

This article was downloaded by: [186.137.54.25]

On: 15 August 2013, At: 06:13

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Contemporary Physics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tcph20>

### The physics of birdsong production

G.B. Mindlin<sup>a</sup>

<sup>a</sup> Department of Physics , FCEyN, University of Buenos Aires , Argentina

Published online: 22 Jul 2013.

To cite this article: Contemporary Physics (2013): The physics of birdsong production, Contemporary Physics, DOI: 10.1080/00107514.2013.810852

To link to this article: <http://dx.doi.org/10.1080/00107514.2013.810852>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

## The physics of birdsong production

G.B. Mindlin\*

*Department of Physics, FCEyN, University of Buenos Aires, Argentina*

*(Received 27 April 2013; final version received 20 May 2013)*

Human babies need to learn how to talk. The need of a tutor to achieve acceptable vocalisations is a feature that we share with a few species in the animal kingdom. Among those are Songbirds, which account for nearly half of the known bird species. For that reason, Songbirds have become an ideal animal model to study how a brain reconfigures itself during the process of learning a complex task. In the last few years, neuroscientists have invested important resources in order to unveil the neural architecture involved in birdsong production and learning. Yet, behaviour emerges from the interaction between a nervous system, a peripheral biomechanical architecture and environment, and therefore its study should be just as integrated. In particular, the physical study of the avian vocal organ can help to elucidate which features found in the song of birds are under direct control of specific neural instructions and which emerge from the biomechanics involved in its generation. This work describes recent advances in the study of the physics of birdsong production.

**Keywords:** birdsong; nonlinear dynamics; biomechanics; bifurcations; complex sounds

### 1. No whistles

When a human imitates a bird, he or she whistles. Actually, the way in which birds vocalise is more similar to the way in which humans produce voiced sounds (i.e. those that involve the modulation of an airflow by oscillating vocal folds, like vowels). In Figure 1 the basic mechanism of birdsong production is described. As the bird exhales, airflow is established through the avian vocal organ, called the syrinx. At the junctures between the bronchi and the tract, there are two pairs of labia. Under the right conditions, the airflow induces oscillations in those labia, generating an acoustic wave. As the wave passes through the trachea and the oroesophageal cavity, it modifies its harmonic content, and eventually emerges as a sound signal. Endoscopic images of the intact songbird syrinx during spontaneous vocalisations were necessary to establish this paradigm [1].

Beyond that description it is important to establish what those ‘right conditions’ are. Moreover, what are the physiological parameters that the bird controls in order to modify the acoustic features of the vocalisations during birdsong production? Biologists and physicists have addressed this issue by complementary approaches.

One can start a modelling effort for birdsong production writing equations for the separation between the labia. Treating a labium as a mass in a spring, we can relate the stiffness of the latter with the labial tension, which the bird controls by tightening or relaxing syringeal muscles. A second important parameter in the problem is the air sac

pressure, which has to overcome a threshold value in order to set the labia into an oscillatory mode.

Let us assume 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 [1] (see the inset displaying the labia in Figure 1). These modes are coordinated in such a way that energy is gained from the airflow in each cycle, making sustained oscillations possible. In order to do so, the labia present a convergent profile when moving away from each other, and a divergent profile during approach to each other. When the labia present a convergent profile, the average pressure between them is close to the air sac pressure. When the labia present a divergent profile, the average pressure between the labia is similar to 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 them to overcome dissipative forces and to transfer energy from the airflow to the labia.

The dynamical description of this mechanism requires writing Newton’s equations for a labium [2,3]. If  $x$  stands for the midpoint position of a labium of unitary mass, then:

$$\begin{aligned}\frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= -k(x)x - \beta y - \gamma x^2 y + a_{av} p_{av},\end{aligned}$$

\*Email: gabo.mindlin@gmail.com

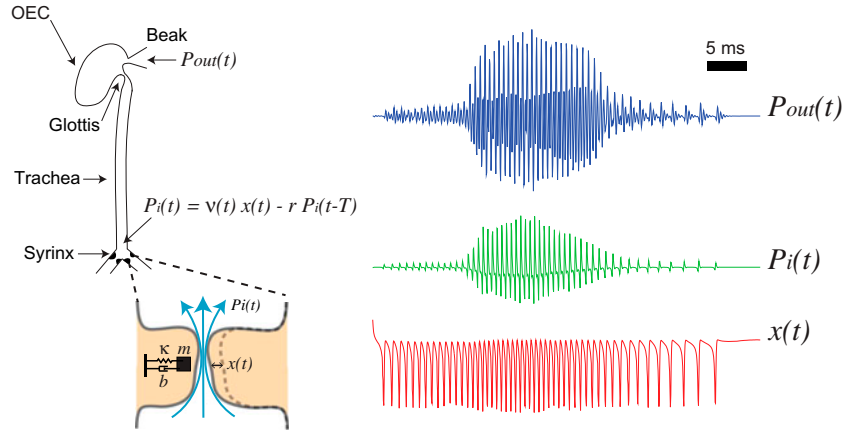


Figure 1. Schematic diagram of the avian vocal organ. Two sources generate sound waves that travel across the trachea and the oro-esophageal cavity (OEC). Each sound source is modelled as a small mass attached to a string, and subjected to inter-glottal air pressure. The labial dynamics determines the modulation of the airflow. The simulated labial position, pressure at the input of the trachea, and output pressure are illustrated (right).

where the first term of the second equation represents the (nonlinear) elastic restitution (3), the second one the linear dissipation, the third one a nonlinear dissipation accounting for the existence of boundaries 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. 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 that describes the time it takes the wave propagating upward in the labia to cover half its length ( $\tau$ ):

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

With these equations it is possible to describe the dynamics of the sound source. For values of the air sac pressure high enough, the labium presents oscillations. For increasing values of the parameters describing the restitution, the oscillations will present higher fundamental frequencies. Moreover, the values of  $x$  can be used to model the fluctuating flow velocity at the entrance of the tract, which can be modelled as a tube. Then, the output of the tube can be used to excite the oro-esophageal cavity, modelled as a Helmholtz resonator [4], and its output will be our model for song. In this way it is possible to synthesise songs with a low dimensional mathematical model whose parameters are easy to interpret in terms of physiological and anatomical observations. In order to synthesise a birdsong syllable, we start in a region of the parameter space where no oscillations occur. If we move our parameters within the phonating region, the labia will start to modulate the airflow, and sound will be generated. During the phonating interval the bird might change its parameters in order to alter the acoustic features of the vocalised sound. For example, it can increase the tension of the labia, while remaining in the phonating region. The end of the vocalisation will occur

as the parameters are moved back to the non-phonating region of the parameter space. Typically there are two very different time scales in the problem: the modulations in the physiological parameters representing the tension of the labia and the pressure occur in the time scale of the syllable (typically between 50 and 100 ms). The labia, on the other hand, will oscillate at a much faster time scale (typical oscillation frequencies being in the order of kHz).

## 2. Bifurcations and sound

Qualitative changes in the dynamics observed when the parameters of a nonlinear system are varied are called bifurcations. In the framework of our problem, it is pertinent to investigate the way in which stationary solutions (corresponding to non-phonating situations) might start to oscillate. Figure 2 displays regions in the parameter space where qualitatively different dynamics take place. Actually, the two ubiquitous ways to turn oscillations on in bi-dimensional dynamical systems are present in this bifurcation diagram: a Hopf bifurcation, for high values of the restitution, and a saddle node in a limit cycle (5).

In the first bifurcation, oscillations are born with zero amplitude and well-defined frequency. For that reason oscillations born in this bifurcation can be used to synthesise tonal sounds. They were in fact used to study canary songs, which are characterised by spectrally pure notes with few overtones. By varying the pressure and tension parameters in time, within the region of the parameter space where Hopf bifurcations take place, it is possible to reproduce the starts, stops, timbre, and continuous changes in pitch of canary song (2). But not every species uses only tonal sounds. The most widely studied songbird, the *Zebra finch*, alternates tonal sounds with syllables that are spectrally very rich; sounds that are made up of many equally spaced harmonics. Moreover, there is a very precise relationship between the

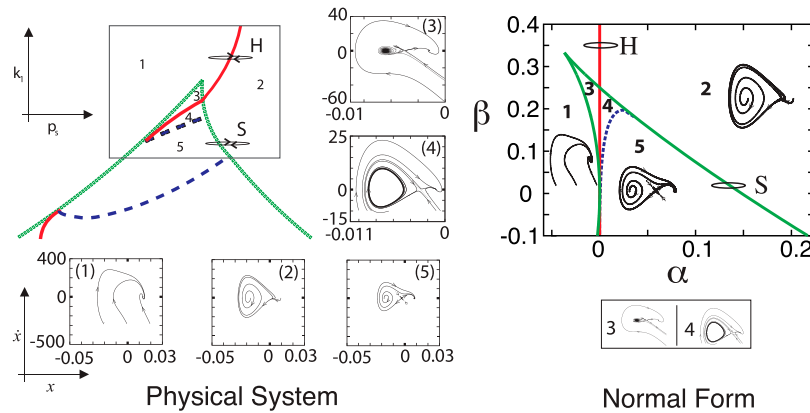


Figure 2. Bifurcation diagrams. For different values of the parameters, the physical model ruling labial dynamics presents qualitatively different regimes (left). A simplified dynamical system, called normal form, presents the same scenarios (right). The panels correspond to the solutions, in phase space, that the models present for parameters in each of the numbered regions. Oscillations occur for sufficiently high values of the pressure. The generation of syllables occurs whenever the parameters drive the system into the regions where oscillations take place. In that way, the labia will oscillate, the airflow will be modulated, and sound will be produced. The lines separating the regions of the parameter space are called bifurcation curves. Qualitatively different ways of starting oscillations exist. In a Hopf bifurcation, an oscillation with a well-defined frequency and zero amplitude is born. In a SNILC (saddle node in limit cycle) bifurcation, an oscillation with zero frequency and finite amplitude is born.

harmonic content and fundamental frequency of the uttered sounds; low frequency sounds are harmonically very rich, while the higher frequency sounds are more tonal. That relationship was investigated in many syllables uttered by different birds, and the functional form relating those acoustic features was the same for all of them [5]. Remarkably, it is precisely what can be expected if a periodic signal is born in a saddle node in a limit cycle bifurcation. In this mechanism, when the periodic oscillation is born, the phase space is left with the ghost of the two fixed points that were annihilated in the bifurcation, which slows down the passing by trajectory. The simplest dynamical system presenting this effect is

$$\frac{d\theta}{dt} = \varpi - \sin \theta,$$

with  $\theta$  a phase. The stationary solutions of this system exist as long as  $\varpi < 1$ . For  $\varpi > 1$ , periodic oscillations exist, but notice that for  $\varpi = 1 + \epsilon$ , the solutions will present a very small phase velocity as  $\theta \approx \pi/2$ . The way in which the period of the oscillations diverge as  $\epsilon \rightarrow 0$  can be analytically computed [6], and this functional form is precisely what is recovered in the vocalisations of the Zebra finch.

The model presented for the labial motion was studied in detail [3]. The equations were integrated for different values of the control parameters (strength of the restitution, air sac pressure), and regions of the parameter space were identified for which the solutions presented qualitatively similar dynamics. The boundaries between those regions are called bifurcation lines. It is possible to simplify the original system of equations, which emerges from physical considerations, to a simpler one capable of displaying the same dynamical regimes. The technique is known as *normal*

*form reduction*, and for the region of the parameter space relevant for our problem leads to [7]:

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

where the parameters  $\alpha$ ,  $\beta$  can be mapped to the air sac pressure and syringeal tension, respectively, and  $\gamma$  is a time constant that allows one to obtain oscillations of the desired spectral content and fundamental frequency for the zebra finch vocalisations [7]. This reduction not only replaces a complicated equation by a simpler one: the parameters involved have been reduced to a minimum. In Figure 2 (right) we display the different solutions that the reduced system can display for different values of the parameters. They are equivalent to the ones present in the original (physical) model and displayed in Figure 2 (left). Paths in parameter space  $(\alpha(t), \beta(t))$  that start in a non-phonating region, visit the phonating region, and return to a non-phonating zone will allow us to synthesise syllables. Remarkably, the trajectories in parameter space can be extremely simple, but the generated sounds will present delicate relationships between different acoustic features that are neither determined by the details of the physiological instructions, nor by the details of the physical model, but by the underlying dynamics.

### 3. What do birds think of all this?

Beyond this sophisticated mathematical analysis, are the synthetic sounds good enough? How can we test the pertinence of this low-dimensional reduction? The process of building confidence on this route of modelling has been

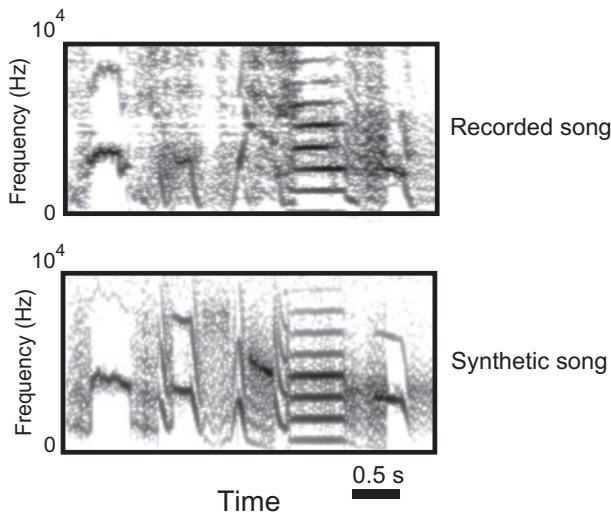


Figure 3. Recorded and synthetic song. From a recorded song (upper panel), it is possible to reconstruct the parameters that would allow the model to generate sounds with similar pitch and spectral content (lower panel).

slow. The first suggestion that even with minimally complex gestures of pressure and tension one could generate a variety of different sounds was proposed in [2]. The first validation involving experimental data required the simultaneous measurement of song, air sac pressure, and tension of different syringeal muscles. In [8] we performed this first test, synthesising song by feeding a low-dimensional model for labial dynamics with physiological variables, acting as time depending parameters ( $k(t)$ ,  $p_{\text{sub}}(t)$ ). The synthetic songs generated with the model were similar to the recorded songs. That conceptual exercise was repeated for different species (including non-songbirds [9]) and for models with the capacity to account for different sounds [10]. In [10], a procedure was described to reconstruct the parameters representing air sac pressure and syringeal tension from the song, and the reconstructed parameters were compared with the measured data. But still the key issue was whether the tests designed to compare the synthetic sounds generated with the model were ‘good enough’. How to define a measure of success? The sonograms look similar, as displayed in Figure 3, and the songs sounded good to us, the researchers working on the problem, but what about for the birds? It was necessary to find a way to ask the birds. And first, it was necessary to ask neuroscientists for the language in which the question had to be formulated.

Dan Margoliash and his group held the key to address this issue. He had previously reported that when a sleeping bird listens to a recording of its own song, neurons in a certain ‘premotor’ part of the brain produce bursts of activity in the same pattern as they did when the bird sang the song [11]. The effect is specific to the bird’s own song (BOS): when a bird listens to another bird’s song, or to its own song played backwards, those neurons don’t fire in the same pattern. In fact, they don’t fire at all. Despite all Zebra

finch songs sharing some acoustic features, and regardless of each individual having learned from a specific tutor, each individual develops a unique song, and these highly sensitive neurons respond, during sleep, to that song only. Besides its intrinsic importance in the process of birdsong learning, this effect provided the perfect opportunity to test the model: would these highly selective neurons respond to synthesis of the song in the same way they respond to the bird’s own song? Ana Amador, working at Margoliash’s lab, performed those experiments. From recorded songs, the motor parameters were reconstructed (see Figure 4). Then, those reconstructed motor gestures were used to drive the low-dimensional model for the labial motion, and its solutions were used to estimate the sound waves generated by the vocal organ [10]. Finally, the synthetic song was played to the sleeping bird whose song was being fitted [12].

In the experimental protocol, each song was played 20 times to the bird while the activity of a selective neuron was recorded. For a given neuron, the BOS would elicit a response at a specific time when the BOS was played. When the synthetic song was played, without changing the location of the electrodes, the recorded neuron would either spike at the same temporal location within the song, or not spike at all. On average, the neurons responded to synthetic songs about 60% of the time. In the protocol, the same neurons were measured also when variations of the synthetic songs were played to the sleeping bird. For

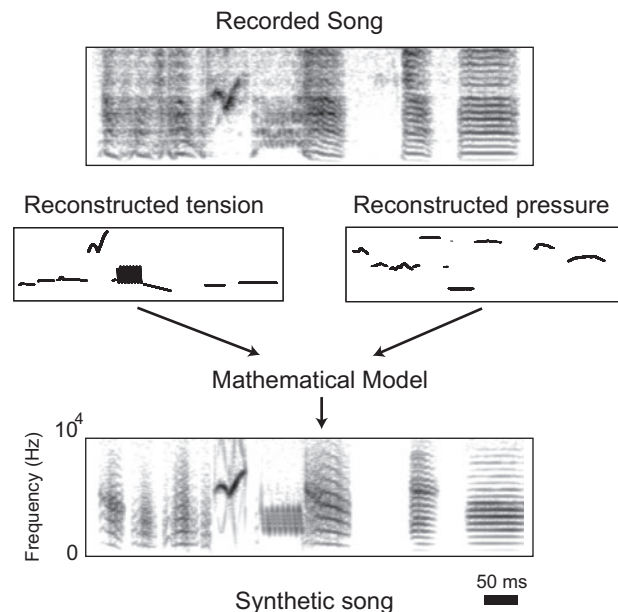


Figure 4. Reconstructed instructions. A recorded song is divided into successive time windows. The parameters of the normal form that allow synthesising sounds with similar pitch and spectral content are reconstructed (middle panels). Then, the reconstructed instructions are fed into the mathematical equations describing labial dynamics. Finally, the synthetic sounds are played to the sleeping bird whose song was modelled in the first place.

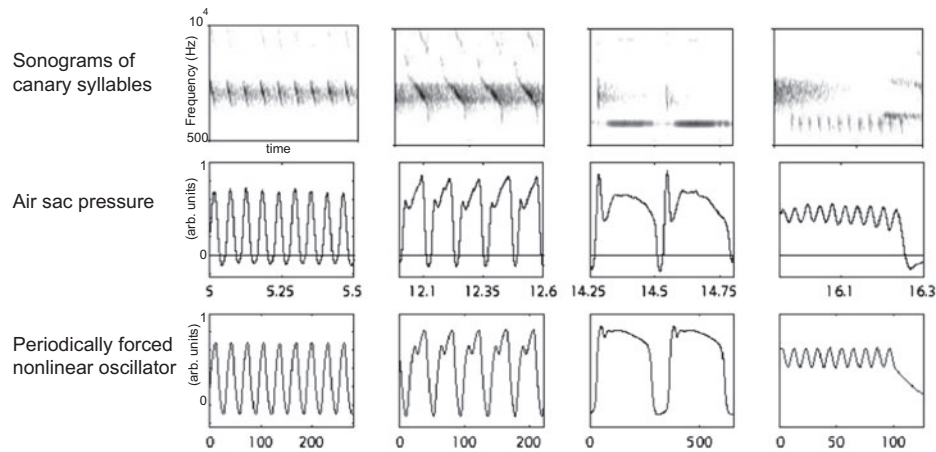


Figure 5. Low-dimensional model for canary air sac pressure. In the case of the domestic canary, the patterns used to generate different syllables are shapes near and dear to nonlinear dynamicists: they are the sub-harmonic responses of a periodically forced nonlinear oscillator. The top panels represent four sonograms of typical canary syllables. Air sac pressure patterns recorded during the production of the sounds are shown in the middle panels. Different solutions of a periodically forced neural oscillator are displayed in the bottom panels.

example, we generated different songs, changing the dissipation parameter in the equations describing the effect of the Helmholtz resonator used to model the oro-esophageal cavity. We found that the response of the neurons decreased dramatically.

#### 4. The instructions

The research described in the previous sections showed that much of the acoustic richness of birdsong could be reproduced when relatively simple physiological instructions operated the nonlinear avian vocal organ. Yet, non-linear dynamics might provide some interesting perspectives at the moment of unveiling the mechanisms behind the generation of the instructions themselves. A well-described network of interconnected brain nuclei generates these motor patterns, which control the syringeal muscles, the upper vocal tract and the respiration [13]. Despite continuous efforts, the role played by the different neural nuclei remains elusive.

The domestic canary (*Serinus canaria*) offers an interesting clue. In Figure 5 we can see the sonograms corresponding to different syllables uttered by a canary (top four panels). Below each syllable, the pressure pattern used for its production is displayed. In order to perform those measurements, a plastic cannula connected to a pressure transducer is inserted into the thoracic air sac [14]. Despite their seemingly arbitrary nature, those patterns have a long history in the field of non-linear dynamics: these are the time series that are expected whenever a nonlinear system presenting a characteristic frequency is periodically forced with different forcing frequencies. As long as the forcing frequency to the natural time scale of the driven system is similar, 1:1 locking is expected to occur. As the forcing frequency is changed, other sub-harmonic responses are obtained [15]. The four panels at the bottom of Figure 5 display the solutions of a mathematical model describing

the dynamics of a forced neural oscillator, for different values of the forcing frequency. It is premature to conjecture where these two time scales are generated in the neural architecture of the canary, although it is plausible that they reflect the interaction between telencephalic instructions and respiratory activity [15,16]. Yet, it is remarkable that the final output of an extremely complex neural architecture, consisting of thousands of highly nonlinear dynamical units such as neurons, is that of a low-dimensional non-linear dynamical system. The emergence of non-trivial, yet low-dimensional dynamics out of coupled non-linear units is an open problem in statistical physics, and birdsong is an ideal field for exploring it.

#### 5. Conclusions

All aspects of behaviour involve instructions generated at the neural level, the response of a peripheral device to both those instructions and the environment, and feedback from the periphery to the nervous system. For that reason it is pertinent to explore the biomechanics involved in a complex behaviour, as part of the strategy to understand the phenomenon. In the case of birdsong, the avian vocal organ is a highly nonlinear device. It is for that reason that it is particularly important to unveil which features require the explicit coordination of delicate instructions, and which will naturally emerge from the nonlinear nature of the biomechanics involved.

It was surprising that a low-dimensional dynamical system for the labial motion, driven by a couple of time-dependent parameters could allow us to synthesise songs that passed the test of highly selective neurons in the brain of songbirds. This dimensionality reduction might also play a role in the process of elucidating how the neural system underlying the behaviour operates. Relating acoustic features to patterns of neural activity might be complex for the

simple reason that the dimensionality of the acoustic description is very high. It might be the case that reconstructing motor gestures ends up providing a system of motor coordinates that allows a more direct interpretation of the neural activity patterns. The observation that bursts of activity in the pre-motor nucleus HVC were locked to significant instances of the motor coordinates [12] suggests that the strategy is worth exploring.

### Acknowledgements

I would like to thank Ana Amador, Yonatan Sanz Perl and Rodrigo Alonso for their comments and criticism. This work describes research partially funded by CONICET, ANCyT and UBA.

### References

- [1] F. Goller and O.N. Larsen, *A new mechanism of sound generation in songbirds*, PNAS 94 (1997), pp. 14787–14791.
- [2] T. Gardner, G. Cecchi, M.O. Magnasco, R. Laje, and G.B. Mindlin, *Simple motor gestures for birdsongs*, Phys. Rev. Lett. 87 (2001), 208101.
- [3] A. Amador and G.B. Mindlin, *Beyond harmonic sounds in a simple model for birdsong production*, Chaos 18 (2008), 043123.
- [4] Y.S. Perl, E.M. Arneodo, A. Amador, and G.B. Mindlin, *Nonlinear dynamics and the synthesis of Zebra finch song*, Internat. J. Bifur. Chaos 22 (2012), 1250235.
- [5] J.D. Sitt, A. Amador, F. Goller, and G.B. Mindlin, *Dynamical origin of spectrally rich vocalizations in birdsong*, Phys. Rev. E 78 (2008), 011905.
- [6] S.H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*, Perseus Books, New York, 2001.
- [7] J.D. Sitt, E.M. Arneodo, F. Goller, and G.B. Mindlin, *Physiologically driven avian vocal synthesizer*, Phys. Rev. E 8 (2010), 031927.
- [8] G.B. Mindlin, T.J. Gardner, F. Goller, and R. Suthers, *Experimental support for a model of birdsong production*, Phys. Rev. E 68 (2003), 041908.
- [9] A. Amador, F. Goller, and G.B. Mindlin, *Frequency modulation during song in a suboscine does not require vocal muscles*, J. Neurophysiol. 99 (2008), pp. 2383–2389.
- [10] 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 (2011), 051909.
- [11] D. Margoliash and M. Konishi, *Auditory representation of autogenous song in the song system of white-crowned sparrows*, PNAS 82 (1985), pp. 5997–6000.
- [12] A. Amador, Y.S. Perl, G.B. Mindlin, and D. Margoliash, *Elemental gesture dynamics are encoded by song premotor cortical neurons*, Nature 495 (2013), pp. 59–64.
- [13] H.P. Zeigler and P. Marler, *Neuroscience of Birdsong*, Cambridge University Press, Cambridge, 2008.
- [14] R.A. Suthers, F. Goller, and R.S. Hartley, *Motor dynamics of song production by mimic thrushes*, J. Neurobiol. 25 (1994), pp. 917–936.
- [15] M.A. Trevisan, G.B. Mindlin, and F. Goller, *Nonlinear model predicts diverse respiratory patterns of birdsong*, Phys. Rev. Lett. 96 (2006), 058103.
- [16] J.M. Méndez, G.B. Mindlin, and F. Goller, *Interaction between telencephalic signals and respiratory dynamics in songbirds*, J. Neurophysiol. 107 (2012), pp. 2971–2983.