University of Buenos Aires (UBA) and National Institute of Health through R01-DC-012859.

Ethical approval All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

References

Alonso R, Goller F, Mindlin GB (2014) Motor control of sound frequency in birdsong involves the interaction between air sac pressure and labial tension. Phys Rev E 89(3):032706

Amador A, Goller F, Mindlin GB (2008) Frequency modulation during song in a suboscine does not require vocal muscles. J Neurophysiol 99(5):2383–2389

Boari S, Perl YS, Amador A, Margoliash D, Mindlin GB (2015) Automatic reconstruction of physiological gestures used in a model of birdsong production. J Neurophysiol 114(5):2912–2922

Boessenecker A, Berry DA, Lohscheller J, Eysholdt U, Doellinger M (2007) Mucosal wave properties of a human vocal fold. Acta Acust United Acust 93(5):815–823

Chiel HJ, Beer RD (1997) The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. Trends Neurosci 20(12):553–557

Doellinger M, Berry DA (2006) Visualization and quantification of the medial surface dynamics of an excised human vocal fold during phonation. J Voice 20(3):401–413

Düring DN, Knörlein BJ, Elemans CPH, Schmidt M, Goller F, Riede T, Rome L (2017) In situ vocal fold properties and pitch prediction by dynamic actuation of the songbird syrinx. Sci Rep 7(1):11296

Elemans CPH, Rasmussen JH, Herbst CT, Düring DN, Zollinger SA, Brumm H, Srivastava K, Svane N, Ding M, Larsen ON, Sober SJ, Švec JG (2015) Universal mechanisms of sound production and control in birds and mammals. Nat Commun 6:8978

Goller F, Cooper BG (2004) Peripheral motor dynamics of song production in the zebra finch. Ann NY Acad Sci 1016(1):130–152

Goller F, Riede T (2013) Integrative physiology of fundamental frequency control in birds. J Physiol Paris 107(3):230–242

Goller F, Suthers RA (1996a) Role of syringeal muscles in controlling the phonology of bird song. J Neurophysiol 76(1):287–300

Goller F, Suthers RA (1996b) Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. J Neurophysiol 75(2):867–876

Jensen KK, Cooper BG, Larsen ON, Goller F (2007) Songbirds use pulse tone register in two voices to generate low-frequency sound. Proc R Soc Lond B Biol Sci 274(1626):2703–2710

Laje R, Gardner TJ, Mindlin GB (2002) Neuromuscular control of vocalizations in birdsong: A model. Phys Rev E 65(5):051921

Mencio C, Kuberan B, Goller F (2017) Contributions of rapid neuromuscular transmission to the fine control of acoustic parameters of birdsong. J Neurophysiol 117(2):637–645

Riede T, Fisher JH, Goller F (2010) Sexual dimorphism of the zebra finch syrinx indicates adaptation for high fundamental frequencies in males. PLoS One 5(6):e11368

Rokni U, Sompolinsky H (2012) How the brain generates movement. Neural Comput 24(2):289–331

Shapiro MB, Kenyon RV (2000) Control variables in mechanical muscle models: A mini-review and a new model. Motor Control 4(3):329–349

Srivastava KH, Elemans CP, Sober SJ (2015) Multifunctional and context-dependent control of vocal acoustics by individual muscles. J Neurosci 35(42):14183–14194

Vicario DS (1991) Contributions of syringeal muscles to respiration and vocalization in the zebra finch. J Neurobiol 22(1):63–73