PHYSICAL REVIEW E **102**, 062415 (2020)

Editors' Suggestion

Dynamics behind rough sounds in the song of the *Pitangus sulphuratus*

Juan F. Döppler,[1,2] Ana Amador,[1,2] Franz Goller,[3,4] and Gabriel B. Mindlin[1,2]
[1]*Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1428, Argentina*
[2]*IFIBA, CONICET, Buenos Aires 1428, Argentina*
[3]*Institute of Zoophysiology, University of Münster, Münster 48143, Germany*
[4]*School of Biological Sciences, University of Utah, Salt Lake City, Utah 84112, USA*

(Received 18 August 2020; accepted 30 October 2020; published 11 December 2020)

The complex vocalizations found in different bird species emerge from the interplay between morphological specializations and neuromuscular control mechanisms. In this work we study the dynamical mechanisms used by a nonlearner bird from the Americas, the suboscine *Pitangus sulphuratus,* in order to achieve a characteristic timbre of some of its vocalizations. By measuring syringeal muscle activity, air sac pressure, and sound as the bird sings, we are able to show that the birds of this species manage to lock the frequency difference between two sound sources. This provides a precise control of sound amplitude modulations, which gives rise to a distinct timbral property.

DOI: 10.1103/PhysRevE.102.062415

I. INTRODUCTION

Birdsong requires an exquisite coordination between motor patterns [1,2]. Birds typically produce sound with an expiratory airstream, which sets two pairs of soft tissue masses (called labia), in oscillatory motion near the tracheobronchial juncture [3–5]. This results in a periodic modulation of the airflow, which generates soundwaves.

The control of the fundamental frequency of a bird's vocalization is an interesting and rich problem [6–9]. A seminal work by Suthers and Goller [10] described how some oscine birds manage this task. Studying the electromyograms (EMG) of syringeal muscles, air sac pressure, and vocal output in singing brown thrashers (*Toxostoma rufum*), they showed that the activity in musculus syringealis ventralis, the largest syringeal muscle, increases exponentially with the fundamental frequency of the ipsilaterally generated sound. Basically, these muscles change the configuration of the syrinx (the avian vocal organ), such that the labia are stretched. The greater the stretching, the higher the oscillation frequency.

However, not all bird species use the same mechanism for frequency control. A study carried out in the Great Kiskadee (*Pitangus sulphuratus*) [11] showed that during spontaneously generated sound, the modulations of the fundamental frequency are highly correlated with the air sac pressure. Moreover, bilateral resection of the tracheosyringeal nerve did not affect the correlation between fundamental frequency and the air sac pressure. Since this nerve carries the information to the syringeal muscles, its resection effectively eliminates any syringeal muscle control. Interestingly, similar experiments performed in oscine birds have shown a significant alteration of song [1]. The anatomical similarities between the Kiskadee's and the oscine syrinx, might suggest similar function of syringeal muscles. In particular, the large obliquus ventralis muscle in kiskadees has a similar location with respect to the syrinx to that of the oscine syringealis ventralis, which is involved in frequency modulation. This raises a natural question: which role does the large obliquus ventralis muscle (ovm) play in the vocalization of this species?

In this work, we address this question by recording electromyography (EMG) from the ovm during song. We find acoustic features that correlate with muscle activity and provide a model that allows us to explain the link between them. Remarkably, a rich dynamical mechanism allows this species to control a low frequency amplitude modulation in parts of the song, thus enriching its timbre.

II. THE USE OF VOCAL MUSCLES IN PITANGUS SULPHURATUS

In order to address how the activity of the ovm affects the acoustic features of the song, we simultaneously measured EMG activity and sound.

Song was recorded in a custom-built acoustic isolation chamber with a microphone (SGC568; Takstar) connected to an audio amplifier and a data acquisition device (DAQ, USB- 6212; National Instruments) connected to a PC. Custom scripts were used to detect and record sounds, including a 1-s pretrigger window.

To measure the EMG activity, bipolar electrodes were custom-built using ultrafine wire (Ø 25 μm, stainless steel 304, heavy polyimide HML insulated, annealed, California Fine Wires Company, item #CFW0013157). These electrodes were implanted in the ovm muscle as previously described [10] and connected to an analog differential amplifier (225 ×) mounted on a backpack previously fitted to the bird. After a 4-h recovery time, birds were placed in the acoustic isolation chamber, and the amplifier mounted to the backpack was connected to the DAQ through a rotator allowing free movement. The amplifier was powered by external 12-V batteries to avoid

©2020 American Physical Society

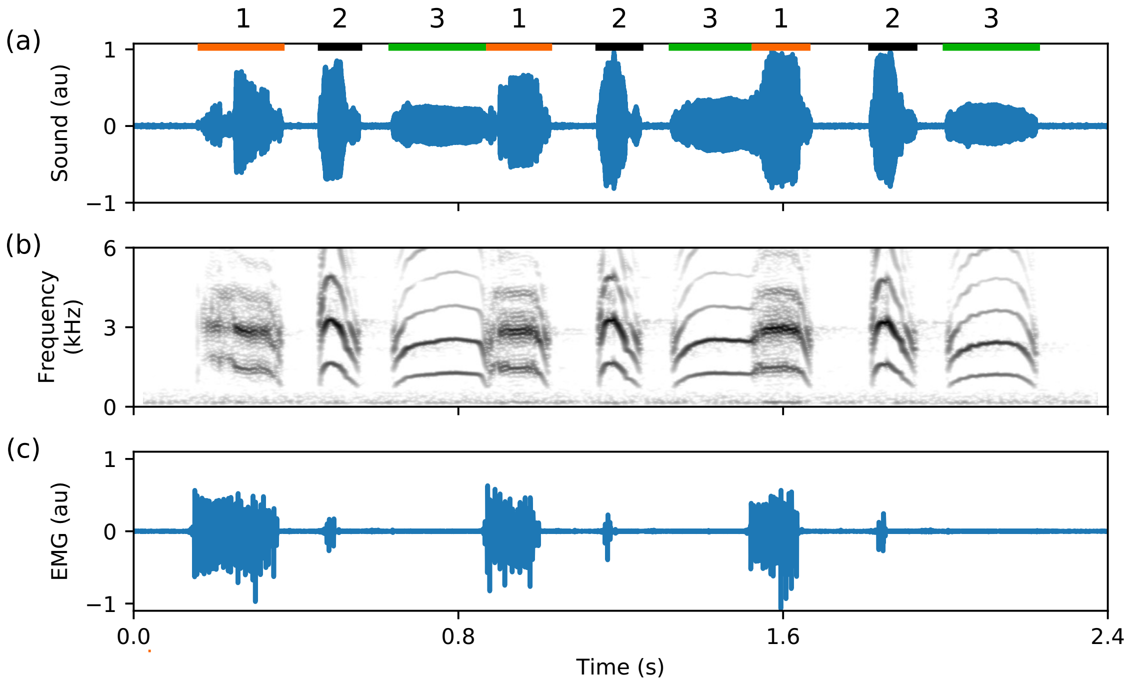

FIG. 1. Simultaneous measurement of sound and EMG activity of the ovm during three renditions of the song. (a) Sound signal. Numbers indicate the syllable number within a song. (b) Spectrogram of the sound. Notice the spectral richness of the first syllable in each repetition of the song. (c) Simultaneous EMG activity. The most prominent activity is produced during the first syllable, while none is present while the third syllable is uttered.

line noise. EMG and audio signals were recorded by a PC running MATLAB, at a sampling rate of 44,150 Hz.

In Fig. 1 we display the sound signal, the spectrogram, and the EMG activity of the ovm during three repetitions of the song, which consists of three syllables. EMG activity is very prominent during the first syllable, and in the second syllable consistently appears during a brief temporal window. In the sonogram we can observe that the first syllable is spectrally richer than the other two, displaying a modulation in its amplitude, which is reflected in its spectral fuzziness. Consistent with the results previously reported [11], the continuous modulations of the fundamental frequency are not reflected in a correlated activation of the ovm. For example, the third syllable presents a clear frequency modulation while there is no EMG activity in the ovm.

In order to unveil the actual role of the ovm we will inspect its activity and corresponding song features in detail. In Fig. 2 we display a detail of Fig. 1, showing the third syllable of a song, followed by the first syllable of the next song. As in Fig. 1, the top panel shows the sound time series, the middle panel the sonogram, and the last panel the EMG. The transition between the two syllables can be characterized by the appearance of a sound amplitude modulation (top panel), which gives rise to a fuzziness of the spectrogram (middle panel). Burstlike activation of the ovm accompanies these song features (bottom panel).

Sound amplitude modulation rate and EMG burst rate appear to be very similar [see Fig. 2(d)], which is confirmed by the quantification of the respective modulation frequencies for syllables from three birds (Fig. 3; $n = 35$). To quantify the frequency of the EMG activity we first computed the Fourier transform of each segment and found the location of its maxima in the range 80–400 Hz. In the case of sound modulation, we first computed its envelope and then proceeded in the same manner. The point size encodes the amplitude of the modulation, computed as the amplitude of the Fourier transform evaluated at the modulation frequency. The computed slope is 1.02 ± 0.07, with $R^2 = 0.87$.

To test the hypothesis that there is a causal relationship between the EMG activity and the modulation of the sound amplitude, we bilaterally denervated the ovm and recorded the songs produced after the surgery. In a previous work [11], we tested that the bilateral denervation did not alter the overall frequency modulation of the syllables. With this new acoustic feature in mind, we explored the vocalizations after the denervation. We found that the modulation did not disappear, and the amplitude of the modulated segment decreased significantly. In Fig. 4(a) we display a representative case, which is the recording after denervation, of the same bird whose song was shown in Fig. 2. While the modulation is still present, the relative amplitude of the modulated segment is reduced compared to song before the denervation. The quantification of the effect is shown in Fig. 4(b). We analyzed the songs from three birds (14 songs from intact birds, and 14 songs after denervation). The amplitude ratio displayed in Fig. 4(b) measures the maximum amplitude of the sound in the first syllable, divided by the maximum amplitude in the third syllable. Notice that in the intact bird, the mean ratio is 1.75, while after denervation it drops below 1.

III. THE TWO-SOUND SOURCE MODEL

The persistence of some level of amplitude modulation in the absence of muscle activity suggests an underlying passive

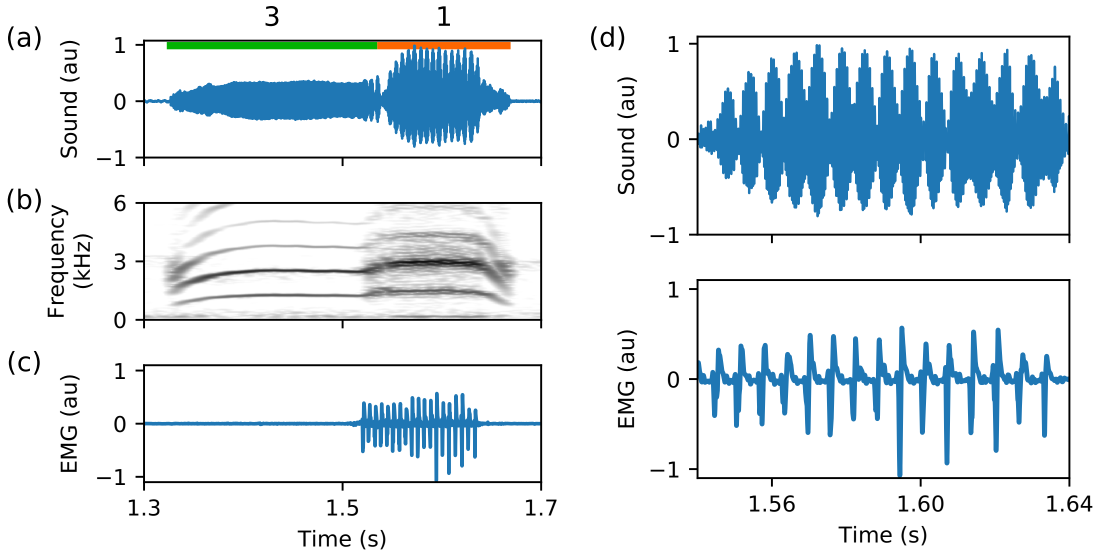

FIG. 2. Comparison of third and first syllables of the song in consecutive renditions. (a) The sound amplitude is modulated during the first syllable of the song. (b) This modulation is observed as a more complex spectrum of the sound. (c) Muscle activity (EMG) consists of a series of peaks at a frequency similar to that of the amplitude modulation and is only present in the first syllable. (d) Detail of the sound and EMG activity during the first syllable.

mechanism whose dynamics is enhanced by the EMG activity in the intact bird. In other words, the role of ovm would be to enhance the effect. To test this hypothesis, we explored a mechanism that (1) could exist in the absence of muscle activity, (2) is present in the first but not the other syllables, and (3) could be enhanced by the muscle activity.

Tyrannid suboscines have two sound sources, each one at the end of the bronchi, right below where they meet to form the trachea [12]. Consistent with previous studies [11,13], we assumed that each of them could be modeled as a nonlinear oscillator. The variables x_1, x_2 stand for the departures from equilibrium of the medial parts of the labia, in each of the two sound sources. The oscillators include nonlinear saturation terms, as well as a cubic nonlinearity in the restitution term describing the labial elastic properties. This allowed us to transduce the slow modulations of the air sac pressure into the slow frequency modulations of the song [14]. Assuming a coupling between the two sound sources (the medial labia are supported by a common bone structure) and a slightly different restitution constant for each of the two labia, we can write the equations for x_1, x_2 as follows:

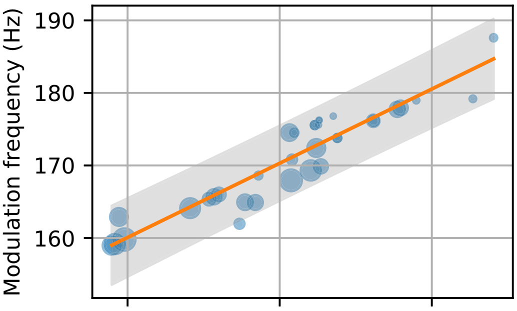

FIG. 3. Sound amplitude modulation frequency is locked to EMG activity frequency. Both frequencies (amplitude modulation and EMG) were calculated as the peaks of their respective Fourier transforms in the range 80–400 Hz. The size of each point is proportional to the amplitude of the modulation, quantified as the amplitude of the Fourier transform at the modulation frequency. The orange line represents the linear fit and the shaded region is the 95% CI of the fit. The slope of the fit (1.02 ± 0.07) shows that both frequencies are not only correlated, but highly similar.

$$\ddot{x}_1 + k\left(1 + \frac{\epsilon\Delta}{2}\right)x_1 + \epsilon\alpha x_1^3 + \epsilon\dot{x}_1\left(cx_1^2 - \mu\right) + \epsilon a(x_1 - x_2) = 0,$$

$$\ddot{x}_2 + k\left(1 - \frac{\epsilon\Delta}{2}\right)x_2 + \epsilon\alpha x_2^3 + \epsilon\dot{x}_2\left(cx_2^2 - \mu\right) + \epsilon a(x_2 - x_1) = 0,$$

where k represents the linear restitution, Δ is the asymmetry between the sources, α is a nonlinear restitution, c is a damping parameter, μ is the air sac pressure, and a is the coupling between sources.

Since we are interested in the amplitude modulation, it is useful to describe each oscillator in terms of an amplitude and a phase [15,16]. In this way, the amplitude modulations in the total sound can be found as the phases of the two oscillators fail to lock. Writing each variable x_i as

$$x_i = r_i \cos(\omega t + \phi_i),$$

assuming $r_1 \sim r_2$ [17], and applying the method of averaging [18], we find for the phase difference $\phi_1 - \phi_2$:

$$\frac{d(\phi_1 - \phi_2)}{dt} = \dot{\chi} = \frac{\epsilon}{2\omega}\left[k\Delta - \frac{3\alpha a}{c}\sin\chi - \frac{a^2}{2\mu}\sin 2\chi\right].$$

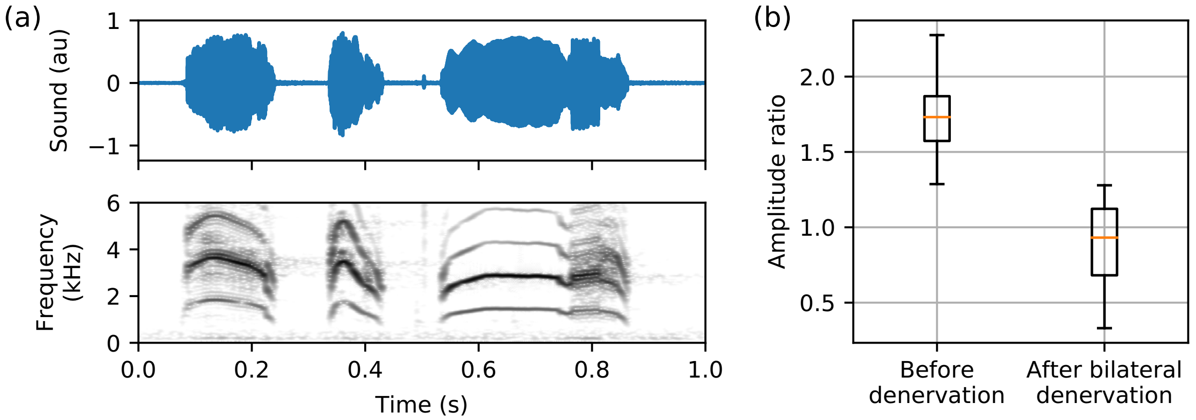

FIG. 4. Amplitude modulation persists after bilateral denervation, but its amplitude is significantly decreased. (a) Sound and spectrogram of a song from a bird after bilateral denervation. Notice how the third syllable is similar in amplitude to the first syllable of the next song. (b) Amplitude ratio of the first and third syllable before and after bilateral denervation. While before denervation this ratio is about 1.75 [e.g., see Fig. 2(a)], it decreases significantly after denervation.

With this equation it is possible to find the conditions under which the sound sources lock [19]:

$$f_1(\chi) \equiv k\Delta - \frac{3\alpha a}{c}\sin\chi = \frac{a^2}{2\mu}\sin 2\chi \equiv f_2(\chi),$$

since when $f_1(\chi) = f_2(\chi)$, $\dot{\chi} = 0$. In Fig. 5 we display in different colors the functions $f_1(\chi)$, $f_2(\chi)$. In Fig. 5(a) we show that for high values of μ (the parameter representing the air sac pressure), the two oscillators fail to lock. Decreasing μ [Fig. 5(b)] fixed points for χ exist, representing locked oscillations. In other words, with this model, high values of air sac pressure lead to the two sound sources running at different frequencies, while low air sac pressure values allow the two sound sources to oscillate at a common frequency.

In Fig. 6 we display synthetic sound segments that are generated by the two sound sources, for the two dynamical regimes described above. The sound is computed as $s(t) = \dot{x}_1 + \dot{x}_2$ [20] (we do not consider in our simulations the filtering effect of the trachea and the oropharyngeal-esophageal cavity). In the first row of Fig. 6 we display the case where the two sound sources drift, while the second row shows the locked case. The panels in the left display the functions $f_1(\chi)$, $f_2(\chi)$, the middle ones show the displacements of the two sound sources (x_1, x_2), while the panels at the right show the synthetic sound. In the first row we describe the case of high air sac pressure (μ), that leads to the sound with modulated amplitude. The second row corresponds to the case with the locked sound sources, for which no modulations occur.

Consistent with these arguments, the air sac pressure used during the production of the first syllable of the song should be higher than the one used during the third syllable, which does not present amplitude modulations. Moreover, this relationship should be valid after denervation, which is a condition in which an amplitude modulation should emerge only from the way in which the sound sources interact, regardless of the effect of the driving provided by the muscles.

Figure 7(a) shows a typical air sac pressure time series data, used to produce the three syllables of the song. The air sac pressure is measured by the insertion of a flexible cannula

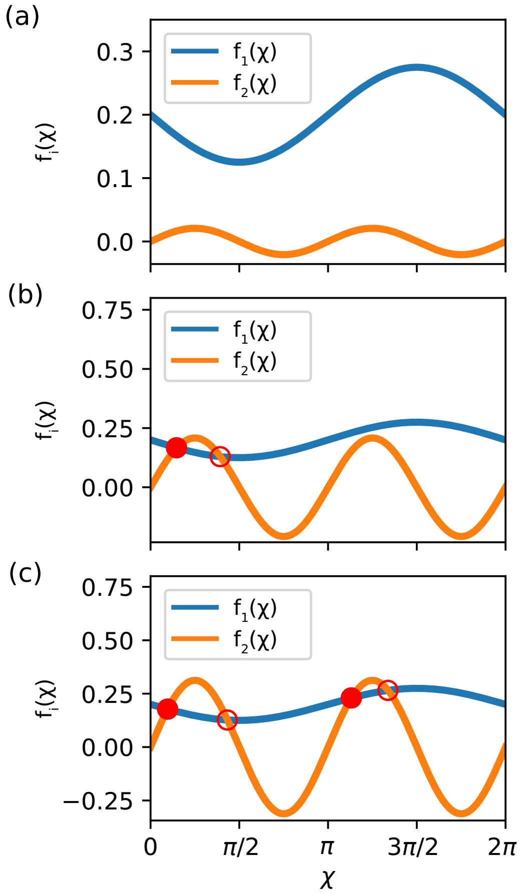

FIG. 5. Plots of $f_1(\chi)$ (blue line) and $f_2(\chi)$ (orange line) for three representative sets of parameters. (a) In this case, a high value of air sac pressure (represented by μ) the system has no fixed point and the sources do not lock. (b) Fixed points for χ (filled dot for stable, empty circle for unstable in the figure) exist when the pressure is decreased. These represent solutions in which the sources are locked (their phase difference remains constant). (c) When pressure is further decreased another pair of fixed points appears. In this regime two stable solutions coexist.

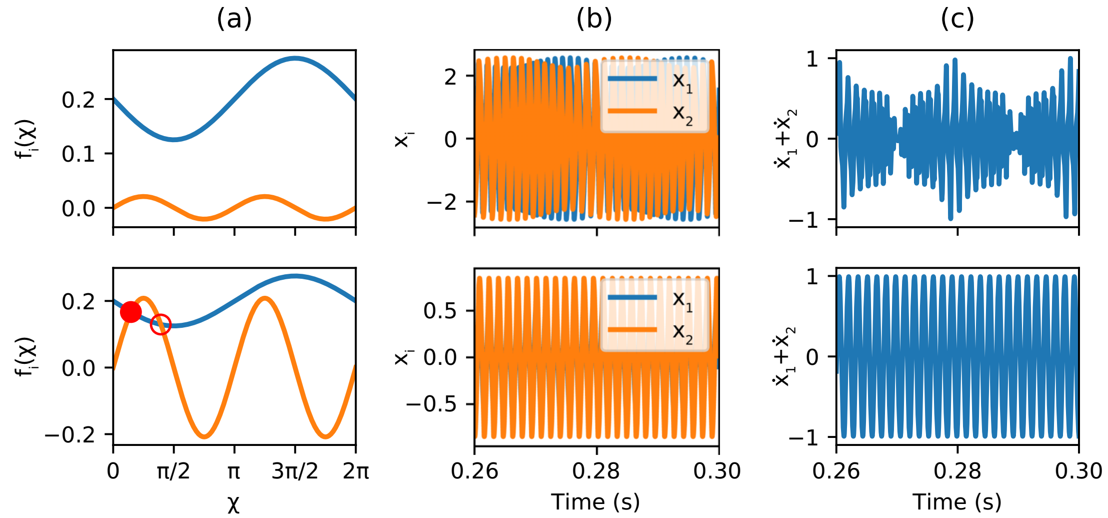

FIG. 6. Numerical integrations of the full model for two representative sets of parameters displaying phase drift (top) and locked sources (bottom). (a) $f_1(\chi)$ (blue line) and $f_2(\chi)$ (orange line). (b) Results of the numerical integration of the model: $x_1(t)$ (blue), and $x_2(t)$ (orange). (c) Sound, computed as $\dot{x}_1(t)+\dot{x}_2(t)$. In the first panel, with parameters in which the sources do not lock, their phase drift causes a slow amplitude modulation in the sound. In the second case, in which the sources are locked, the amplitude of the sound remains constant.

through the abdominal wall just posterior to the last rib, so that it extended a few millimeters into a thoracic air sac. The free end of the cannula was connected to a miniature piezoresistive pressure transducer (Fujikura model FPM-02PG), which was mounted on the bird's back, whose output was later recorded by a PC. In Fig. 7(b) we display the ratio between the maximum of the air sac pressure for the first syllable and the maximum pressure of the third syllable, before and after denervation. We analyzed 28 songs from three birds (six songs uttered by intact birds, and 10 after denervation). Notice that the first syllable always uses a larger pressure, even for vocalizations produced after denervation. This is consistent with the first syllable in the song presenting amplitude modulations (with the two sound sources failing to lock, as in the first row in Fig. 6), and the last syllable being generated by two locked sound sources (second row of Fig. 6).

The analysis described above shows how to generate a sound with amplitude modulations before the action of the syringeal muscles is considered (i.e., for the denervated birds). In order to study the sound alteration that follows the presence of activity in the syringeal muscle, we need to estimate how an EMG activity pattern modifies the parameters in our model. In particular, we expect the EMG to change the configuration of the syrinx, stretching the labia, and therefore increasing the restitution constant. In a previous work [21] it was shown that a sequence of pulses generated at rate r is translated into a modulation of k (at rate r) mounted on a DC level.

In order to explore how this modulation is translated into acoustic features, we introduced $k(t)=k_m+k_A\cos(\omega_f t)$ into the equations for $x_1,\ x_2$, and inspected their evolution for different values of ω_f. Figure 8 shows the most noticeable effect of the forcing: the amplitude of the modulations increases.

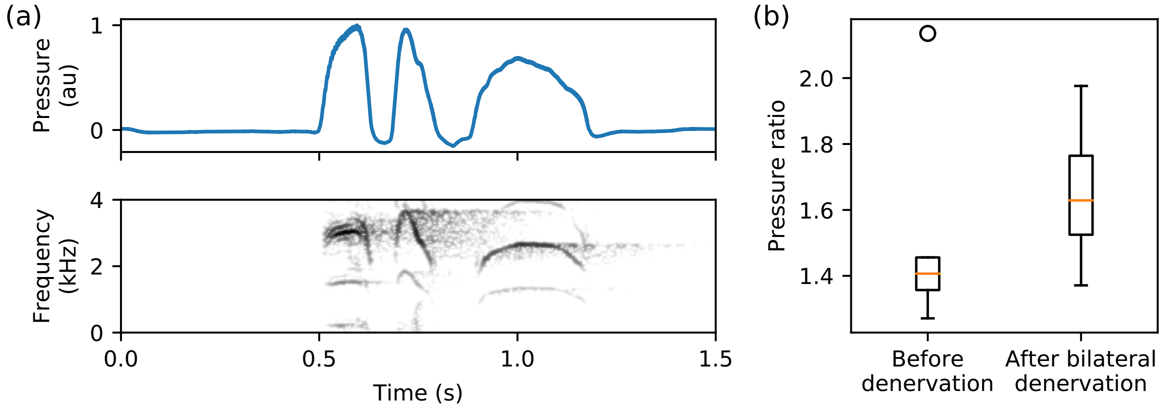

FIG. 7. Air sac pressure is greater for the first syllable. (a) A simultaneous measurement of air sac pressure (top) and sound (bottom, represented as a spectrogram). Pressure on the first syllable (in which amplitude modulations occur) is significantly greater than in the third syllable (in which no EMG activity is observed). (b) Pressure ratio between the first and third syllable before and after denervation. In all cases it is significantly greater than 1, indicating that the pressure in the first syllable is greater than that of the third one.

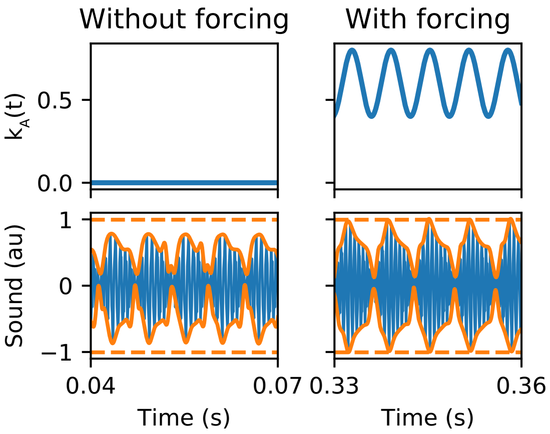

FIG. 8. The proposed model reproduces the increase of sound amplitude modulation when forced. Top: $k_A(t)$ for unforced (left) and forced (right) cases. Bottom: sound (blue) obtained from integration of the model for each case. The full orange line represents the envelope of the sound. Orange dashed lines in the bottom panel indicate the maximum and minimum amplitude for the forced case.

The panels at the top indicate the parameters $k_A(t)$, while the bottom panels display the synthetic sound obtained from integrating the model. Notice that we prepared our model to present amplitude oscillations for sufficiently large air sac pressure, by assigning slightly different natural frequencies to the sound sources.

Interpretation: Amplitude increase, and locking of the modulation

Yet, our simulations indicate that for a forcing frequency close to the modulation frequency, the later locks to the forcing. In Fig. 9 we show the frequency of the amplitude modulations that appear in a syllable generated with pressure high enough so that even without external forcing, the two frequencies do not lock, giving rise to amplitude modulations. As the time dependent forcing of k is introduced, the amplitude modulation locks to the forcing frequency f_f. We observe this effect by plotting the amplitude modulation frequency for different values of the forcing frequency [22]. When the forcing frequency is similar to the amplitude modulation frequency of the unforced system, the resulting sound displays amplitude modulations at the forcing frequency.

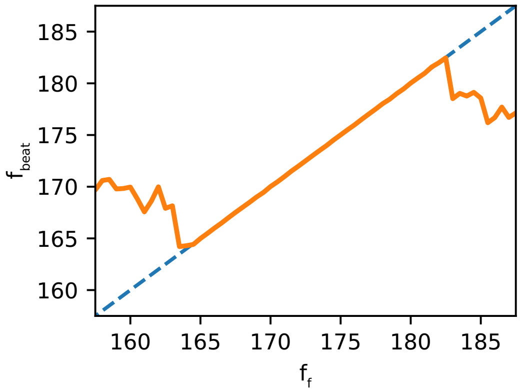

FIG. 9. Sound amplitude modulation frequency versus forcing frequency for the model (orange line). The blue dashed line represents the identity. Notice that for a range of frequencies, the modulation frequency locks to the forcing frequency.

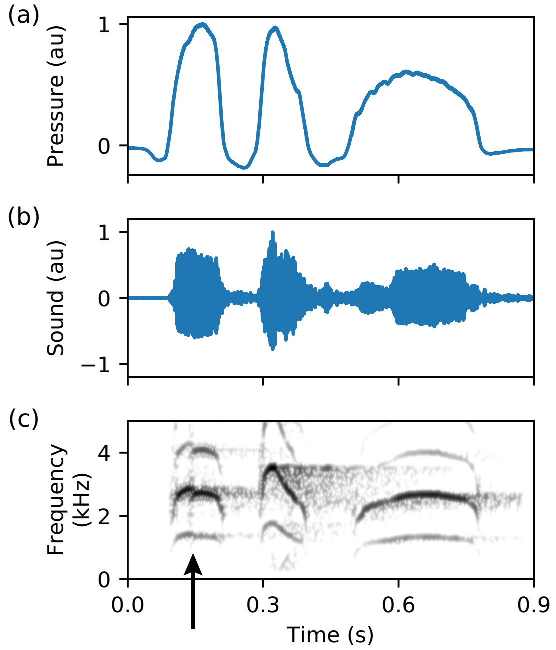

FIG. 10. A frequency jump observed in a nonmanipulated kiskadee: (a) air sac pressure, (b) recorded sound, and (c) spectrogram. The black arrow indicates the occurrence of the frequency jump.

It is interesting that the sound source interaction that we propose is also capable of accounting for other acoustic features observed in *kiskadee* song, such as frequency jumps. These are discontinuities in the fundamental frequency of the song, which occur frequently in the first syllable of the song and in some calls (see Fig. 10). We conjecture that the coexistence of solutions found in the model for low enough pressure [see Fig. 5(c)], might lead to sudden transitions as the pressure is smoothly increased. This would be the result of one of the stable stationary solutions disappearing in a saddle node bifurcation as the air sac pressure is increased (see Fig. 5), while another stable solution remains. If the system is initially at the vanishing solution (where the oscillators are in antiphase), it would evolve to the remaining attracting state, in which the oscillators are in-phase. In this way, a mechanism proposed to account for some acoustic features explains other observations as well.

IV. CONCLUSIONS

Much of the work on birdsong focuses on birds that learn their vocalization, i.e., on those approximately 4000 species that require some level of exposure to a tutor in order to acquire the characteristic songs of the species. In the last years, evidence started to build supporting the thesis that nonlearners (as it is the case for the species analyzed in this work) find their niche in the acoustic landscape by means of a wide variety of anatomical adaptations [23,24].

In this work we presented experimental data and a theoretical analysis that builds confidence on a model on how *Pitangus sulphuratus* (a nonlearning tyrannid from the Americas) achieves the complex timbral features characteristic of some of its vocalizations. We found that during the production of the first syllable, and a brief time interval during the production of the second syllable, the ovm displays activity bursts at a low frequency. Locked to these bursts, the sound amplitude is modulated. This mechanism therefore identifies a functional role of the syringeal musculature during song production in this tyrannid and thus explains the puzzling results from denervation studies [11]. The combination of experimental data and modeling allows us to conclude that the activity of the ovm enhances the amplitude of a low frequency modulation of the sound.

We presented a model capable of capturing several properties of the song: the survival of low amplitude oscillations in the first syllable after denervation, the locking between the fluctuations in the EMG and the amplitude fluctuations, and the occurrence of spontaneous frequency jumps. We found the mechanism by which that modulation is achieved: the forcing locks to a frequency drift between the two sound sources. The computational model used to test these hypotheses is consistent with a previously proposed model for kiskadee song, which reproduced the correlation between the modulations of the fundamental frequency with the air sac pressure [11].

The present work illustrates that nonlearners do not only use morphological adaptations to enrich their vocalizations: interesting dynamical effects in conjunction with neural control patterns contribute to this enrichment as well.

ACKNOWLEDGMENTS

The authors acknowledge support from Universidad de Buenos Aires and Agencia Nacional de Promoción de la Investigación el Desarrollo Tecnológico y la Innovación (Argentina). The experiments were conducted following the protocols of the CICUAL (committee for care and use of animals in the laboratory) of the School of Sciences of the Universidad de Buenos Aires, c.n.113, under permits from the state of Buenos Aires (DI-2018-GCDEBA-DFYMAGP). The birds in this study were released after recovery, following the protocols described in the last permit.

---

[1] D. S. Vicario, J. Neurobiol. **22**, 63 (1991).
[2] S. Nowicki, M. Westneat, and W. Hoese, Semin. Neurosci. **4**, 385 (1992).
[3] F. Goller and O. N. Larsen, J. Exp. Biol. **200**, 2165 (1997).
[4] R. A. Suthers, F. Goller, and C. Pytte, Philos. Trans. R. Soc. London Ser. B. **354**, 927 (1999).
[5] R. A. Suthers, in *Nature's Music: The Science of Birdsong*, edited by P. Marler and H. W. Slabbekoorn (Elsevier, New York, 2004), pp. 272–295.
[6] G. J. L. Beckers, J. Exp. Biol. **206**, 1833 (2003).
[7] C. P. H. Elemans, R. Zaccarelli, and H. Herzel, J. R. Soc. Interface **5**, 691 (2008).
[8] A. Amador and D. Margoliash, J. Neurosci. **33**, 11136 (2013).
[9] F. Goller and T. Riede, J. Physiol. Paris **107**, 230 (2013).
[10] F. Goller and R. A. Suthers, J. Neurophysiol. **76**, 287 (1996).
[11] A. Amador, F. Goller, and G. B. Mindlin, J. Neurophysiol. **99**, 2383 (2008).
[12] P. L. Ames, *The Morphology of the Syrinx in Passerine Birds* (Peabody Museum of Natural History, Yale University, New Haven, CT, 1971).
[13] R. G. Alonso, C. Kopuchian, A. Amador, M. de los, A. Suarez, P. L. Tubaro, and G. B. Mindlin, J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol. **202**, 361 (2016).
[14] R. Alonso, F. Goller, and G. B. Mindlin, Phys. Rev. E **89**, 032706 (2014).
[15] T. Chakraborty and R. H. Rand, Int. J. Nonlinear. Mech. **23**, 369 (1988).
[16] A. P. Kuznetsov, N. V. Stankevich, and L. V. Turukina, Physica D **238**, 1203 (2009).
[17] G. Dangelmayr and E. Knobloch, Nonlinearity **4**, 399 (1991).
[18] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, New York, 1983).
[19] V. Anishchenko, S. Nikolaev, and J. Kurths, Chaos **18**, 037123 (2008).
[20] G. B. Mindlin, Chaos **27**, 092101 (2017).
[21] J. F. Döppler, A. Bush, F. Goller, and G. B. Mindlin, J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol. **204**, 209 (2018).
[22] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2003).
[23] S. M. Garcia, C. Kopuchian, G. B. Mindlin, M. J. Fuxjager, P. L. Tubaro, and F. Goller, Curr. Biol. **27**, 2677 (2017).
[24] G. Uribarri, M. J. Rodríguez-Cajarville, P. L. Tubaro, F. Goller, and G. B. Mindlin, Phys. Rev. Lett. **124**, 098101 (2020).