Nonlinear dynamics in the study of birdsong

Cite as: Chaos **27**, 092101 (2017); https://doi.org/10.1063/1.4986932
Submitted: 07 June 2017 • Accepted: 17 August 2017 • Published Online: 19 September 2017

Gabriel B. Mindlin

Nonlinear dynamics in the study of birdsong

Gabriel B. Mindlin
Departamento de Física, FCEyN, Universidad de Buenos Aires IFIBA, CONICET, Argentina

(Received 7 June 2017; accepted 17 August 2017; published online 19 September 2017)

Birdsong, a rich and complex behavior, is a stellar model to understand a variety of biological problems, from motor control to learning. It also enables us to study how behavior emerges when a nervous system, a biomechanical device and the environment interact. In this review, I will show that many questions in the field can benefit from the approach of nonlinear dynamics, and how birdsong can inspire new directions for research in dynamics. *Published by AIP Publishing.*
[http://dx.doi.org/10.1063/1.4986932]

Biologically inspired problems pose deep challenges to a dynamicist. A biological problem needs to be framed within the theory of evolution, with the profound complexity it entails. Therefore, it is difficult to travel the road of dimensionality reduction, or identify the *basic mechanisms* behind the phenomenon under consideration, both of which are core strategies followed by dynamicists to study a problem. In this review, I will describe a dynamical approach to one specific biological problem: birdsong production.

I. INTRODUCTION

Birdsong is an interesting model in neuroethology, i.e., the neurobiology of behavior.[1] Among the several reasons that contribute to its appeal are its complexity, its stereotypy and why not, its beauty. For neuroscience, a particular interest emerges from the observation that some degree of learning is involved in song production in approximately forty percent of the known bird species.[2] Since learned vocal production occurs rarely in the animal kingdom, songbirds (which account for the majority of forty percent of birds that learn their vocalizations) constitute a favorite animal model to study the neurobiology of vocal learning.

The approach from neuroethology stresses how this complex behavior emerges from the interaction between the nervous system, the body and the environment.[3] And, it is precisely at the interaction between the nervous system and the biomechanics that extremely interesting and pertinent questions for dynamicists naturally emerge. For example, how much of the acoustical complexity is due to the complexity of the instructions that the nervous system sends to the periphery, and how much is due to the nonlinear nature of the vocal organ? Can the avian vocal organ respond in complex ways to relatively simple physiological instructions? How many of the acoustic features are independently controlled at the level of the nervous system? Do some features arise together due to the nonlinear nature of the oscillations responsible for the sound production mechanism? Unveiling these issues can provide neuroscientists with a map of what is worth studying at the level of the nervous system, and which features are conditioned by the dynamics of the periphery.

It is also interesting to address the complexity of the physiological instructions driving the avian vocal organ. Its operation requires the delicate interplay between the physiological instructions that control its configuration,[4] and the airflow necessary for song production, controlled by the respiratory system.[5] Remarkably, those physiological instructions can themselves be understood as the solution of reasonably low-dimensional dynamical systems. Could these provide a link between the known anatomy of the parts of the nervous system involved in birdsong production and their macroscopic functionality? This observation can be framed in an interesting debate: do cortical neural patterns represent movement parameters, or do they constitute a dynamical system that generates and controls motion?[6]

In this review, I will address both issues: how much of the complexity found in birdsong is conditioned by the nonlinear nature of the vocal organ, and what we can learn about the parts of the nervous system involved in song production from the way the vocal organ is controlled. Dynamics will provide a common language for our study.

II. THE ANATOMY OF THE AVIAN VOCAL ORGAN

The avian vocal organ is the syrinx. In songbirds, it is a bipartite structure at the junction between the bronchi and the trachea.[7,8] At each junction, there is a pair of labia which can be set into oscillatory motion when strong enough airflow is established between them. In this respect, each of the two sides of this bipartite structure behave somewhat similarly to the human vocal folds, which can be set in oscillatory motion when voiced sounds are uttered. Labial oscillations modulate the airflow at each of the syringeal sides and, therefore, there is a periodic injection of air into the trachea (assumed to be a tube of volume V_0). The rate of mass injection for a unit of volume q can be written in terms of air velocity v, density ρ and lumen's area A as

$$q = \rho A v / V_0.$$

The density perturbations induced by this mass injection, at the base of the trachea, are ruled by

$$\frac{\partial^2 \rho}{\partial t^2} - {c_0}^2 \nabla^2 \rho = \frac{\partial q}{\partial t},$$

with c_0 being the sound velocity.[9] Therefore, labial dynamics are responsible for the temporal evolution of the lumen's area A and act as a sound source in the linear approximation of the acoustic problem. The lumen's area will be the product of a transverse, constant length and a variable, whose (nonlinear) dynamics describe the labial motion. The nonlinear nature of labial dynamics is responsible for the harmonically rich content of the acoustic sound, which gets filtered by different passive cavities as the sound waves find their way from the sound source to the beak. In Fig. 1(a), we display the basic elements of this description. The air passing through the syrinx is injected into the trachea, which connects to the oroesopharingeal cavity, which opens towards the exterior through the beak.

The syringeal configuration can be modified by the action of muscles attached to it.[4,10] In Fig. 1(b), we show two ventral muscles (syringealis ventralis, vS, and syringealis tracheobronchialis, vTB) and two dorsal ones (tracheobronchialis dorsalis, dTB, and syringealis dorsalis, dS). The muscle vS attaches at the second cartilaginous bronchial ring. Since the labia are housed in the inner part of the bronchi, between the second and third rings, the contraction of vS stretches the labia, affecting the frequency at which they oscillate under a given airflow. For this reason, the muscle vS is involved in the modulation of the fundamental frequency of a song. The muscle vTB controls the active separation of the tissues that oscillate during phonation, while the dorsal muscles dTB and dS control the active closing of the lumen.

In light of this, there are two clearly distinct timescales in the problem. The syllabic timescale involves the modulations of the airflow below and above a phonating threshold, as well as the modulations of the acoustic features achieved by the activation of the muscles that alter the configuration of the syrinx. These modulations are therefore actively controlled and present periods from 50 ms to 400 ms, depending on the species and the syllable type. The labial timescale is related to the self-induced oscillations taking place when energy from the airflow is transferred to the labia (typically, in the order of kHz). Due to the separation between these two timescales, one can analyze the generation of birdsong by first studying the labial dynamics under stationary parameters (i.e., the bifurcation diagram for the dynamical system ruling the behavior of the labia) and then explore how sound is affected as the parameters are slowly modulated.[11]

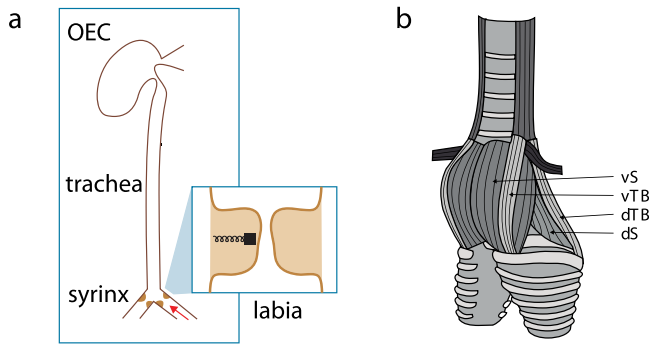

FIG. 1. The syrinx, the trachea and the oroesopharingeal cavity (a). Muscles controlling the oscine syrinx. See text for full description (b).

III. THE FUNCTIONALITY. THE EQUATIONS OF A SIMPLE MODEL

Let us start with a first attempt at modeling the labial dynamics, whose oscillations are responsible for the airflow modulations. For the moment, we will analyze the behavior of one sound source, and we will assume that its two labia (medial and lateral) are synchronized. We are postponing the discussion about the circumstances under which this is a realistic hypothesis for Sec. VII.

The variable in our first model will be x, describing the departure of the labial midpoint position from rest. Its dynamics will be ruled by Newton equations, where the forces are (i) the elastic restitution, depending on the labial's departure from its rest position, (ii) the linear dissipation, which is proportional to the labial velocity and includes both a negative contribution accounting for the transfer of energy from the airflow to the mass as a mucosal wave propagates along the labium and a positive contribution due to loss, (iii) the nonlinear dissipation, which is responsible for bounding the oscillations, and represents either the labia collapsing against the containing walls or against each other, and (iv) forces representing active adduction (pulling the labia together) and abduction (labial separation). These last forces do not depend on x, or its time derivative. This model, then reads[12]

$$\frac{dx}{dt} = y$$

$$\frac{dy}{dt} = -k(t)\gamma^2(x+\epsilon x^3) - \big(\beta(t)-\beta_0\big)\gamma y - c\gamma x^2 y + \gamma^2(f_{add}(t) - f_{abd}(t)),$$

where k(t) describes the restitution (proportional to the tension of the labia), $\beta(t) - \beta_0$ is the negative dissipation (proportional to the air sac pressure responsible for establishing the airflow through the lumen), and γ is the problem's timescale. There are two sources of nonlinearities in this first model: at the restitution and at the losses. As soon as the labia depart from equilibrium, they would collide with each other or against the containing walls, dramatically losing their energy. This is modeled through a nonlinear dissipation term (i.e., a term which is important when the variable is away from its equilibrium position).

As soon as we introduce nonlinearities, the modulation of the fundamental frequency will involve both β and k. Let us show that this is the case by analyzing the bifurcation diagram of this model (operating at $f_{add} = f_{abd} = 0$).

Scaling the time through $t \equiv \Gamma\tau$, $\Gamma = \frac{1}{\gamma\sqrt{k}}$, defining

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} z \\ z^* \end{pmatrix},$$

and keeping the resonant terms, we get

$$\frac{dz}{d\tau} = \left(\frac{\beta}{\sqrt{k}} + i\right) z + \left(\frac{3}{2}\epsilon i - \frac{1}{2\sqrt{k}}\right) z|z|^2.$$

Analyzing the radial part of this complex equation and its phase, we get that, at the Hopf bifurcation, the frequency of the oscillation being born is:

$$\omega = \gamma\sqrt{k}(1 + 3\epsilon\beta).$$

This means that the isofrequency curves (i.e., curves in parameter space (β, k), defined as $k = k(\beta)$, for which solutions with a given fundamental frequency exist) meet the Hopf bifurcation curve with negative slopes. For higher values of the parameter β, it is possible to show that the isofrequency curves present positive slopes (since $\omega \approx k/\beta$). In Fig. 2, we show a numerically obtained diagram of isofrequency curves for this simple model, for the region of the (β, k) parameter space with oscillations. With the help of this diagram, it is easy to infer that for a given value of β, we need to increase the parameter k in order to obtain oscillations of higher fundamental frequency . However, modulating the frequency can be trickier. In a region of parameter space where the isofrequencies have negative slopes, it is possible to keep the labial tension constant and increase the fundamental frequency by increasing β. It could even be possible to reduce the tension slightly, while the pressure is being increased, and achieve an increase of the fundamental frequency, as long as the slope of the curve in the parameter space is smaller than the slope of the isofrequency. Modulo these subtleties, one can use this static bifurcation diagram to obtain good insight into how to translate physiological gestures into acoustic features. The parameter β will have to be moved from the region where no oscillations exist, through the oscillating region in order to start the phonation. In the course of this trajectory, the parameter k can be moved, so that the fundamental frequency is properly modulated. The syllable will end as soon as the parameter β is returned to the non-oscillating region of the parameter space. In Fig. 2, we display a schematic path in parameter space that would lead to an upsweep syllable, i.e., the frequency would increase its value monotonically during the phonation.

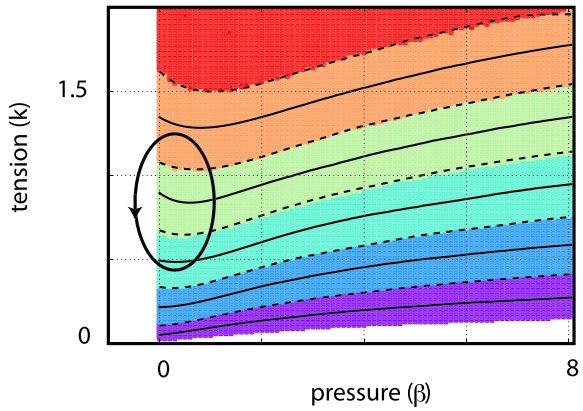

FIG. 2. Isofrequencies in the parameter space of a simple model for labial dynamics. The slowly changing parameters during the production of a syllable are the coordinates of a trajectory in the parameter space. Pressure and tension are (β, k) in the model.

IV. EXPERIMENTAL EVIDENCE

Testing this paradigm required a long experimental program. The identification of the oscillating labia, as the sound source required nothing less than direct visualization of the labia during phonation. Goller and Larsen carried out this experiment, which ended a long debate on the origin of the sound. It was more complicated to build confidence on the role played by the physiological parameters involved in birdsong production and control, namely, the air sac pressure and the activity of muscles controlling the configuration of the syrinx. Let us review this research program.

The sub-glottal pressurization is a necessary condition for establishing an airflow strong enough to start the self-sustained labial oscillations responsible for sound production. Birds have rigid lungs, and the air is passed through them as air sacs (connected to the lungs) are compressed in a coordinated fashion. It is relatively non-disruptive, then, to measure the level of sub-glottal pressurization by inserting a flexible cannula into one of these air sacs, with the cannula's free end connected to a transducer. In Fig. 3, we show three syllables of a canary song.[13] In the first panel, we display the sound, as recorded by a microphone. In the second panel, the air sac pressure is measured using a cannula inserted through the abdominal sac with its free end connected to a miniature piezoresistive pressure transducer mounted on the bird's back. The third panel in Fig. 3 corresponds to the electromyographic EMG activity pattern as measured by a very thin wire implanted on the syringealis ventralis muscle (vS). Details on both procedures can be obtained from the work of

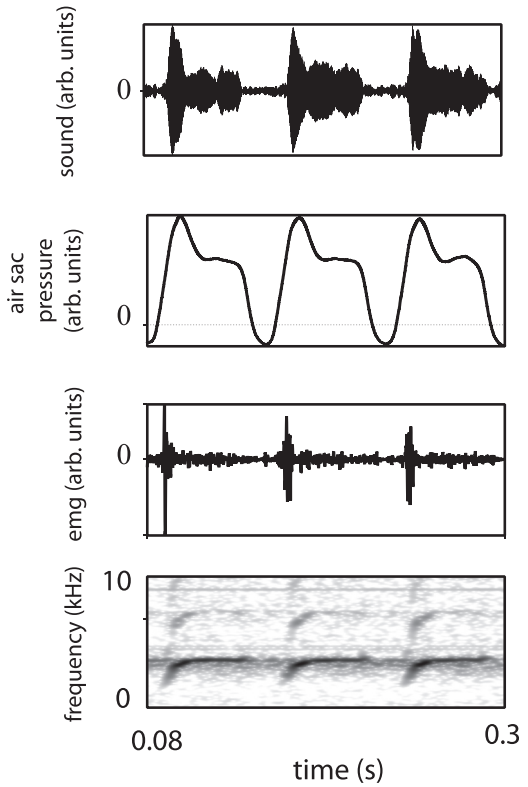

FIG. 3. Experimental recordings during the production of a canary syllable. The sound (first panel), the air sac pressure (second panel), the electromyography of the right muscle syringealis ventralis (third panel) and the sonogram (fourth panel).

Suthers and Goller, who developed these techniques originally.[4,10]

The first test of the simple dynamical model discussed before precisely consisted of integrating the dynamical equations, using as (slowly) time dependent parameters, the envelopes of the EMGs recorded as a bird song.[14] The rationale behind the procedure was that the contraction of the muscle would stretch the labia, and that stretched labia would present a higher restitution (k). Remarkably, the synthetic sound generated by the model driven in this way shared acoustical features with the song produced while the driving EMG was recorded. In particular, it closely followed the modulation of its fundamental frequency.

V. A MODEL FROM FIRST PRINCIPLES AND ITS DYNAMICS

In the previous model, the parameter β was used to turn on the labial oscillations, by parameterizing the energy transfer to the labia. A model based on first principles would allow us to find the actual link between this parameter and the air sac pressure.

Let us assume that for sufficiently high values of airflow, the labia start to oscillate with a wavelike upward motion.[15] In order to describe this wave, we introduce two basic modes: a lateral displacement of the labia and a flapping motion, as displayed in Fig. 4. This leads to an out-of-phase oscillation of the top and bottom portions of the labia. As before, a variable x will describe the medial position of a labium and allows us to write a kinematic description of the modal motion. This will be necessary to compute the actual force between the labia. If a_1, a_2 stand for the half separation between the lower and upper edges of the labia, we can write them as:

$$a_1 = a_{10} + x + \tau \frac{dx}{dt},$$

$$a_2 = a_{20} + x - \tau \frac{dx}{dt},$$

where a_{10}, a_{20} are the half separations in the resting state and τ is the time the wave propagates along the labium takes to traverse half the labial vertical size. In this geometry, the average pressure between the labia for a given value of subglottal pressure p_{sub} is:

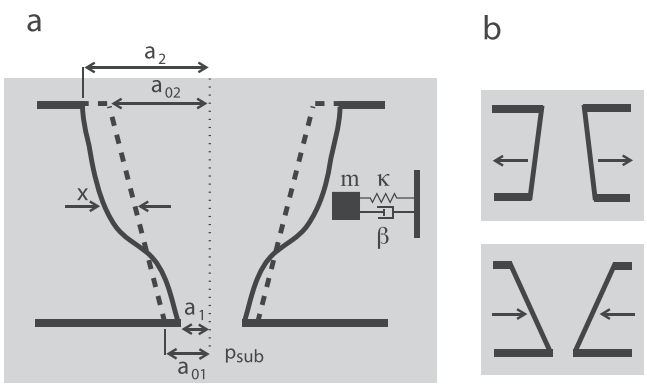

FIG. 4. The modes of labia during one cycle.

$$p_{average} = p_{sub}\left(1 - \frac{a_2}{a_1}\right),$$

allowing us to write Newton's equations for the labia as:

$$\frac{dx}{dt} = y,$$

$$m\frac{dy}{dt} = -kx - by - cx^2 y + a_{lab} p_{sub}\left(1 - \frac{a_2}{a_1}\right) + f_0,$$

which, expressing the half separations as functions of the labial position, leads to

$$\frac{dx}{dt} = y,$$

$$\frac{dy}{dt} = \frac{1}{m}\left(-k(x)x - \beta(y)y - cx^2 y + f_0 + a_{lab} p_{sub}\left(\frac{a_{10} - a_{20} + 2y\tau}{a_{01} + x + \tau y}\right)\right).$$

In these equations, a_{lab} stands for the lateral labial area and f_0 for the difference between the externally controlled adducting and abducting forces. Now, the physical mechanism needed in order to obtain self-sustained oscillations is clearer. The described kinematics prescribes that the labia are moving away from each other when they present a convergent profile, and moving towards each other when they present a divergent one. Furthermore, a convergent profile means that the average pressure between the labia is similar to the air sac pressure (and larger than the atmospheric pressure), while a divergent one makes the pressure between the labia closer in value to the atmospheric pressure. Therefore, when the labia move away from each other, there is a net force in the direction of the velocity that is larger than when the labia move towards each other. This leads to energy transfer from the air flow to the labia.

Figure 5 shows a bifurcation diagram for this dynamical system.[16] For small values of pressure, as expected, the labia do not oscillate. High values of pressure lead to oscillations. Interesting enough, the model presents a cusp bifurcation, and the line in the parameter space where a Hopf bifurcation takes place is tangent to one of the saddle nodes of the cusp, at a Takens-Bogdanov co-dimension two bifurcation. Interestingly enough, this organization of the parameter space leads to the existence of a saddle node in limit cycle (SNILC) bifurcation (involving part of the second branch of the saddle node bifurcation line), where the oscillations are born with zero frequency and a finite amplitude. In this way, labial oscillations can be born at Hopf bifurcations for high values of tension, or in the saddle node in limit cycles for smaller values. The last two panels in Fig. 5 show the sonograms of a recorded song and its synthetic replica using this model.

The spectral content of a sound signal is not only determined by the nature of oscillations of the labia modulating the airflow. The sound source $s(t) = \frac{\partial q}{\partial t}$ plus the backpropagating wave contribute to the pressure fluctuations $p_i(t)$ at the input of the trachea of length L

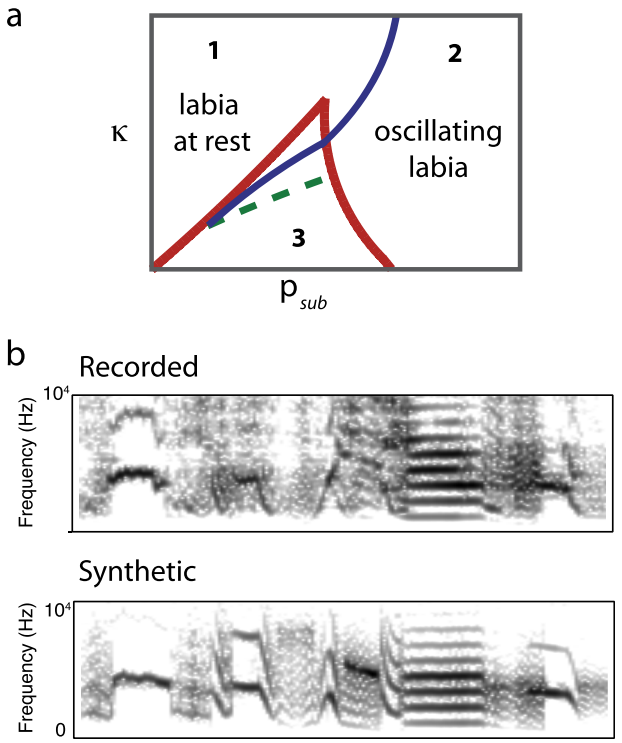

FIG. 5. Parameter space for the model described in the text. The red curves correspond to saddle node bifurcations and the blue curve to a Hopf (a). The sonogram of a recorded song (top panel) and the sonogram of a synthetic sound produced by the model (bottom panel) (b).

$$p_i(t) = s(t) + p_{back}\left(t - \frac{L}{c}\right),$$

$$p_{back}(t) = -\gamma p_i\left(t - \frac{L}{c}\right),$$

and the transmitted pressure wave still has to excite the oroesopharingeal cavity (modeled as a Helmholtz resonator) before constituting a reasonable approximation to birdsong. But, it is clear that the source of spectral content richness lies at the sound producing mechanism, and the fact that a saddle node in limit cycle oscillation exists in the problem predicts interesting features. For the range of parameters where this bifurcation can be expected, low frequency sounds should be spectrally rich, and there should be a precise relationship between the spectral content of the signal and its fundamental frequency.

One can proceed to measure the spectral richness of a sound segment in this way. Given a sound segment, it is possible to compute its fast Fourier transform (FFT) and reconstruct two parameters from it.[17] The first one is f_{aff}, the average fundamental frequency (AFF), and the second one, the mean spectral frequency (MSF), is defined as:

$$f_{MSF} = \sum_i \omega_i \epsilon_i / E,$$

where ω_i is each frequency component of the spectrum, ϵ_i is its energy, and E stands for the total energy in the spectrum. With these parameters, we compute the spectral content index (SCI) as $SCI \equiv f_{MSF}/f_{aff}$. This definition allows us to compare spectrally different syllables. In Fig. 6, we display

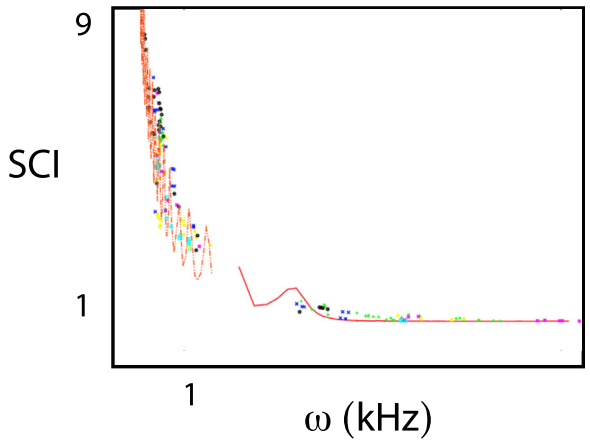

FIG. 6. The relationship between the spectral content and the frequency for syllables, recorded from different birds. The continuous curve is obtained with our model: an oscillation born in a SNILC (saddle node in limit cycle), filtered with a tube.

the computation of this index for 172 sound segments, sung by 6 different birds. The computed values are plotted as color points in a *SCI* vs the *AFF* space. The continuous line, on the other hand, is obtained through the numerical integration of our physical model. To obtain different values of the curve, the value of p_s was changed, and the sounds filtered with a 20 mm tube representing the trachea. The value of k for the simulations was chosen so that the oscillations were born in a saddle node in limit cycle (SNILC) bifurcation. In this way, the growth of the curve for small values of frequency reflects the "explosive" nature of oscillations born in a SNILC.

In this model, modulating the fundamental frequency strongly depends on the region of the space parameter where the system operates. For high values of k, the isofrequency curves in the (p_s, k) space have small local slopes.[18] Therefore, the modulation of frequencies will depend on the modulation of k. On the contrary, for smaller values of the restitution constant k, where the SNILC takes place, the isofrequency curves will be more or less parallel to the SNILC bifurcation curve, which has strong negative local slopes in the (p_s, k) space. Therefore, one can increase the fundamental frequency by increasing p_s for constant values of k. This prediction was actually tested with birds implanted with a valve capable of depressurizing the air sac at selected parts of a given syllable.[19] In these experiments, birds singing harmonic stacks (characterized by long sounds with constant fundamental frequency) had their air sac pressure manipulated through the activation of the valve at the end of the syllables. As expected, a pressure drop induced a decrease in the syllable's fundamental frequency.

VI. TESTING THE PERTINENCE OF THESE MODELS

How can we test the hypothesis that it is the dynamics and not the details of the forces involved what brings together the acoustic features that *better* characterize a birdsong?

Biologists and physicists can have heated discussions over this point. A reductionist model will always make a biologist uneasy: once you set up your mind to study every

biological problem in the framework of evolution (as biologists must do), the simplifications involved in any dynamical model always seem suspicious.

We will start to address this issue by first writing an even simpler set of equations that display the same dynamics in its parameter space than our physical model. Since in the region of interest, there is a Takens Bogdanov bifurcation (where a Hopf line is tangent to one of the two saddle node curves of the cusp bifurcation),[16] we chose its normal form as a simplified dynamical model.[18] Given that three fixed points participate in the dynamics of our physical model, we included cubic terms in the normal form. We chose the signs of the two quadratic terms, and the signs of the two cubic ones, so that the same bifurcations were present in the original problem and in the normal form. Scaling the time through a constant γ, the system reads

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = \alpha\gamma^2 + \beta\gamma^2 x - \gamma^2 x^3 - \gamma x^2 y + \gamma^2 x^2 - \gamma xy,$$

where α and β stand for the unfolding parameters. We adjusted the time constant γ by taking close to thirty song segments from four birds, and computing the SCI and fundamental frequency of each segment. Then, for a given value of γ, we searched for the best synthetic approximation to the original sounds, minimizing the SCI indices and frequencies of synthetic and original sounds. For each γ, the smallest distance (smallest χ^2) was plotted, and we chose its value so that χ^2 presented a minimum.

But, how can we test the *pertinence* of this model? This model is built to generate sounds in which fundamental frequencies and spectral content are related such as in the songs of zebra finches. But does a bird care for those features? We performed a series of experiments in birds, taking advantage of the amazing selectivity that some cortical neurons present to the bird's own song (BOS). This phenomenon, reported by Konishi and Margoliash in the early eighties[20] consists in comparing the activity of neurons in a region of a sleeping bird's brain, while it listens to a recording of its own song, with the activity present while other songs are played. Classical controls include the reverse song, or the song of a conspecific animal. Remarkably, the neurons will spike much more when the BOS is played. It is interesting that in the case of the reverse song i.e., a song played backwards, the same frequencies are present, and yet no response is elicited. Therefore, a natural test for our model would be to compare the response of the bird to its BOS with its response to a synthetic song generated with our normal form.

In order to create a synthetic version of a zebra finch song, we proceeded as follows. We decomposed the song to be copied in 20 ms successive segments, and for each one, its fundamental frequency and SCI were computed.[18] Then, a search in the parameter space (α, β) of the normal form was performed over a grid, so that the synthetic sounds produced would match the fundamental frequencies and the SCI of the segment to be fitted. Then, over that set, a search was carried out so that the SCI of the synthetic sound matched the SCI of the sound segment. That procedure leads, for a segment starting at t_i, to a pair $(\alpha(t_i), \beta(t_i))$. Figure 7 shows an example. Spectrally, the agreement is so good that it came as a bitter surprise that this synthetic song could elicit no neural response whatsoever when the bird was exposed to it. It was not until the filter was better approximated that the selective neurons in the cortex started to respond to the synthetic songs.[21] The last two pieces of the puzzle that were necessary to include were 1. the oroesopharingeal cavity, which lies between the trachea and the beak (see Fig. 1), modeled as a Helmholtz oscillator of approximately $6 \times 6 \times 6\,\text{mm}^3$ following physiological dimensions, and is acoustically responsible for enhancing frequencies close to the 4 kHz for zebra finches and 2. some amount of noise added to the physiological parameter representing the labial tension. It is interesting to observe that, in physics, when the fundamental equations of the problem are known, it is mathematically algorithmic to improve an approximated model. In biology, it is not necessarily easy to know how to enrich a simplified model that did not work. Do we know *a priori* that a hierarchy of importance exists that allows us to establish an order in the simplifications? This question is at the core of the difficulties in modeling biological systems, and there are no recipes beyond close interdisciplinary work.

With a model that was capable of eliciting responses of highly selective neurons to the birds' own songs, we explored the sensitivity of the neural responses to modifications in the model's parameters. We exposed sleeping birds to synthetic songs produced by the model in which different parameters were changed.[21] In Fig. 8, we illustrate the procedure. The first three panels show the sonograms (top) and the accumulated number of spikes that were elicited in 20 trials, each one consisting of exposing the bird to three copies of one song. In the panel on the left, the stimulus was the bird's own song. In the other two panels, we used synthetic sounds. The different stimuli in the experiment were generated by varying the amount of noise added to the parameters and the dissipation of the Helmholtz resonator. The result is shown at the bottom of Fig. 8, where the response to each stimulus is plotted as a fraction of the response to the BOS. Each point on the grid corresponds to a different set of parameters (noise and dissipation). On the

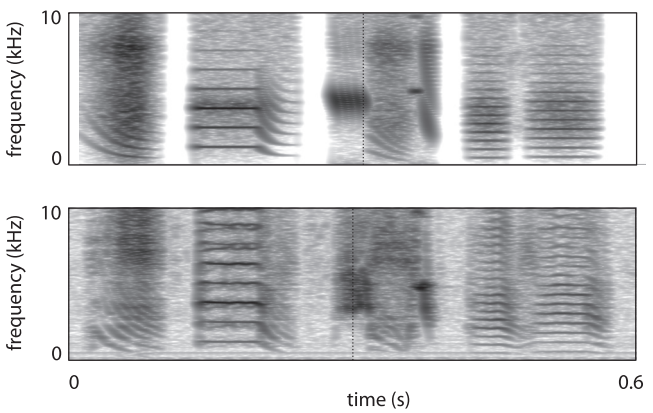

FIG. 7. A recorded song (top panel) and a synthetic sound (bottom panel) generated by the model described in the text.

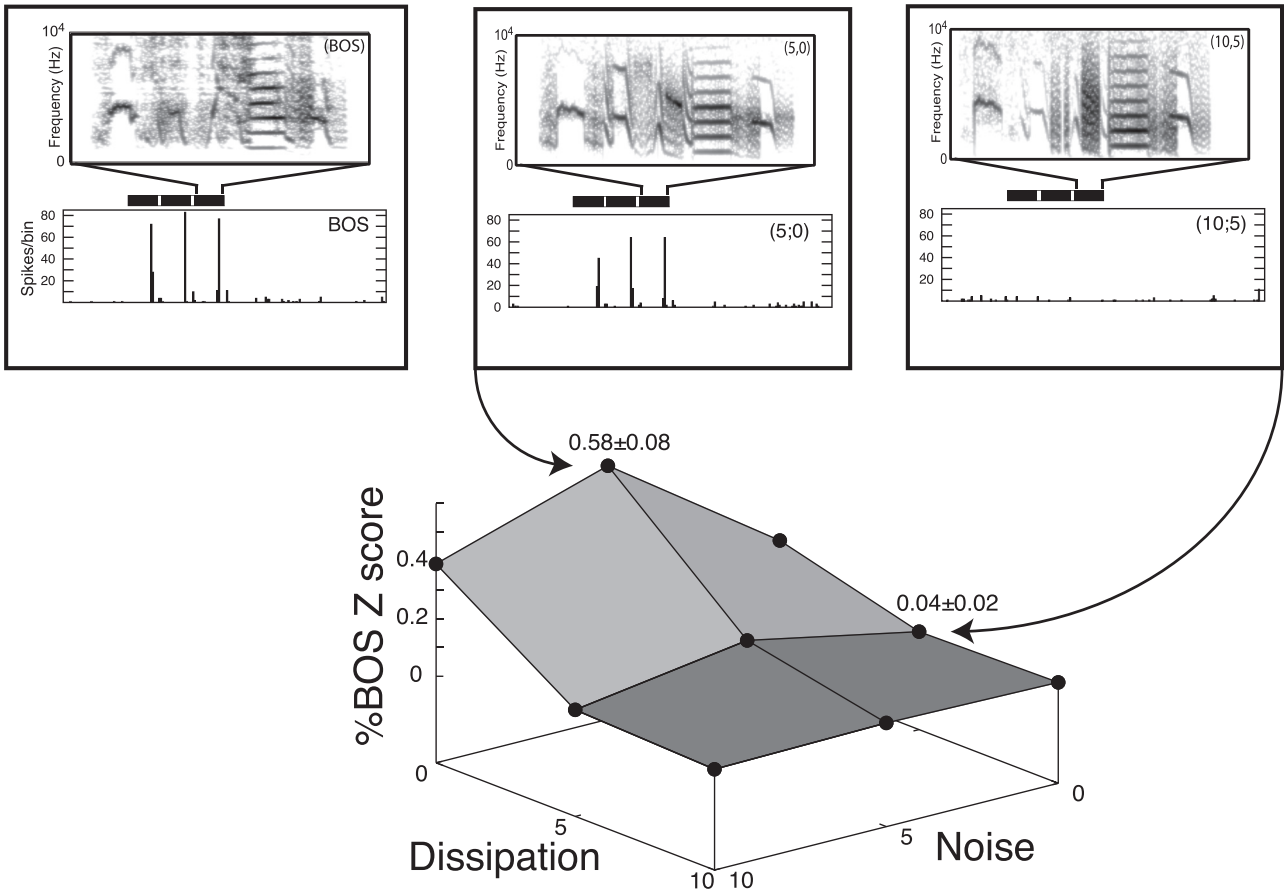

FIG. 8. Testing the response of highly selective neurons to synthetic sounds generated by the model, with slightly different parameters.

one hand, the best response is 60% of that obtained using BOS as a stimulus. On the other hand, we could verify that the system is exquisitely sensitive to changes in the parameters, and rapidly ceases to elicit any response as soon as we depart from the optimal values.[21]

VII. OTHER SIGNATURES OF NONLINEARITY

In the previous models, we assumed that both labia of a sound source were locked and therefore we described the dynamics of the lumen in terms of the mean position of one of the two labia. It is a hypothesis that deserves to be inspected with care, since the two labia will in general be different, and it has been shown to enrich the dynamics.[22,23] In the case of oscine birds, the syrinx is a bipartite structure, and each of the two sound sources is itself asymmetric: the lateral and medial labia are slightly different. Even in birds with a tracheal syrinx (only one sound source, consisting of two opposed labia at the trachea), the labia are not identical. Could it be possible that, in this case, the dynamics are richer than what we discussed so far?

If there were an asymmetry between the labia, the dynamics of their midpoint positions would be ruled by the following dynamical system[23]

$$M_{(l,r)}\frac{d^2\xi_{(l,r)}}{dt^2}+B_{(l,r)}\left(1+\eta_{(l,r)}\xi^2_{(l,r)}\right)\frac{d\xi_{(l,r)}}{dt}+K_{(l,r)}\xi_{(l,r)}=P_g$$

with $M_{(l,r)}$, $B_{(l,r)}$ and $K_{(l,r)}$ being the mass of the tissue, the dissipative constant and the coefficient that accounts for the restitution force of a displaced membrane, respectively (notice that l and r stand for left and right). The constant $\eta_{(l,r)}$ is a nonlinear coefficient that accounts for energy dissipation at large labial displacements. As before, the average pressure can be written in terms of the lower and upper cross-sectional areas a_1, a_2 as

$$P_g = P_s\left(\frac{a_1-a_2}{a_1}\right),$$

where the areas can be written as functions of displacements as

$$a_{(1,2)}(t)=L[\xi_0+\xi_r(t\pm\tau_r)]+L[\xi_0+\xi_l(t\pm\tau_l)],$$

with L being the membranes' length, ξ_0 half of the lumen's width at rest, and $\tau_{r,l}$ the times that the right and left wave-like motions of the respective membrane take to travel half its vertical size. These kinematics of the labia (that assume, as in our previous models, that a lateral mode and a wave-like mode are active), allow us to write (for small values of τ_l, τ_r) the pressure at the lumen in terms of midpoint displacements:

$$P_g\approx\frac{P_s}{\xi_0}\left(\tau_r\frac{d\xi_r}{dt}+\tau_l\frac{d\xi_l}{dt}\right)\equiv\beta\left(\frac{d\xi_r}{dt}+\frac{d\xi_l}{dt}\right).$$

Now, we can write the final dynamical system for the left and right labia as

$$\frac{d\xi_{(l,r)}}{dt} = \nu_{(l,r)},$$

$$\frac{d\nu_{(l,r)}}{dt} = \gamma\frac{(2\beta - B)}{M}\nu_{(l,r)} - \gamma\frac{\beta\eta}{M}\xi^2_{(l,r)}\nu_{(l,r)} - \gamma^2\frac{K_{(l,r)}}{M}\xi_{(l,r)} + \gamma\frac{\beta}{M}\left(\nu_{(r,l)} - \nu_{(l,r)}\right).$$

In order to explore the solutions of this four-dimensional dynamical system as a function of the system's symmetry, we introduced a detuning parameter Q at the level of the linear part of the restitution, i.e., $K_{(l,r)} = K_{0(l,r)} + K_{1(l,r)}\xi^2_{(l,r)}$, with $K_{0(r)} = QK_{0(l)}$. The effect of this asymmetry is summarized in Fig. 9, where we display the different locking regimes as a function of air sac pressure and detuning parameter Q. Region I in the figure corresponds to 1:1 locked solutions, i.e., the two labia oscillate with the same period, while parameters in region II lead to solutions that do not lock. If we describe the equations in terms of a radial and a phase component, the equations of the radial parts will contain a term that depends on the phase difference between the left and right oscillations. Therefore, for the region of the parameter space where the phase difference is time dependent, the solutions will present sidebands in their spectrogram.[24] The reason is that the time dependent phase difference induces the radial parts of the variables to oscillate, giving rise to amplitude modulations.

There are two sister species of pigeons that allowed us to test these ideas.[24] One is the spot-winged pigeon *Patagioenas maculosa* and the other is the picazuro pigeon *P. picazuro*, hereafter called maculosa and picazuro. The temporal patterns and the modulation of the fundamental frequency in their songs are extremely similar, and yet the timbre of their sounds is very different. In Fig. 9, we show their sonograms, where we can see the appearance of sidebands in the song of maculosa. We were able to reproduce these bands by assuming an asymmetry, and, more importantly, we managed to test the hypothesis by examining the syrinx of these two species comparatively. As predicted, the syrinx of maculosa is significantly asymmetric, while picazuro's syrinx is not. In maculosa, all individuals examined had larger left oscillating tissues.

VIII. BEYOND LOW DIMENSIONAL DYNAMICS: VORTEX SOUND

All the acoustic effects described so far were the result of modulating the airflow by means of a valve. The mechanism is similar to the one used by humans when vocalizing voiced sounds, like vowels. And just as humans also use unvoiced sounds (such as "s," or "f"), some bird species alternate syllables of regular oscillations with extremely noisy sounds.

Howe developed the field of aero-acoustics.[9] He considered the source terms that are neglected in the linear approximation of acoustics as follows:

$$\frac{D^2p}{Dt^2} - c_0{}^2\nabla^2 p = \rho_0\nabla.(\overline{w} \times \overline{v}),$$

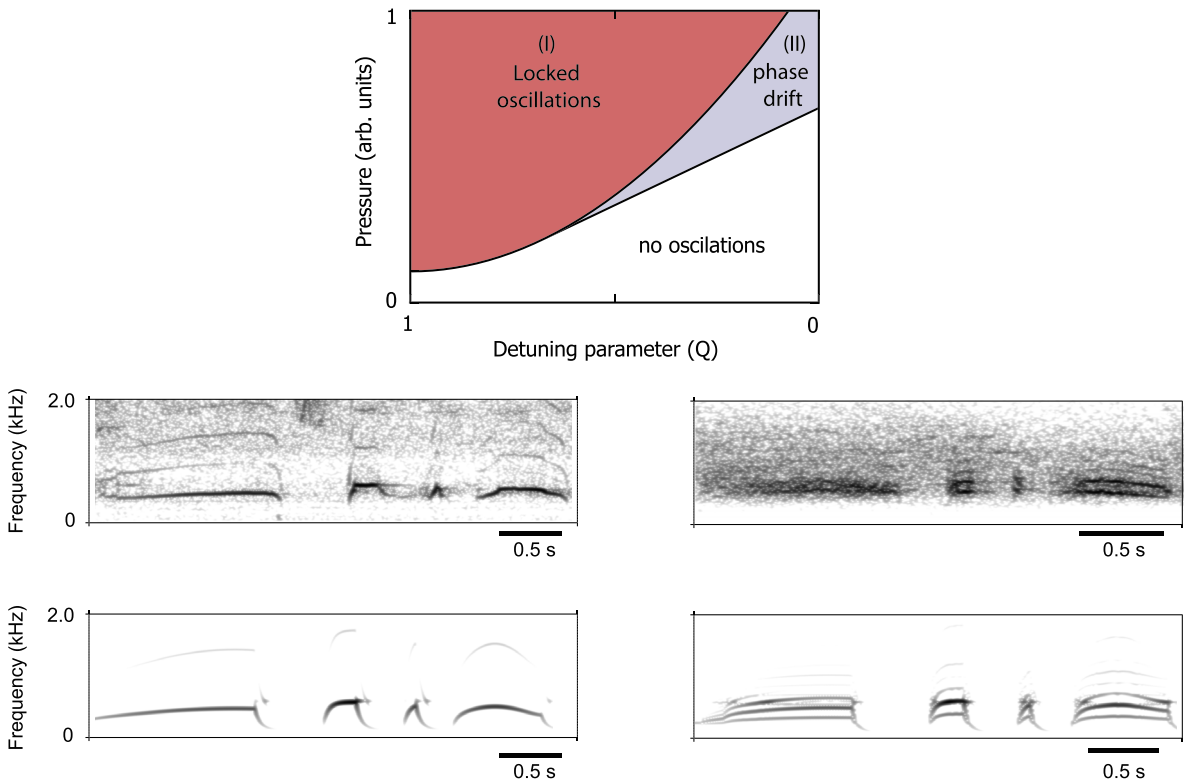

FIG. 9. Asymmetric labia can generate sounds with distinctive timbre (notice the side bands in the sonograms). In the top panel is the bifurcation diagram. The recorded sonogram of the picazuro pigeon song and the synthetic sound generated with symmetric labia (left), compared with the song of maculosa pigeon and sound generated with asymmetric labia (right).

with D being the material derivative, ρ_0 the ambient undisturbed density, $\overline{w}$ the vorticity and $\overline{v}$ the local velocity. Notice that the vector product involves quadratic contributions of velocity to the source of sound. In this way, this expression is a higher order correction to the homogeneous wave equation studied in linear acoustics, so the effects that emerge from this description should be added to those eventually generated by the fluctuating mass injection that results from the modulation of airflow by a valve.

When the flow crosses the lumen before being injected into the trachea, it separates from the walls forming a jet. This is a focused region surrounded by stagnant air, and, in the border of each of these regions, the air particles undergo a rotation. This is quantified using the vorticity of the velocity field. In this way, the shear layers between the jet and the walls coalesce into irregular structures that travel along the trachea. When these vortices arrive at the glottal contraction, sound is generated. The solution of the wave equation can be written in integral form as[9]

$$p(x,t) = -\rho_0 \frac{sgn(x-y)}{2A(1+M)} \int_A \int_y \left[(\overline{\omega} \times \overline{v}).\overline{U}^*\right] dAdy$$
$$\equiv \frac{sgn(x-y)}{A} S\left(y, t - \frac{|x-y|}{c_0(1+M)}\right),$$

where A is the cross-section at the observation point, M is the Mach number and $\overline{U}^*$ stands for the ideal flow velocity that would be present if the duct was filled with a uniform and steady flow. The brackets indicate that the integral is computed at a retarded time $t_{ret} = t - |x-y|/(c_0(1+M))$.

The actual computation is particularly difficult, since it requires the computation of flows and vorticity in a complicated configuration. Yet, a qualitative approximation can be obtained.[25] The vector $\overline{\omega} \times \overline{v}$ points radially when a vortex travels axially. The unperturbed flow $\overline{U}^*$, on the other hand, presents a negative radial component right before the constriction and a positive one right after it. Therefore, the source function S consists of a pulse whose duration is given by $\Delta t \approx H/U_c$, where H is a characteristic size of the constriction and U_c is the flow speed at the constriction. We can estimate an order of magnitude for this speed considering that the volume of an air sac is expelled in the duration of a typical syllable, and that the constriction is a fraction of the tracheal radius.[26] To account for the effect of a train of vortices, we compute the convolution of this source term with an arrival function that consists of a series of delta functions at arrival times $\{T^{arrival}\}$:

$$p_{vortex}(t) = \int S(\tau) \sum \Gamma_n \delta\left(t - T_n{}^{arrival} - \tau\right) d\tau$$

The amplitude of the nth contribution Γ_n depends on its circulation, which is assumed to be proportional to the difference between the nth and $(n-1)$th arrival times. The rationale behind this hypothesis is that a large time between two consecutive vortices means that the second one was generated far away from the constriction, having gained circulation during its travel towards it. Figure 10 illustrates the spectrum of a recorded song, and the spectra of two synthesized versions.[26] For the case displayed in the bottom panel, we followed the procedure described in Sec. VI. For the case displayed in the middle panel we added to that synthesis vortex sound generated as we described above. For frequencies above 10 kHz, the spectrum of the synthesis with the vortex sound is a better approximate to the spectrum of the recorded sound. Since the spectrum of a convolution is the product of the spectra, and the arrival times are taken from a uniform distribution, features like the minima marked with arrows in Fig. 10 originate in the pulse shape. Remarkably, sensible time and geometric estimations lead to spectra with minima in the right spectral values. These contributions carry a very small fraction of the energy, but it will be interesting to test whether a synthesis that takes this effect into account improves the response of selective neurons to surrogate synthetic sounds.

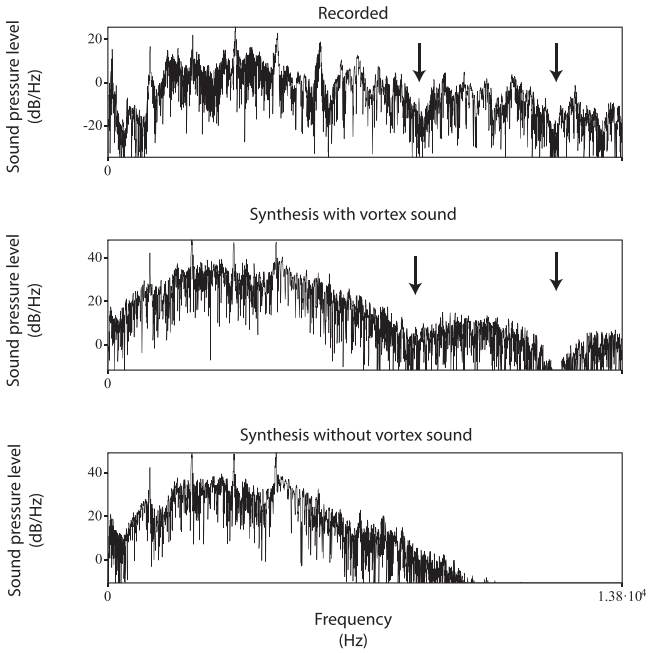

FIG. 10. The spectrum of a song (first panel), synthetic sound with an added vortex sound (middle panel) and the synthetic sound without a vortex sound (bottom panel). There is a significant difference for frequencies higher than 10 kHz.

IX. THE DYNAMICAL ORIGIN OF PHYSIOLOGICAL INSTRUCTIONS

We have discussed how birdsong emerges when a highly specialized biomechanical device, the syrinx, is controlled by a set of exquisitely coordinated physiological instructions (in particular, those controlling syringeal configuration and respiration). During the song, the normal respiratory activity patterns are changed for those that establish the proper airflow necessary for phonation. These physiological instructions are generated in songbirds by a set of interconnected neural nuclei known as the song system.[27,28]

The song system includes the dorsomedial nucleus (DM), which projects to the respiratory pacemaker. This "pacemaker" includes nucleus retroambigualis (RAm), related to expiration, and nucleus parambigualis (PAm), which is active mostly during inspiration. These nuclei are present in other orders of birds

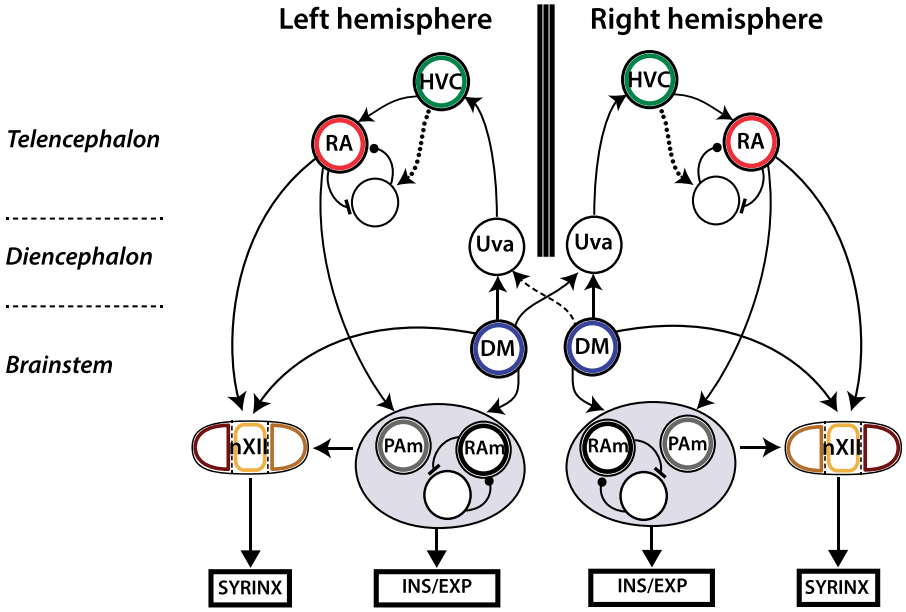

FIG. 11. A schematic diagram of the nuclei in the oscine song system.

as well. But, in songbirds, the DM projects to the nucleus uvaeformis (UVA) in the thalamus, which connects to the telencephalic nucleus HVC (used as a proper name). This nucleus, in turn, projects to the robust arcopallial nucleus (RA), which closes the loop by projecting back to the respiratory nuclei. The nucleus that projects to the syringeal muscles is called nXII, which receives inputs from the telencephalon (RA), as well as from the respiratory nuclei and the DM. This architecture is found symmetrically in the left and right hemispheres of the bird's brain. A schematic diagram summarizing this description is displayed in Fig. 11. To add complexity to the problem, most of these nuclei are built from inhibitory interneurons and excitatory neurons (which can interconnect within the nucleus or sometimes project from one nucleus to another, forming a pathway). Each nucleus can have tens of thousands to several hundreds of thousands of neurons. Considering that unveiling how the system works requires understanding clearly how both interneurons and projection neurons work, it is necessary to study the neural activity of the system with electrodes of high enough impedance so that individual neurons can be isolated and identified. This is an extremely difficult problem, and therefore it is not surprising that even in the case of the nucleus HVC, whose proximity to the scalp makes recording its activity somewhat easier, its code is a matter of heated debate.

What is surprising is that in the song system's output, at least in terms of respiratory gestures, one can identify signatures of low-dimensional nonlinear dynamics.[29,30] Figure 12 displays the pressure patterns measured in the air sac of a canary during song. Four different patterns can be identified: a small oscillation mounted on a DC level, a regular harmonic-like oscillation, an oscillation whose spectrum will display two clear peaks, and a wider pattern with a slow decay. These four basic patterns have been found in different birds, and used to generate a variety of different syllables.[31] There is a range of syllabic rates that can be associated to these different patterns.

In the panel below, we display four solutions of a low dimensional nonlinear dynamical system. Each solution was obtained by driving (differently) a two-dimensional dynamical system, in which one of its variables (let us call it x) was made to behave as the recorded air sac pressure.[32] This dynamical system was built in such a way that in a two-dimensional parameter space, we could find a Hopf

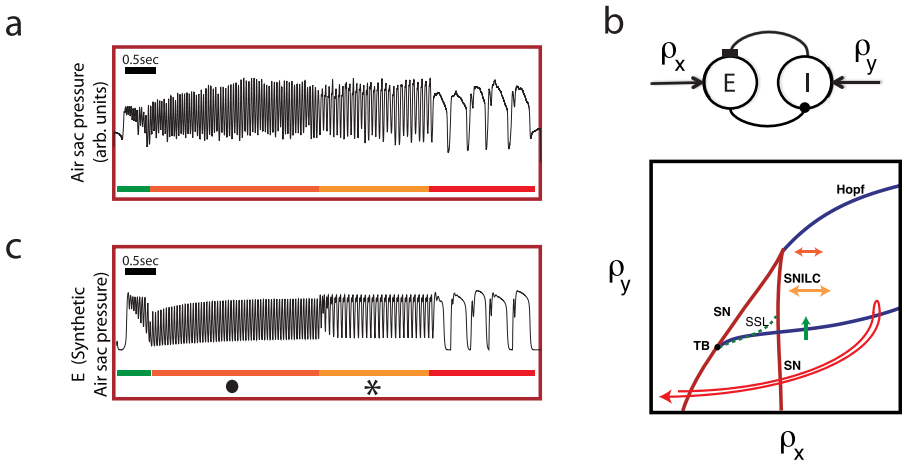

FIG. 12. Recorded air sac pressure in a singing canary (a). A neural oscillator, i.e., a population of excitatory neurons coupled to a population of inhibitory ones (top). The bifurcation diagram showing the different regions of the parameter space presenting qualitatively different flows (bottom). The parameters are the inputs to the neural oscillator (b). The synthetic time series emulating air sac pressure (c). The syllables are generated changing the parameters along the trajectories displayed in (b).

bifurcation and a cusp, with the Hopf bifurcation line tangent to one of the saddle node curves of the cusp.[30] With this dynamical architecture, it is possible to reproduce the observed air sac pressure patterns driving the dynamical system with simpler time dependent parameters.[32,33] For example, the low rate solutions are obtained by driving the model with the red path in Fig. 12(b). This path consists of a fast change of ρ_x and ρ_y that takes the system from a region of the parameter space where there is a fixed point of low x value to a region with a fixed point of high x value. Then, the parameters ρ_x, ρ_y are slowly decreased, and the solution decreases first slowly, because the value of the fixed point decreases. Then, when the saddle fixed point kills the attractor at the leftmost saddle node branch, the solution rapidly decreases as the system collapses towards the only attractor of the system at the left of the cusp: the original fixed point of low x value. The tiny peaks at the beginning of the pulse can be generated as transients, since the path that we described reaches a region of the parameter space where the eigenvalues of the fixed point of high x value are complex due to the vicinity of the Hopf bifurcation line. The rest of the patterns can be generated within this architecture as well. The small and large amplitude oscillations with one maximum are found by placing the dynamical system in two different points within the region of the parameter space with oscillations born at the Hopf bifurcation. Both solutions can be obtained with stationary parameters ρ_x, ρ_y, or with $\rho_x(t)$, $\rho_y(t)$ fluctuating periodically, as long as the frequency of those fluctuations is similar to the natural frequency of the oscillations born in the Hopf bifurcation. Remarkably, if there is indeed a forcing frequency, and we increase it, the system can undergo a period doubling, responding with one pressure pulse to two peaks of the forcing. This is the pattern marked with an asterisk in Fig. 12(c).

It is suggestive that this dynamical skeleton can be found in the simplest model of an excitatory neural population coupled to an inhibitory one. Using the mathematical description proposed by Wilson and Cowan, let us call x and y the average activities of the excitatory and inhibitory populations, respectively. We can describe their dynamics through:

$$\frac{dx}{dt} = -x + S(\rho_x + a_1 x + a_2 y),$$
$$\frac{dy}{dt} = -y + S(\rho_y + a_3 x + a_4 y)$$

with

$$S(x) = \frac{1}{1+e^{-x}}.$$

The bifurcation diagram for this system is precisely the diagram displayed in Fig. 12(b) for $(a_1, a_2, a_3, a_4) = (10, -10, 10, 2)$.

A natural question is whether the forcing parameters $\rho_x(t)$, $\rho_y(t)$ can themselves be the solutions of low dimensional systems with variables associated to the activities of other parts of the song system. These time dependent parameters now summarize the inputs from both RA (the telencephalic nucleus that projects to the respiratory nuclei) and the inputs from other parts of the brainstem, such as the inspiratory related nucleus PAm (which does have a subpopulation of neurons that fire during expiration), or nucleus DM.

Figure 13 shows how to generate the four basic respiratory patterns used in a canary song with solutions of a dynamical model embedding our minimal toy description of the neural oscillator.[32,33] Since one of the variables of that minimal two-dimensional dynamical system was made to represent the air sac pressure, our neural oscillator is a toy model for the expiratory related area RAm. Yet, the global dynamical system in which our model of RAm is embedded (the circular architecture displayed in Fig. 13) makes specific predictions on the kind of activity patterns that can be expected in other areas of the song system. For example, the model predicts that in this species, when the long expiratory

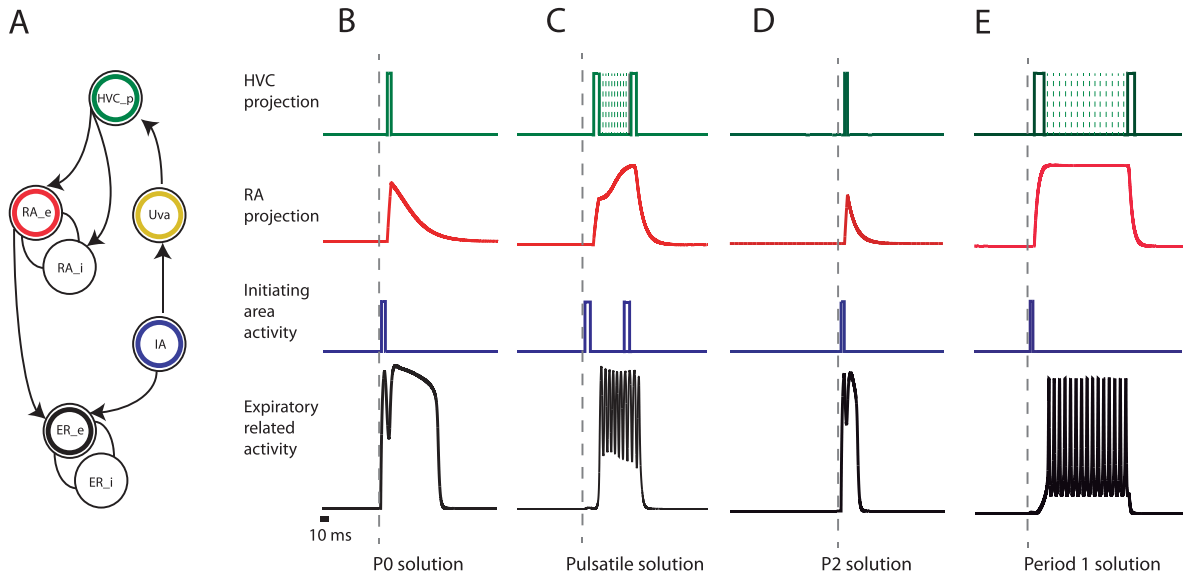

FIG. 13. One-way of generating synthetic pressure gestures similar to the recorded ones. The dynamical system is written in terms of the activities of interconnected nuclei.

patterns are generated, the activity of the HVC projection neurons will be heterogeneous, with a continuous component superimposed on peaks of activity close to the beginning of the syllables. Notice that the respiratory patterns emerge as the dynamics of the brainstem is affected by signals from the telencephalon. But, it is not until they interact nonlinearly that the richness of the solutions emerges. Anatomy supports the idea of a circular functional architecture, as do experiments in which thermal manipulation affects times scales at different parts of the song system.[34–36] However, this picture is under construction, and it is most likely to evolve as experiments test its predictions. For example, the degree of heterogeneity in zebra finches' HVC activity has been quantified by the measurement of hundreds of individual neurons.[37] It is likely that similar studies will be performed on other species such as canaries in the near future, and that other nuclei will be measured in singing birds as well. Whatever the picture that emerges in the end, it is remarkable that non-trivial yet low dimensional dynamics are behind the physiological gestures necessary for birdsong production. This offers an extremely interesting example on how to build low dimensional dynamical systems from first principles, for nontrivial neural architectures.

X. TOWARDS MODELS FOR AVERAGE ACTIVITIES

The results described in Sec. IX point to an interesting issue for nonlinear dynamics and statistical physics. How to establish the connections between the different description timescales of out-of-equilibrium systems is an active field of research. This is particularly relevant in neuroscience, where the central nervous system is in charge of the physiological instructions that control peripheral systems and, therefore, of the interaction of the organism with the macroscopic world. As there is no comprehensive theory to deal with out-of-equilibrium statistical mechanics, macroscopic models of nervous systems are usually built phenomenologically and not statistically. Having said that, there have been very important advances recently. In part, this is due to the capacity to perform biologically informed numerical simulations.[38] But, there have been important advances in the theory of collective dynamics of out of equilibrium units. In 2008, Ott and Antonsen[39] reported a method to compute the dynamics of a macroscopic observable describing the degree of synchronization of a set of interconnected phase oscillators.[40,41] Their result is framed into a very active line of research: the study of globally coupled phase oscillators. They showed that it is possible, for a class of systems, to write a low dimensional dynamical system to describe the asymptotic behavior of the system's order parameter. Let us briefly review the main ideas of this development, in the framework of the coupling between an excitatory and an inhibitory neural population (the basic elements of a neural oscillator architecture we used in Sec. IX).[42]

Let us assume that each excitable unit is described in terms of its phase, and that its dynamics is modeled through Adler's equation. Let us also assume that the units can be classified as excitatory and inhibitory, and that the populations are coupled:

$$\frac{d\theta_i}{dt} = \omega_i - \cos(\theta_i) + I\Big(\{\theta_j\},\{\tilde{\theta}_j\}\Big),$$
$$\frac{d\tilde{\theta}_i}{dt} = \omega_i - \cos(\tilde{\theta}_i) + \tilde{I}\Big(\{\theta_j\},\{\tilde{\theta}_j\}\Big)$$

with

$$I\Big(\{\theta_j\},\{\tilde{\theta}_j\}\Big) = \frac{k_E}{N}\sum_{j=1}^{N}(1-\cos(\theta_j)) - \frac{k_I}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\Big(1-\cos\Big(\tilde{\theta}_j\Big)\Big),$$
$$\tilde{I}\Big(\{\theta_j\},\{\tilde{\theta}_j\}\Big) = \frac{\tilde{k}_E}{N}\sum_{j=1}^{N}(1-\cos(\theta_j)) - \frac{\tilde{k}_I}{\tilde{N}}\sum_{j=1}^{\tilde{N}}\Big(1-\cos\Big(\tilde{\theta}_j\Big)\Big).$$

In these equations, the parameters $k, \tilde{k}$ describe the coupling, which consists of an impulse when $\theta_j \sim \pi$, and approaches zero as $\theta_j \sim 0$. The first approximation to the macroscopic description of the system is through its order parameters:

$$z = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}, \quad \tilde{z} = \frac{1}{\tilde{N}}\sum_{j=1}^{\tilde{N}} e^{i\tilde{\theta}_j}.$$

One can describe the system though the distributions of its phases. It is in the calculation of those distributions that Ott and Antonsen made a major contribution. These have to obey a continuity equation (a partial differential equation), and therefore a mode decomposition of the distributions would lead to an infinite number of ordinary differential equations for the mode amplitudes. Yet, for a large class of systems, all those equations can be satisfied provided that the first mode of each distribution is ruled by a simple dynamical system. This allows us to simplify our problem, which can be completely described by solving the following dynamical system:

$$\frac{dz}{dt} = [-\Delta + i(\omega_0 + I(z,\tilde{z}))]z - \frac{i}{2}(1+z^2),$$
$$\frac{d\tilde{z}}{dt} = \Big[-\tilde{\Delta} + i\big(\tilde{\omega}_0 + \tilde{I}(z,\tilde{z})\big)\Big]\tilde{z} - \frac{i}{2}(1+\tilde{z}^2),$$

where the parameters $(\omega_0, \Delta, \tilde{\omega}_0, \tilde{\Delta})$ describe the centroids and dispersions of distributions of the parameters $(\omega_i, \tilde{\omega}_i)$.

The order parameters do not correspond to average activities; they are not the variables of the phenomenological models used in Sec. IX. The order parameters measure the degree of synchronization of the populations. Yet, we can compute the average activities $(\phi(t), \tilde{\phi}(t))$ of the populations in terms of them.[42] By definition, the average activity of the populations can be computed as:

$$\phi(t) = \int_{-\infty}^{\infty} f(\theta,\omega,t)(\omega - \cos(\theta) + I(\{z\},\{\tilde{z}\}))|_{\theta=\pi} d\omega,$$

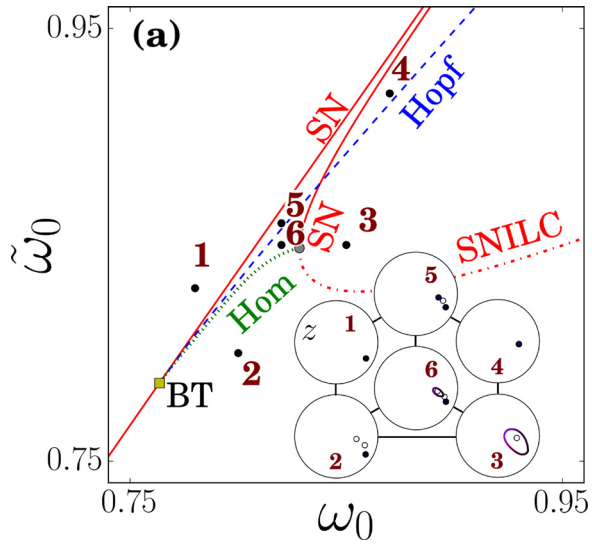

FIG. 14. Parameter space regions of the average model.

$$\tilde{\phi}(t) = \int_{-\infty}^{\infty} \tilde{f}(\theta,\omega,t)(\omega - \cos(\theta) + \tilde{I}(\{z\},\{\tilde{z}\}))|_{\theta=\pi} d\omega,$$

where $f(\theta,\omega,t)$ and $\tilde{f}(\theta,\omega,t)$ stand for the distribution density of phases for the two populations. Expressing these distributions as functions of the order parameters, we get:

$$\phi(t) = \frac{1}{\pi}\left(\frac{1+Re\,z}{|1+z|^2} - \frac{1}{2}\right)(\omega_0 + 1 + I(z,\tilde{z})) + \frac{\Delta}{\pi}\left[\frac{Im\,z}{|1+z|^2}\right],$$

$$\tilde{\phi}(t) = \frac{1}{\pi}\left(\frac{1+Re\,\tilde{z}}{|1+\tilde{z}|^2} - \frac{1}{2}\right)(\tilde{\omega}_0 + 1 + \tilde{I}(z,\tilde{z})) + \frac{\tilde{\Delta}}{\pi}\left[\frac{Im\,\tilde{z}}{|1+\tilde{z}|^2}\right],$$

which are quantities that can be directly compared with the variables (x, y) of the phenomenological model studied in the previous section.

The dynamical elements needed in the phenomenological model to reproduce the respiratory patterns are also present in the model derived from first principles (see Fig. 14).[42] It is remarkable that a very simple phenomenological model such as the Wilson-Cowan oscillator can capture the subtle features that a population of coupled excitable units displays macroscopically. Needless to say,, our calculation assumes serious simplifications. One of them is that we have considered an all-to-all coupling between the units. It is likely that as soon as more complex and informed topologies are used, larger phase spaces will have to be explored. This might lead to more complex solutions, which will lead to predictions that need to be compared with behavioral observations.

Since different biological mechanisms can be described in terms of the same dynamical elements, modeling will always require a continuous dialogue between dynamics and experiments. But, having described the physiological observables in a dynamical language, we can start to explore the—long—road from first principles to behavior.

XI. CONCLUSIONS

This review was organized around a few questions. Is it possible that part of the complexity of the sounds used in a birdsong is due to the nonlinear nature of the avian vocal organ? Could it be that reasonably simple instructions driving this highly nonlinear device create this amazing richness of sounds? How complex are those instructions? How can we determine that a low dimensional model is biologically pertinent? We got some answers to these questions by studying specific problems. We found, for example, that the relationship between the fundamental frequency and the spectral content of the sounds in the zebra finch song is the signature of the bifurcation that gives rise to the oscillations responsible for airflow modulation. We described the dynamical origin of subtle spectral features such as sidebands in the sonogram of a song: it is the result of the loss of locking between the parts of the syringeal valve for a very asymmetric syrinx. But, behind all those specific stories, there is one important question, near and dear to the heart of any dynamicist: how relevant is this simplifying approach to a biological problem? Are these approximations pertinent?

When a dynamicist starts interacting with biologists, it is always complicated. A first thought is that what lies at the root of the difficulty is the biologist's lack of mathematical background, or the dynamicist's lack of knowledge on biology. In fact, it is even more complicated: the very same idea of what it means to be rigorous is different in the two communities. A dynamicist searches for depth and elegance in minimal mechanisms. A biologist, on the other hand, will not find either of those virtues without framing the problem under study within its evolutionary context. Therefore, a biologist will be, *a priori*, suspicious of reductions and simplifications, just as a dynamicist will lose interest in a problem as "details" start to add up. It is for these reasons that our experiments on testing the models using selective neurons were so important in our research program. Granted, there were a minimal number of elements that had to be included for the selective neurons to start reacting to our synthetic sounds as they did to the bird's own song. But, there was a hierarchy of importance in the different elements in the model: changes in different parameters would lead to response losses at different rates. This opens a wide opportunity for evolutionary interpretations.

It is remarkable that the diversity of pressure gestures used in birdsong, emerging from a real brain, can be so beautifully reproduced by integrating a periodically forced two-dimensional dynamical system. We know that the topological organization of different orbits is a distinctive signature of the mechanisms underlying the dynamics. Therefore, even if the particular model proposed for generating the pressure gestures evolves as experiments inform us on the song system, the identified dynamical skeleton will have to be present in eventual upgrades to the model. A thorough and direct testing of these neural models will take time. The connections between the nuclei involve only part of the neurons (the projecting ones), that require to measure with high impedance electrodes, so that individual neurons can be identified. Only through direct collision experiments, can the nature of each neuron under study be verified. And, of course, enough neurons have to be measured so that the actual physiological instructions can be reconstructed. In the meantime, we can improve or refute our models through indirect measurements

that we can carry out, thanks to the existence of dynamical models, such as the degree of heterogeneity in the average activity of a given nucleus, or by testing the precise and quantitative predictions on the effect of cooling. In other words, dynamical models are not only summaries of observations, but ways to carry out nontrivial and precise quantitative predictions. Interestingly enough, we are now closer to obtaining these low dimensional models from mean field computations of coupled excitable systems. Recent developments in the field of synchronization have provided us with tools to advance forward from a pure phenomenological ansatz.

In this review, I have focused my attention on birdsong production, but studying the deep connection between the biomechanical periphery and the nervous system is a strategy that will allow us advancing in our understanding of many different problems in neuroscience, particularly those requiring sensory motor integration. For example, the developmental dynamics of marmoset monkey vocal production was investigated by studying in parallel the maturation of the vocal apparatus and the effect of parental feedback.[43] It is likely that recent descriptions of the functional organization of the human sensory-motor cortex for speech articulation will allow us linking the average activity of specific neural populations with the parameters needed to synthesize human speech.[44]

But, behavior emerges not only from the subtle interaction between the nervous system and the biomechanics of the periphery. The environment interacts with the biomechanics as well, and in non-trivial ways. Some animal bodies present stable motions that require relatively little activation or control, and it is only when the whole system (central nervous system-biomechanics-environment) is analyzed that the behavior can be understood. Lamprey dynamics during swimming constitutes a stellar example of this phenomenon.[45] Moreover, proprioception as a mechanism of feedback from the body to the neural circuits involved in the behavior is the key to understand stability and robustness. It is clear that the more integrated a model gets, the stronger the need for a dynamical description.

Quantitative models will enrich our study and understanding of biological problems. Yet, their construction requires subtle work. The difference in approach to a problem in physics, dynamics and biology is deeply rooted in the essence of each discipline, and the path of working across disciplines is hard. Nevertheless, the stories are worth the effort, and they will probably be written in the language of dynamics.

ACKNOWLEDGMENTS

I would like to thank Marcelo Magnasco and Guillermo Cecchi for letting me know about this field. Many of my adventures in it would not have been possible without the help of Franz Goller, who taught me how to measure many of the physiological parameters that I needed for the modeling efforts, as well as for endless hours of fun talks on biology, dynamics, and nature. Dan Margoliash opened the door to neural recordings for the people in our lab, and I still learn about biology and good scientific taste from him. Finally, I would like to thank Ana Amador, who probably has been the best experimentalist in our team since she was a student. These stories are the result of a fruitful interaction with many students and post docs in the lab, and I deeply thank them for their effort and commitment. I would like to specially thank Ceci Herbert for her careful reading. Finally, the University of Buenos Aires, Conicet, Agencia and NIH supported our research continuously.

[1]*Neuroscience of Birdsong*, edited by H. P Zeigler and P. Marler (Cambridge University Press, Cambridge, 2008), pp. 5–31.
[2]P. Marler, "Birdsong and speech development: Could there be parallels? There may be basic rules governing vocal learning to which many species conform, including man," Am. Sci. **58**(6), 669–673 (1970).
[3]H. J. Chiel and R. D. Beer, "The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment," Trends Neurosci. **20**(12), 553–557 (1997).
[4]R. A. Suthers, F. Goller, and C. Pytte, "The neuromuscular control of birdsong," Philos. Trans. R. Soc. B: Biol. Sci. **354**(1385), 927 (1999).
[5]M. F. Schmidt and F. Goller, "Breathtaking songs: Coordinating the neural circuits for breathing and singing," Physiology **31**(6), 442–451 (2016).
[6]M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy, "Neural population dynamics during reaching," Nature **487**(7405), 51–56 (2012).
[7]R. W. Warner, "The anatomy of the syrinx in passerine birds," J. Zool. **168**(3), 381–393 (1972).
[8]T. Riede and F. Goller, "Functional morphology of the sound generating labia in the syrinx of two songbird species," J. Anat. **216**(1), 23–36 (2010).
[9]M. S. Howe, *Theory of Vortex Sound* (Cambridge University Press, 2003), Vol. 33.
[10]F. Goller and R. A. Suthers, "Role of syringeal muscles in gating airflow and sound production in singing brown thrashers," J. Neurophysiol. **75**(2), 867–876 (1996).
[11]T. Gardner, G. Cecchi, M. Magnasco, R. Laje, and G. B. Mindlin, "Simple motor gestures for birdsongs," Phys. Rev. Lett. **87**(20), 208101 (2001).
[12]R. Alonso, F. Goller, and G. B. Mindlin, "Motor control of sound frequency in birdsong involves the interaction between air sac pressure and labial tension," Phys. Rev. E **89**(3), 032706 (2014).
[13]J. A. Alliende, J. M. Méndez, F. Goller, and G. B. Mindlin, "Hormonal acceleration of song development illuminates motor control mechanism in canaries," Dev. Neurobiol. **70**(14), 943–960 (2010).
[14]G. B. Mindlin, T. J. Gardner, F. Goller, and R. Suthers, "Experimental support for a model of birdsong production," Phys. Rev. E **68**(4), 041908 (2003).
[15]I. R. Titze, "The physics of small amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**(4), 1536–1552 (1988).
[16]A. Amador and G. B. Mindlin, "Beyond harmonic sounds in a simple model for birdsong production," Chaos **18**(4), 043123 (2008).
[17]J. D. Sitt, A. Amador, F. Goller, and G. B. Mindlin, "Dynamical origin of spectrally rich vocalizations in birdsong," Phys. Rev. E **78**(1), 011905 (2008).
[18]Y. S. Perl, E. M. Arneodo, A. Amador, F. Goller, and G. B. Mindlin, "Reconstruction of physiological instructions from Zebra finch song," Phys. Rev. E **84**(5), 051909 (2011).
[19]A. Amador and D. Margoliash, "A mechanism for frequency modulation in songbirds shared with humans," J. Neurosci. **33**(27), 11136–11144 (2013).
[20]D. Margoliash and M. Konishi, "Auditory representation of autogenous song in the song system of white-crowned sparrows," Proc. Natl. Acad. Sci. **82**(17), 5997–6000 (1985).
[21]A. Amador, Y. S. Perl, G. B. Mindlin, and D. Margoliash, "Elemental gesture dynamics are encoded by song premotor cortical neurons," Nature **495**(7439), 59–64 (2013).
[22]I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal fold model," J. Acoust. Soc. Am. **97**(3), 1874–1884 (1995).
[23]J. C. Lucero and J. Schoentgen, "Modeling vocal fold asymmetries with coupled van der Pol oscillators," in *Proceedings of Meetings on Acoustics ICA2013* (ASA, 2013), Vol. **19**, No. 1, p. 060165.
[24]R. G. Alonso, C. Kopuchian, A. Amador, M. de los Angeles Suarez, P. L. Tubaro, and G. B. Mindlin, "Difference between the vocalizations of two sister species of pigeons explained in dynamical terms," J. Comp. Physiol. A **202**(5), 361–370 (2016).
[25]M. H. Krane, "Aeroacoustic production of low-frequency unvoiced speech sounds," J. Acoust. Soc. Am. **118**(1), 410–427 (2005).

[26]G. B. Mindlin, "Avian vocal production beyond low dimensional models," J. Stat. Mech.: Theory Exp. **2017**(2), 024005 (2017).
[27]F. Nottebohm, T. M. Stokes, and C. M. Leonard, "Central control of song in the canary, Serinus canarius," J. Comp. Neurol. **165**(4), 457–486 (1976).
[28]R. C. Ashmore, J. M. Wild, and M. F. Schmidt, "Brainstem and forebrain contributions to the generation of learned motor behaviors for song," J. Neurosci. **25**(37), 8543–8554 (2005).
[29]M. A. Trevisan, G. B. Mindlin, and F. Goller, "Nonlinear model predicts diverse respiratory patterns of birdsong," Phys. Rev. Lett. **96**(5), 058103 (2006).
[30]L. M. Alonso, J. A. Alliende, F. Goller, and G. B. Mindlin, "Low-dimensional dynamical model for the diversity of pressure patterns used in canary song," Phys. Rev. E **79**(4), 041929 (2009).
[31]A. Amador, S. Boari, and G. B. Mindlin, "From perception to action in birdsong production: Dynamics of a whole loop," Current Opinion in Systems Biology **3**, 30–35 (2017).
[32]R. G. Alonso, M. A. Trevisan, A. Amador, F. Goller, and G. B. Mindlin, "A circular model for song motor control in Serinus canaria," Front. Comput. Neurosci. **9**, article 41, 1–9 (2015).
[33]R. G. Alonso, A. Amador, and G. B. Mindlin, "An integrated model for motor control of song in Serinus canaria," J. Physiol.-Paris **110**(3) Part A, 127–139 (2016).
[34]K. Hamaguchi, M. Tanaka, and R. Mooney, "A distributed recurrent network contributes to temporally precise vocalizations," Neuron **91**(3), 680–693 (2016).
[35]M. A. Goldin, L. M. Alonso, J. A. Alliende, F. Goller, and G. B. Mindlin, "Temperature induced syllable breaking unveils nonlinearly interacting timescales in birdsong motor pathway," PLoS One **8**(6), e67814 (2013).
[36]M. A. Goldin and G. B. Mindlin, "Evidence and control of bifurcations in a respiratory system," Chaos **23**(4), 043138 (2013).
[37]G. F. Lynch, T. S. Okubo, A. Hanuschkin, R. H. Hahnloser, and M. S. Fee, "Rhythmic continuous-time coding in the songbird analog of vocal motor cortex," Neuron **90**(4), 877–892 (2016).
[38]L. Gibb, T. Q. Gentner, and H. D. Abarbanel, "Brain stem feedback in a computational model of birdsong sequencing," J. Neurophysiol. **102**(3), 1763–1778 (2009).
[39]E. Ott and T. M. Antonsen, "Low dimensional behavior of large systems of globally coupled oscillators," Chaos **18**(3), 037113 (2008).
[40]Y. Kuramoto, "Collective synchronization of pulse-coupled oscillators and excitable units," Phys. D: Nonlinear Phenom. **50**(1), 15–30 (1991).
[41]S. H. Strogatz, "From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators," Phys. D: Nonlinear Phenom. **143**(1), 1–20 (2000).
[42]J. Roulet and G. B. Mindlin, "Average activity of excitatory and inhibitory neural populations," Chaos **26**(9), 093104 (2016).
[43]D. Y. Takahashi, A. R. Fenley, Y. Teramoto, D. Z. Narayanan, J. I. Borjon, P. Holmes, and A. A. Ghazanfar, "The developmental dynamics of marmoset monkey vocal production," Science **349**(6249), 734–738 (2015).
[44]K. E. Bouchard, N. Mesgarani, K. Johnson, and E. F. Chang, "Functional organization of human sensorimotor cortex for speech articulation," Nature **495**(7441), 327 (2013).
[45]E. D. Tytell, P. Holmes, and A. H. Cohen, "Spikes alone do not behavior make: Why neuroscience needs biomechanics," Curr. Opin. Neurobiol. **21**(5), 816–822 (2011).