EPJ B
Condensed Matter and Complex Systems
EPJ.org
your physics journal

Eur. Phys. J. B (2014) 87: 300 DOI: 10.1140/epjb/e2014-50566-5

Low dimensional dynamics in birdsong production

Ana Amador and Gabriel B. Mindlin

edp sciences
Società Italiana di Fisica
Springer

Eur. Phys. J. B (2014) 87: 300
DOI: 10.1140/epjb/e2014-50566-5

THE EUROPEAN PHYSICAL JOURNAL B

Colloquium

Low dimensional dynamics in birdsong production

Ana Amador[a] and Gabriel B. Mindlin

Department Physics, FCEyN, University of Buenos Aires and IFIBA-CONICET, 1428 Buenos Aires, Argentina

Received 20 August 2014 / Received in final form 11 November 2014
Published online 10 December 2014 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2014

Abstract. The way in which information about behavior is represented at different levels of the motor pathway, remains among the fundamental unresolved problems of motor coding and sensorimotor integration. Insight into this matter is essential for understanding complex learned behaviors such as speech or birdsong. A major challenge in motor coding has been to identify an appropriate framework for characterizing behavior. In this work we discuss a novel approach linking biomechanics and neurophysiology to explore motor control of songbirds. We present a model of song production based on gestures that can be related to physiological parameters that the birds can control. This physical model for the vocal structures allows a reduction in the dimensionality of the behavior, being a powerful approach for studying sensorimotor integration. Our results also show how dynamical systems models can provide insight into neurophysiological analysis of vocal motor control. In particular, our work challenges the actual understanding of how the motor pathway of the songbird systems works and proposes a novel perspective to study neural coding for song production.

1 Introduction

Birdsong is a complex behavior, which emerges out of the interaction between a nervous system, a peripheral biomechanical device, and the environment. At least for approximately forty percent of the known species, some degree of learning is involved in the process of acquiring the species-specific vocalizations. This task is quite rare in the animal kingdom and so birdsong has emerged as an important animal model for studying how a nervous system reconfigures itself during the acquisition of a complex behavior through learning [1–3].

Between the nervous system and the actual song stands a peripheral system, including the respiratory system, the vocal tract and the sound source. The last one is a complex biomechanical device called the *syrinx*. This is a bipartite structure between the bronchi and the trachea that holds two pairs of internal labia, which modulate the airflow generating sound waves. The configuration of this device can be controlled by the activation of specific muscles, and those changes are ultimately transduced in acoustic modulations of the uttered sounds. Since the generation of sound requires also establishing airflow between the internal labia, a bird has to exquisitely coordinate respiratory and syringeal muscles to provide the uttered sounds with specific acoustic features [4].

Those physiological instructions in charge of controlling the syrinx are the output of a nervous system, which dedicates well-defined neural nuclei to their production. Each nucleus is set of a few thousands of interconnected neurons, and the sets are also connected among themselves, constituting a structure known as the song system. A variety of techniques have been applied to unveil how the different parts of the song system end up generating the necessary motor patterns in control of the syrinx. Despite important efforts in the last decades, the problem has shown to be quite elusive.

The problem seems discouragingly complex and yet, in this work we describe three instances where low dimensional nonlinear dynamics allows unveiling important aspects of birdsong production. First, we show that many acoustic features present in birdsong arise when relatively simple physiological instructions drive a simple model that captures the essential dynamics of the syrinx. We tested this model by exposing both anaesthetized and sleeping birds to replicas of their own songs, and compared the responses of neurons selective to the bird's own song to the synthetic stimuli generated with a low dimensional nonlinear model. The similarity in the neural responses allowed us to build confidence on the pertinence of a low dimensional description of the biomechanics involved in birdsong production. Afterwards, we move one step closer to the nervous system, and show that those relatively simple physiological gestures needed to drive the syrinx can also be generated as the solutions of a low dimensional dynamical system representing a simple neural architecture.

The work is organized as follows. In Section 2 we describe the biomechanics of birdsong production, and the mechanisms by which relatively simple physiological instructions (air sac pressure and labial tension) are

[a] e-mail: anita@df.uba.ar

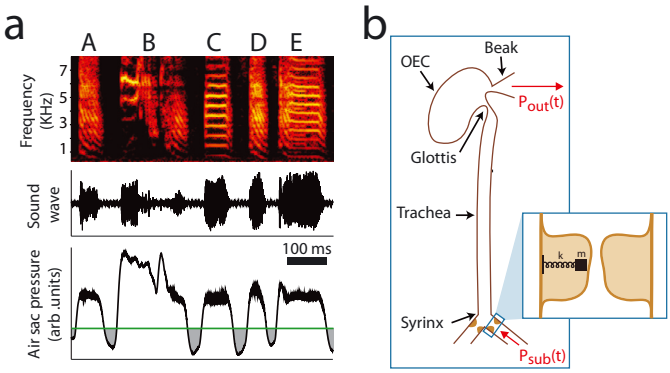

Fig. 1. Typical zebra finch song and vocal production model. (a) The top panel shows the spectrogram of the sound wave (middle panel) corresponding to a zebra finch utterance. The pressure pattern used to generate this song is shown in the lower panel, where each expiratory pulse (pressure above green line) defines a syllable indicated with letters in the top panel. The grey area indicates inspiratory pulses. (b) The syrinx is positioned at the base of the trachea, followed by the glottis, the oro-esophageal cavity (OEC) and the beak. All these structures form the vocal tract that acts as a filter of the sound wave generated at the syrinx. The syringeal labia are modeled as a nonlinear oscillator.

transduced into complex sounds. Section 3 describes air sac pressure dynamics in the case of canaries. We show that pressure patterns can be obtained as a the result of the nonlinear competition between two time scales, and describe experiments that we performed to manipulate one of these time scales in order to test our prediction. In Section 4 we show that the same neural architecture used to generate the previously described physiological instructions is capable of generating the labial tension in a different species. In Section 5 we discuss these results.

2 Generating acoustically complex signals with reasonably simple instructions

Song spectrograms have been the most extended tool for describing singing behavior and relating neural recordings to singing (see top panel of Fig. 1a).

These are obtained by performing a time-frequency analysis of the sound signal. This linear technique allows inferring, graphically, several acoustic features of the song, as the fundamental frequency, the harmonic content of the time trace, or the relative emphasis of frequencies (color coded in the sonogram). It is natural to ask whether the bird has to code (and eventually learn) all those features independently, or if the nonlinear nature of the avian vocal organ constrains those features. Ultimately, it is the dynamics of the structures – principally the dynamics of the intrinsic syringeal musculature controlling the tension of the syringeal labia and of the thoracic and abdominal musculature controlling air sac pressure – that produces the airflow fluctuations which after being filtered by the upper vocal structures generates the sound.

In an extensive development over many years, we have tested the hypothesis that much of the acoustical complexity present in oscine song was the result of relatively simple trajectories in a low dimensional physiological parameter space of pressure and tension driving a highly nonlinear biomechanical device (e.g., [5–10]). This hypothesis has been tested in several species with very different types of songs, both spectral and temporal wise: rufous-collared sparrow (*Zonotrichia capensis*), canary (*Serinus canaria*), northern cardinal (*Cardinalis cardinalis*) brown thrashers (*Toxostoma rufum*) and zebra finch (*Taeniopygia guttata*). Moreover, the nonlinear nature of the sound source allows fundamental frequency modulation through pressure modulation. This fact was a prediction of the model and was tested experimentally in a suboscine species, the great kiskadee (*Pitangus sulphuratus*) [11]. Before this remarkable result, it was assumed that frequency and pressure were two physiological parameters controlled independently by the central nervous system. The nonlinear model allowed a more integrative perspective.

Similarly to human speech, birdsong is typically generated during expiration (see Fig. 1a). During song production a stereotyped pattern of expiratory air sac pressure pulses, alternate with brief, deep inspirations, called mini-breaths (shaded region in the lower panel of Fig. 1a). In this way, the respiratory activity determines the coarse temporal structure of the song (syllable sequence). Therefore, a model for birdsong production requires inspecting the dynamics of syringeal labia during those expiratory pulses.

In order to generate a dynamical model for the vocal source, we propose that the labia support two modes of vibration: an upward propagating wave and a lateral displacement around their midpoint positions, as suggested by videography of the folds during phonation [12]. These modes are coordinated in such a way that energy is gained from the airflow in each cycle, making auto-sustained oscillations possible. In order to do so, the labia present a convergent profile when moving away from each other, and a more planar profile during approaching. When the labia present a convergent profile, the average pressure between the labia is close to the air sac pressure. When the labia are approaching, the average pressure between them is similar to the atmospheric pressure. In this way, the pressure is high when the labia are moving away from each other, and low when the labia are approaching each other. This allows to overcome dissipative forces and to transfer energy from the airflow to the labia [4,5,13].

The dynamical description of this mechanism requires writing Newton's equations for a labium [4,5]. If x stands for the midpoint position of a labium of unitary mass, they read:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -k(x)x - \beta y - \gamma x^2 y + a_{av} p_{av} \end{aligned} \qquad (1)$$

where the first term of the second equation represents the (nonlinear) elastic restitution ($k(x) = k_1 + k_2 x^2$) [14], the second one the linear dissipation, the third one a nonlinear dissipation accounting for the existence of boundaries

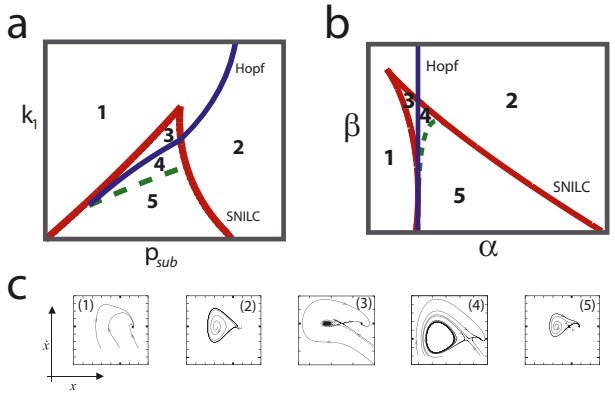

Fig. 2. Bifurcation diagram of birdsong production model. Qualitative differences in the dynamics are displayed in the numbered regions. The physical model (a) and its normal form (b) exhibit analogous bifurcation diagrams. The integration of the model for each region is shown in the numbered insets in (c). In this set of parameters, a Takens-Bogdanov bifurcation occurs, where a saddle-node bifurcation (brown line) is touched tangentially by a Hopf bifurcation (blue line) and a homoclinic bifurcation (dashed green line). A SNILC bifurcation occurs between region 2 to 5, and a Hopf bifurcation between regions 1 and 2.

for the oscillations, and p_{av} is the spatial average of the inter labial pressure. This last term is responsible for the energy transfer from the airflow to the labium and it can be written in terms of the sub syringeal pressure p_{sub}, the midpoint position of the labium x, its velocity y, the resting positions of the edges of the labium (x_{01} and x_{02}), and a parameter describing the time τ that takes the wave propagating upward in the labia to cover half its length:

$$p_{av} = p_{sub}(x_{01} - x_{02} + 2\tau y)/(x_{01} + x + \tau y). \quad (2)$$

These equations allow describing the dynamics of the sound source that can be explored in a two-parameter space. In Figure 2a we show the bifurcation diagram of the system moving parameters (p_{sub}, k_1). If the parameters' values are in region 1, only one attracting fixed point exists. This corresponds to non-oscillating labia, i.e., silent regime. In region 2 we find that the fixed point is unstable against a limit cycle. Oscillations of the labia correspond to sound production. Regions 3, 4, and 5, bounded by the saddle node curves (full brown line in Fig. 2) that converge to the cusp bifurcation point, have three fixed points. Between regions 2 and 5, a saddle node in a limit cycle (SNILC) bifurcation takes place, and between regions 1 and 2 a Hopf bifurcation (full red line in Fig. 2) [15]. The dashed green line indicates a homoclinic bifurcation.

This simple model captures the physiological mechanism of sound initiation: when the bird intends to vocalize, it increases air sac pressure to pass through a threshold in order to generate auto-sustained oscillations that would modulate the airflow generating sound (see recorded air sac pressure signal and sound wave in Fig. 1a). In the model, this could happen in two distinctive dynamical ways: through a Hopf or a SNILC bifurcation. The former generates tonal sounds: sinusoidal oscillations, born with a defined frequency and zero amplitude. The latter generates oscillations born with infinite period and spectrally rich: the saddle node remnant generates a spike-like wave, providing a substantial amount of energy to the harmonics. This rich variety of sounds covers a lot of the birdsong diversity.

The other physiological feature that is described by the model is that for increasing values of the parameter describing the lineal restitution (k_1), the oscillations will present higher fundamental frequencies. This parameter is related to the activity of syringeal muscles that increase the tension of syringeal labia. It has been shown that increasing muscle activity results in increasing values of fundamental frequency [16].

Another way of controlling the fundamental frequency is generating the oscillations through the SNILC bifurcation, where frequency values depend on the distance from the bifurcation point. In the model, this can be controlled not only by the restitution parameter but also by the pressure parameter, which provides a dynamical explanation for fundamental frequency modulation through air sac pressure [11,14,17].

In order to minimize the number of parameters involved in the description we performed a reduction of the model to a normal form: a minimal mathematical description of the dynamics capable of presenting the same bifurcation diagram as the original model [15,18]. We derived a standard equation which presents the bifurcation diagram of the physical model. It is a two dimensional dynamical system presenting a cusp and a Hopf bifurcation, obtained as a third order expansion around a Takens-Bogdanov linear singularity [18], which reads:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\alpha(t)\gamma^2 - \beta(t)\gamma^2 x - \gamma^2 x^3 - \gamma x^2 y \\ &\quad + \gamma^2 x^2 - \gamma x y. \end{aligned} \quad (3)$$

Comparing the bifurcation diagram of the physical model defined by equation (1) (see Fig. 2a) with that of the normal form defined by equation (3) (see Fig. 2b), a proper mapping of air sac pressure and tension into the unfolding parameters allows to recover the qualitative dynamics. In reference [19] the relationship was explicitly computed; the basic operations involved were (i) a translation of the values of (p_{sub}, k_1) where a Takens Bogdanov bifurcation takes place in the physical model to $(\alpha, \beta) = (0, 0)$, (ii) a multiplicative scaling, and (iii) a rotation of π. The normal form equations describe dynamical behavior in the regime of interactions between Hopf and a saddle node in limit cycle (SNILC) bifurcations. For zebra finches, most phonatory oscillations start in SNILC bifurcations, as the air sac pressure is increased (going from region 5 to region 2 in the parameter space shown in Fig. 2). Such dynamical constraints help to define the universe of zebra finch sounds. For example, there is a tight relationship between spectral content of a sound and its fundamental frequency: the lower the frequency, the higher the spectral content [20]. Small changes in parameters can produce radically different changes in sound output, depending on where those movements occur in the bifurcation space.

So far, we have described the sound source, which sits at the base of the vocal tract (see Fig. 1b). In order to reproduce realistic sounds, we modeled the trachea as a tube and then its output was used to excite the oro-esophageal cavity (OEC) which is modeled as a Helmholtz resonator [10,21], and its output will be our synthetic song.

In this way it is possible to synthesize songs with a low dimensional mathematical model whose parameters are easy to interpret in terms of physiological and anatomical observations. The time-dependent parameters can change smoothly in order to reproduce the acoustic properties of the recorded song that the model for song production attempts to describe. These time traces represent "motor coordinates", and they potentially provide a more natural bridge between the activity of the central nervous system and behavior than do sound spectrograms [8]. We tested this hypothesis by generating a synthetic copy of the recorded birdsong (e.g., Fig. 3). To go beyond the striking similarity between synthetic and natural songs, we used a highly accurate neurophysiological measurement in order to assess the biological relevance of the low dimensional model. We recorded neural responses to these auditory stimuli in song motor nuclei, demonstrating strong positive results [8].

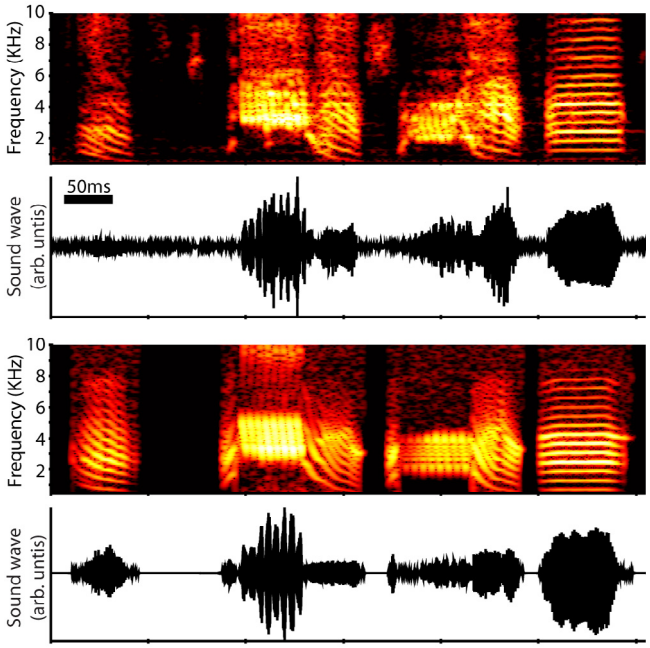

Fig. 3. The output of a low dimensional model for song production is able to reproduce a complex natural behavior. An algorithmic procedure allows to obtain synthetic copy of a zebra finch song by fitting two acoustic features of the recorded song. The top panel shows the sound wave and spectrogram of the recorded song and the lower panel corresponds to the synthetic song.

3 The nature of the instructions: the case of pressure for canary song

As we have discussed above, a central aspect of the motor control of birdsong production is the capacity to generate diverse respiratory rhythms, which determine the coarse temporal pattern of song. The neural mechanisms that underlie this diversity of respiratory gestures and the resulting acoustic syllables are largely unknown.

The neural network responsible for the generation of the gestures involved in birdsong has been well described. It includes two forebrain nuclei: HVC (used as a proper name) and the robust nucleus of the arcopallium (RA) that generates the coordinated motor patterns that help to shape the instructions driving the muscles controlling respiration, the vocal organ and upper vocal tract structures. However, it has been debated to what degree this telencephalic motor program contains direct instructions for detailed patterns such as the various timescales of the behavioral output [22]. One model, based in studies performed in zebra finches, proposes a direct control by the telencephalic song control area, such that all timescales present in the song arise directly from the output signal of this neural nucleus. Support for this model was derived from observations of sparsely coding output neurons in HVC as well as experiments in which song was stretched by cooling of HVC [23,24]. On the other hand, in the case of canaries, an interactive model has been proposed, where motor instructions emerge from nonlinear interaction between timescales of different components of the motor control network [7,25]. Two pieces of evidence point towards this integrative picture. One of them is the morphology of the pressure patterns used to generate the different syllables used in canary song. Figure 4a illustrates a typical pressure pattern during canary song production [9,26]. The different parts of this temporal pattern can be obtained as a minimally complex neural architecture that is driven by simple instructions. Namely, integrating the following dynamical system:

$$\begin{aligned} \dot{x} &= \lambda\,(-x + S(\rho_x + ax - by)), \\ \dot{y} &= \lambda\,(-y + S(\rho_y + cx - dy)), \\ \text{with} \qquad S(x) &= (1 + e^{-x/x_0})^{-1}, \end{aligned} \tag{4}$$

where x stands for the activity of a population of excitatory units, y for the activity of a population of inhibitory neurons, and $\rho_x(t)$, $\rho_y(t)$ for the inputs to those populations. The constants a, b, c and d describe the architecture of the array. The top panel of Figure 4b shows a caricature of this model, with interconnected nucleus of excitatory and inhibitory populations. A good intuition on the dynamics of this system can be gained by inspecting the different behavior displayed by this system under stationary values of the input parameters and ρ_x, ρ_y [27]. Figure 4b shows the bifurcation diagram of the system. Note the similarity between Figures 4b and 2.

In order to generate different patterns, the nature of the driving signal has to experience some changes, typically the amplitude and frequency of the forcing. In this way, different shapes in the time series can be generated as a unique neural architecture driven by slightly different instructions. An example of this is shown in Figure 4c where all the pressure patterns corresponding to a canary song

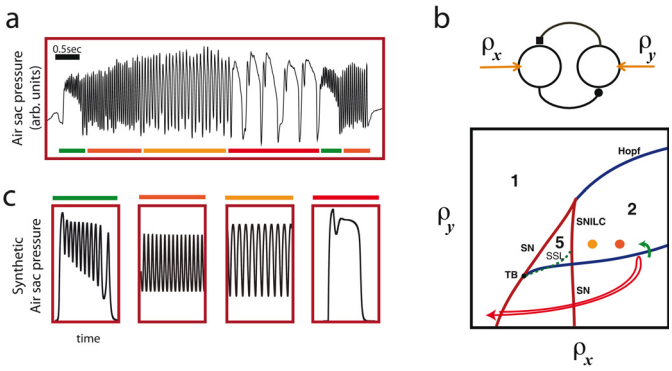

Fig. 4. Respiratory patterns during canary song. (a) Recorded air sac pressure during singing. The colored lines indicate different syllable types. (b) Diagram of neural nuclei of excitatory and inhibitory populations and the bifurcation diagram of the system. Note the similarity between this figure and Figures 2a and 2b (the numbers indicate the same dynamical regime of Fig. 2). (c) Synthetic pressure patterns obtained integrating the dynamical systems described by equation (4) with different values of (ρ_x, ρ_y), shown as colored dots and lines in the bifurcation diagram of (b).

are generated using the model described by equation (4), using different combinations of (ρ_x, ρ_y). The colored dots and lines in the bifurcation diagram of Figure 4b correspond to a particular pressure pattern in Figure 4c.

Since any neural architecture is intrinsically nonlinear, even slight changes in the driving instructions can lead to qualitatively different patterns. In this way, different patterns have been interpreted as sub-harmonic solutions of simple instructions [7,9].

The other evidence suggesting the integration of different time scales in the generation of the pressure patterns used in canary song comes from the results obtained as one of the time scales is manipulated [25]. It has been well established that the thermal manipulation of HVC affects the timing of the song [24]. This experiment showed that at least one of the pertinent time scales could be perturbed by thermally tempering with that telencephalic nucleus.

A simple integrative model, which can also reproduce the shapes of the pressure temporal patterns, predicts that a small change in the frequency of an average instruction coming from HVC may provoke a drastic change in the temporal pattern of resulting syllables. This facilitates a control mechanism where simple neural instructions interacting with downstream neural architecture can result in complex rhythms in the output patterns. Manipulation of the telencephalic timescale through local cooling results in the predicted effects of initial stretching and then "breaking" of syllables when the cooling range is extended. These novel "syllable-breaking" patterns can be interpreted in terms of bifurcations of the model [25].

In the paradigm of nonlinear interaction, HVC, or a circuit including HVC, does play an important role in the generation of the respiratory patterns. This paradigm challenges the idea of a "*look up table*" linking bursts of HVC activity with brief segments of song, with HVC bursts determining a unique timescale in the birdsong system[22]. We do not propose a specific location in the neural architecture of the songbird system for the second timescale necessary to explain the shape and rate of the measured pressure patterns and their breaking. Instead, we show that a minimal computational model (two neural populations: interconnected excitatory and inhibitory neural nuclei) is capable of transducing very simple instructions into the specific patterns measured in canaries.

4 The nature of the instructions: the case of tension in finches

The initial computational models of birdsong production suggested that very simple gestures for pressure and labial tension could account for many of the features found in song. For example, the nature of the frequency modulation was thought to be controlled by the phase difference between those gestures, which could be as simple as harmonic fluctuations [5]. The actual measurement of the air sac pressure in canaries unveiled a subtler scenario: different syllables required different pressure patterns. Yet, the diversity of gestures was not a capricious collection of shapes: they can be obtained as the different solutions presented when a minimal neural architecture is driven by simple driving patterns [9,25]. Is this also the case for the patterns corresponding to the tension?

This issue is slightly more difficult to test. The smooth time series that is measured by a pressure transducer connected through a cannula to an air sac has no parallel when it comes to estimating labial tension. The songbird syrinx is a complex anatomical structure composed by cartilaginous rings and half-rings, muscles and labia. Six pairs of syringeal muscles provide mechanical control in the syrinx, four of which have both insertion sites on the syrinx (intrinsic muscles), being attached to the cartilaginous structures. The phonating labia are attached to the internal part of modified cartilaginous half-rings. In this way, specific syringeal muscles would contract to move the cartilage that would ultimately modify the properties of the labia. Moreover, very little is known about the histology or elastic properties of the labia in songbirds [28]. The closer one gets to measure a proxy for the labial tension is to measure the electrical activity of the syringeal muscles innervating the syrinx. Unfortunately, the relationship between the EMG of the different syringeal muscles and the actual physiological parameters in the model is still not completely unveiled. In other words, there is no unique relationship between EMG activity of individual syringeal muscles and the parameters controlled by them. Moreover, it is not yet solved how the different muscles synergistically control the acoustic properties.

Despite these difficulties, it was recently shown that a reconstruction of the gestures necessary to drive a simple biomechanical model to generate realistic song was plausible. In order to validate the reconstruction procedure two strategies were followed. First, the reconstructed air sac pressure was compared with the experimentally recorded

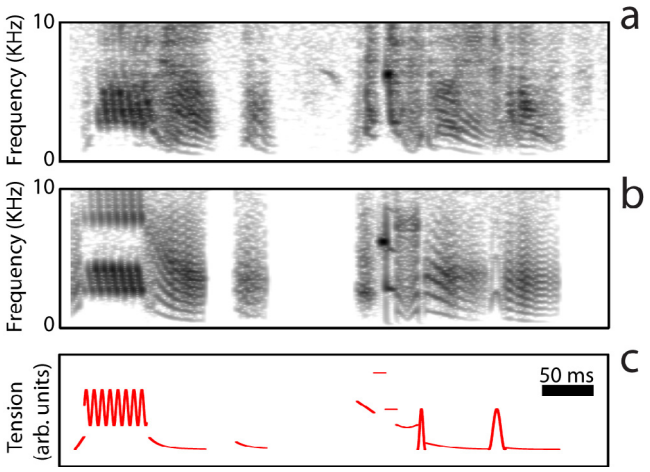

Fig. 5. Simple motor gestures generate complex behavior. Recorded zebra finch song (a) is well reproduced by the output of a low dimensional model (b) when the parameters are in the vicinity of bifurcations. Panel (c) shows the temporal evolution of one of the two parameters of equation (3), namely, $\beta(t)$ that can be related to the tension of the syringeal labia.

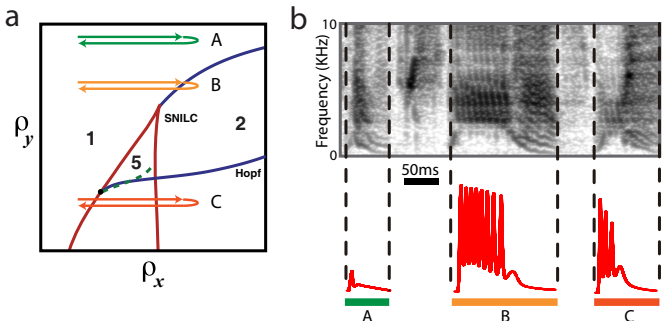

Fig. 6. Tension of the syringeal labia of zebra finches generated by a simple neural architecture. (a) Bifurcation diagram of the dynamical system described by equation (4). Simple parameter explorations near bifurcation regions generate a diversity of tension gestures that could be used to generate synthetic song (b). The colored lines and numbers show the correspondence between the exploration in parameter space and the temporal evolution of the system's output.

one, obtaining a high correlation between the time traces [10]. More spectacularly, synthetic song produced with the model, when driven by the reconstructed physiological gestures, was played to sleeping birds while the activity of neurons selective to the bird's own song was recorded [8]. Despite the extraordinary selectivity of these neurons to the bird's own song [29,30], the synthetic song was capable of eliciting responses. Thereby, we performed electrophysiological experiments in order to obtain a highly accurate biological measurement to validate the birdsong production model.

Assuming the biomechanical model described in Section 2 (a model capable of starting oscillatory behavior through a SNILC bifurcation), we created synthetic versions of the songs that our test birds sang. Time-dependent parameters of the model describing the labial dynamics were reconstructed to account for the time-dependent acoustic properties of the sound. We used an algorithmic procedure to reconstruct unique functions for the air-sac pressure ($\alpha(t)$) and the tension of syringeal labia ($\beta(t)$) [10]. The result of the procedure for one song is illustrated in Figure 5, showing that many features observed in the spectrogram of the recorded song (Fig. 5a) were also present in the synthesized song (Fig. 5b). Remarkably, relatively simple time traces of reconstructed pressure and tension arose from fitting the bird's song (see Fig. 5c for an example of reconstructed tension). In fact, song was described by the sequence of these pressure-tension trajectories, which we call gestures, with gesture onsets and offsets defined as discontinuities in either the pressure or tension functions.

For the songs reported in reference [8], we found that most of the parameters could be approximated well by fractions of sine functions, exponential decays, constants or a combinations of these. Most remarkably, those patterns can also be found when a unique neural architecture is driven by essentially the same simple instruction: a kick followed by a slow decay. In Figure 6a we display the bifurcation diagram of a low dimensional dynamical system ruling the average dynamics of a population of excitatory units coupled to a population of inhibitory ones (the same system described with Eq. (4)). The lines represent bifurcation curves: curves in parameter space that separate parameter regions where the system presents qualitatively different behaviors. The three curves ending in arrows represent trajectories in the parameter space of the driving of excitatory and inhibitory populations (ρ_x, ρ_y). The lower panel of Figure 6b shows the solutions generated by the neural network under the driving showed by the curves A, B and C in Figure 6a. No quantitative calculation is really needed in order to recognize that the solutions correspond to the reconstructed patterns (as the one shown in Fig. 5c). The same neural architecture (an excitatory and an inhibitory population of neurons), and similar temporal patterns driving it (kicks with exponential decays) were necessary in order to generate the different tension patterns.

5 A common theme: variability around linear degeneracies

There is a hierarchy of complexity in this phenomenon. The sounds are complex indeed, with rich and subtle relationships between different acoustic properties like fundamental frequency and spectral content. Indeed, most of the acoustic features of birdsong come in package. The origin of this is the nonlinear nature of the vocal organ, which bounds the spectral relationships between the uttered sounds. Moreover, these do not depend on the details of the biomechanics, but on the basic family ("normal form") of dynamical system the model of the sound source belongs to. Because of this, it is quite straightforward to generate a diversity of songs using the same biomechanical model, even modeling songs belonging to birds of very distant groups (e.g. oscines and suboscines).

Moreover, many acoustic features are found both in birds and humans, which share the basic biomechanical principles of the sound sources, but details of musculature and structure are indeed very different [17,28].

There is structure at the level of the instructions driving the syrinx itself. In canaries, the different pressure patterns used to utter different syllables can be obtained by driving a basic neural architecture with very simple time series. It is remarkable that those simple driving instructions are transduced into specific patterns: small oscillations mounted on a DC level, period solutions followed by a period doubled ones, or larger pulses consisting of rapid oscillations followed by a slowly decaying pulse (see Fig. 4c). Even more surprising: those patterns deform as predicted by the dynamical model when one of the time scales involved is modified using an experimental setup to control temperature [25].

To complete the picture, in zebra finches, for which the other key physiological parameter driving the syrinx has been reconstructed from theoretical means, the labial tension can also be generated as the same basic neural architecture is driven by very simple time series. The exponential decays, the rapid increases, the oscillations followed by exponential decays, the constant frequencies followed by a small decaying tail: all the features reconstructed in the process of attempting to synthesize realistic sounds, could be found by driving a simple neural architecture.

These three instances point to a unique strategy: the transduction of somewhat simple time series into richer ones. Probably high in the song system (where high is taken as a synonym of distant to the peripheral vocal device) the instructions are coded in very simple terms, and richness is gained as the instructions are integrated into other systems. After all, the bird first breathes, then vocalizes, and finally learns to sing. It is therefore likely that the final outcome of the songbird's vocal system is the integration of those systems.

In the three scales described in this manuscript, the diversity of gestures can be achieved easily, due to the fact that the system was operating in the neighborhood of a linear singularity [31]. From the point of view of the modeling, this allows us to change rapidly between different solutions. Interesting enough, the regions in parameter space that converge in a linear singularity are open regions of the parameter space. This allows us to have both the benefits of flexibility and robustness, two of the most remarkable ingredients of behavior.

6 Conclusions

In this paper we analyzed the generation of birdsong from a dynamical perspective. Much of the complexity of the sounds could be tracked to the nonlinear nature of the avian vocal organ. Specifically, many acoustic features ended up emerging in "packages" controlled by the underlying dynamics of the biomechanics. A typical example is the relationship between spectral content and fundamental frequency in the song of the zebra finch. Furthermore, nonlinear dynamics seems to play a role beyond the periphery: in the canary song, the diversity of syllables requires respiratory patterns that are sub-harmonic responses of a simple dynamical system driven by simple instructions. Suggestively, the dynamical system can be trivially interpreted as the one ruling the dynamics of a simple neural architecture, overwhelmingly present in the neural system needed for birdsong production. Remarkably, the same dynamical system, driven by simple time series, also reproduces the tension gestures that have been reconstructed in the song of zebra finches.

We are not yet capable of identifying where in the song system (i.e., the specific neural architecture needed to generate song), but the idea that the physiological instructions emerge out of the interaction between different levels of complexity is already a breakthrough. Birds have developed song, building on a pre-existing system: the respiratory one. Even non-learners have developed the ability to coordinate their respiratory activity with the proper morphological modifications of the air pathways, controlling constrictions that allowed the modulation of airflow and therefore, the production of song. Oscine birds, which constitute approximately 40 percent of the known bird species, have built on top of this structure a sophisticated set of neural nuclei that help to learn the physiological patterns controlling the syrinx and the respiration. The integrated circuit then is a complicated neural structure that includes brainstem nuclei that project to the neurons in charge of controlling respiration and syringeal muscle, as well as to the telencephalic nuclei which integrate these signals with auditory and sensory information, to finally project back to the brainstem. This integrated anatomy is consistent with our description of progressively complex instructions interacting nonlinearly, and challenge a widespread view of telencephalic nuclei (HVC) in charge of a look up table where all the behavior is coded in brief (approximately 10 ms) commands. The bias towards stressing the role of telencephalic nuclei in the system is also a result of the relative simplicity to perform neural recordings in this region, compared to other parts of the song system. It is one of the great challenges in the field to overcome this difficulty, in order to unveil the relative role of different parts of the song system in the generation and learning of this complex and rich behavior.

We would like to thank Santiago Boari for comments to this manuscript. This work describes research partially funded by CONICET, ANCyT, UBA and NIH through RO1-DC-012859.

References

1. A.J. Doupe, P.K. Kuhl, Ann. Rev. Neurosci. **22**, 567 (1999)
2. H.P. Zeigler, P. Marler, *Neuroscience of Birdsong* (Cambridge University Press, Cambridge, 2012)
3. R. Mooney, Learn. Memory **16**, 655 (2009)
4. G.B. Mindlin, R. Laje, *The Physics of Birdsong* (Springer-Verlag, Berlin, 2005)
5. T. Gardner, G. Cecchi, M. Magnasco, R. Laje, G.B. Mindlin, Phys. Rev. Lett. **87**, 208101 (2001)
6. G.B. Mindlin, T.J. Gardner, F. Goller, R. Suthers, Phys. Rev. E **68**, 41908 (2003)

7. M.A. Trevisan, G.B. Mindlin, F. Goller, Phys. Rev. Lett. **96**, 58103 (2006)
8. A. Amador, Y. Sanz Perl, G.B. Mindlin, D. Margoliash, Nature **495**, 59 (2013)
9. L.M. Alonso, J.A. Alliende, F. Goller, G.B. Mindlin, Phys. Rev. E **79**, 41929 (2009)
10. Y.S. Perl, E.M. Arneodo, A. Amador, F. Goller, G.B. Mindlin, Phys. Rev. E **84**, 051909 (2011)
11. A. Amador, F. Goller, G.B. Mindlin, J. Neurophysiol. **99**, 2383 (2008)
12. F. Goller, O.N. Larsen, Proc. Natl. Acad. Sci. USA **94**, 14787 (1997)
13. I.R. Titze, J. Acoust. Soc. Am. **83**, 1536 (1988)
14. A. Amador, G.B. Mindlin, Chaos **18**, 043123 (2008)
15. S.H. Strogatz, *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry and Engineering* (Perseus Books, Cambridge, 1994)
16. F. Goller, R.A. Suthers, J. Neurophysiol. **76**, 287 (1996)
17. A. Amador, D. Margoliash, J. Neurosci. **33**, 11136 (2013)
18. J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, 1997)
19. J.D. Sitt, E.M. Arneodo, F. Goller, G.B. Mindlin, Phys. Rev. E **81**, 31927 (2010)
20. J.D. Sitt, A. Amador, F. Goller, G.B. Mindlin, Phys. Rev. E **78**, 011905 (2008)
21. N.H. Fletcher, T. Riede, R.A. Suthers, J. Acoust. Soc. Am. **119**, 1005 (2006)
22. T.W. Troyer, Nature **495**, 56 (2013)
23. R.H.R. Hahnloser, A.A. Kozhevnikov, M.S. Fee, Nature **419**, 65 (2002)
24. M.A. Long, M.S. Fee, Nature **456**, 189 (2008)
25. M.A. Goldin, L.M. Alonso, J.A. Alliende, F. Goller, G.B. Mindlin, PLoS One **8**, e67814 (2013)
26. R.S. Hartley, R.A. Suthers, J. Neurobiol. **21**, 1236 (1990)
27. F.C. Hoppensteadt, E.M. Izhikevich, *Weakly Connected Neural Networks* (Springer, 1997)
28. T. Riede, F. Goller, Brain Lang **115**, 69 (2010)
29. D. Margoliash, J. Neurosci. **6**, 1643 (1986)
30. D. Margoliash, M. Konishi, Proc. Natl. Acad. Sci. USA **82**, 5997 (1985)
31. M.O. Magnasco, O. Piro, G.A. Cecchi, Phys. Rev. Lett. **102**, 258102 (2009)