

Avian vocal production beyond low dimensional models

This content has been downloaded from IOPscience. Please scroll down to see the full text.

J. Stat. Mech. (2017) 024005

(<http://iopscience.iop.org/1742-5468/2017/2/024005>)

[View the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 157.92.4.71

This content was downloaded on 18/04/2017 at 15:33

Please note that [terms and conditions apply](#).

You may also be interested in:

[Depth-kymography of vocal fold vibrations](#)

Frits F M de Mul, Nibu A George, Qingjun Qiu et al.

[Experimental study of self-oscillations of the trachea–larynx tract by laser doppler vibrometry](#)

G Buccheri, E De Lauro, S De Martino et al.

[On the single-mass model of the vocal folds](#)

M S Howe and R S McGowan

[The physicist's guide to the orchestra](#)

Jean-Marc Bonard

[A lightweight feedback-controlled microdrive for chronic neural recordings](#)

A Jovalekic, S Cavé-Lopez, A Canopoli et al.

[Aero-thermoacoustics \(The Generation of Sound by Turbulence and by Heat Processes\)](#)

O K Mawardi

[Acoustic tube model of the human vocal tract: formants and vowels](#)

Herbert M Urbassek

[Numerical simulation of the influence of the orifice aperture on the flow around a teeth-shaped obstacle](#)

J Cisonni, K Nozaki, A Van Hirtum et al.

[Discrete-time modelling of musical instruments](#)

Vesa Välimäki, Jyri Pakarinen, Cumhur Erkut et al.

Special Issue on Statphys 26

Avian vocal production beyond low dimensional models

Gabriel B Mindlin

Physics Department, University of Buenos Aires, Buenos Aires, Argentina
E-mail: gabo@df.uba.ar

Received 22 October 2016, revised 13 December 2016

Accepted for publication 18 December 2016

Published 16 February 2017



Online at stacks.iop.org/JSTAT/2017/024005
[doi:10.1088/1742-5468/aa54d8](https://doi.org/10.1088/1742-5468/aa54d8)

Abstract. Birdsong is an active field of research in neuroscience, since songbirds learn their songs through a process similar to that followed by humans during vocal learning. Moreover, many of the vocalizations produced by birds are quite complex. Since the avian vocal organ is nonlinear, it is sensible to explore how much of that complexity is due to the neural instructions controlling the vocal organ, and how much to its nonlinear nature. In this work we first review some of the work carried out in the last years to address this problem, and then we discuss the existence of noisy sound sources in the avian vocal organ. We show that some spectral features of the song produced by the *Zebra finch* (one of the most widely studied species) can only be explained when vortex sound is taken into account.

Keyword: computational biology

J. Stat. Mech. (2017) 024005

Contents

1. Introduction	2
2. Methods	3
2.1. Sound sources I: mass injection	3
2.2. Low dimensional models for oscillating tissue	3
2.3. The one mass model	4
2.4. The two mass model	5
2.5. The flapping simplification	5
2.6. The dynamical simplification	7
2.7. Birdsong: measuring the parameters and testing for salience	8
2.8. Filters	9
3. Sound sources II: vortex sound	10
3.1. The effect of vortex sound in timbre	12
4. Conclusions	13
Acknowledgments	14
References	14

1. Introduction

Behavior emerges from the interaction between a nervous system, a peripheral system, and its environment [1]. That is certainly the case in vocal production, where many vertebrates generate highly complex communication signals through sound generating mechanisms involving different parts of the motor system. Interesting enough, species evolutionary distant as humans and songbirds produce their communication sounds through the same basic processes, engaging the motor systems in charge of controlling respiration, the vocal organ and the upper vocal tract [2, 3] (Titze 1994). In the case of birdsong production, respiratory gestures generate airflow, which induces connective tissue masses to oscillate. Muscles innervating the holders of these tissues can reconfigure them. This affects the tissues themselves, modifying some of the spectral features of the modulated airflow. This is finally filtered by the upper vocal tract, which can also be dynamically adjusted. This conceptually simple scenario presents one caveat: the dynamics of the vocal organ, and in particular the dynamics of the oscillating tissue is highly nonlinear. For this reason, it is difficult to predict, even knowing the activity of the motor systems involved in phonation, which acoustic features are to be expected in the output sound. *It is difficult to unveil which of the sound features are explicitly coded in the nervous system, and which emerge as eventually simple physiological instructions operating on the nonlinear vocal organ.* This is one of the reasons for which computational modeling is instructive in the field of phonation.

The modulation of airflow by oscillating tissue is not the only way to generate sound that vertebrates use. Many interesting and subtle phenomena are engaged in the process of sound production. Yet, the modeling in the case of oscillating tissue is where progress has been more significant in the last decades. Despite its complexity, low dimensional models (i.e. models involving a low number of variables, whose time evolution is ruled by a small number of equations) have been implemented and tested for their perceptual salience [15]. In this work we review this aspect of the problem, and move beyond it to describe the noisy aspects of birdsong production.

2. Methods

2.1. Sound sources I: mass injection

Figure 1 displays the vocal phonating device of oscine birds (a set of approximately four thousand species that share with humans the need of a tutor in order to learn the species-specific vocalizations). At the juncture of the bronchi and the trachea, pairs of labia are capable of modulating the airflow. Therefore, during phonation, there is a periodic injection of air into the trachea. The rate of mass injection for unit of volume q can be written in terms of the air velocity v , density ρ and the lumen's area A as

$$q = \rho A v.$$

The dynamics of the density perturbations induced by this mass injection, at the base of the trachea, will be given by

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

with c_0 the sound velocity [18]. Therefore, it is the labial dynamics that is responsible for the time dependence of the lumen's area A , and therefore acts as a sound source in the linear approximation of the problem. The lumen's area will be the product of a transverse, constant length, and a variable whose dynamics describes the labial motion. In the following section we describe models proposed to account for the labial oscillations.

2.2. Low dimensional models for oscillating tissue

We can start a model for oscillating tissues by representing them as ‘blocks’ attached to springs, which are introduced to portray the elastic nature of the tissues [4] (Flanagan and Ishizaka 1973). In this way, the problem of modeling the self sustained oscillations that can give rise to phonation is reduced to solving two specific issues: the nature of the additional forces on the masses capable overcoming the dissipation, and the effects responsible for bounding the motion that is established when the new forces overcome the losses.

One way to compensate the energy losses of a mass attached to a spring in the presence of viscosity is by applying a force whose direction changes together with the direction of the velocity. This is the opposite strategy followed by the dissipation in

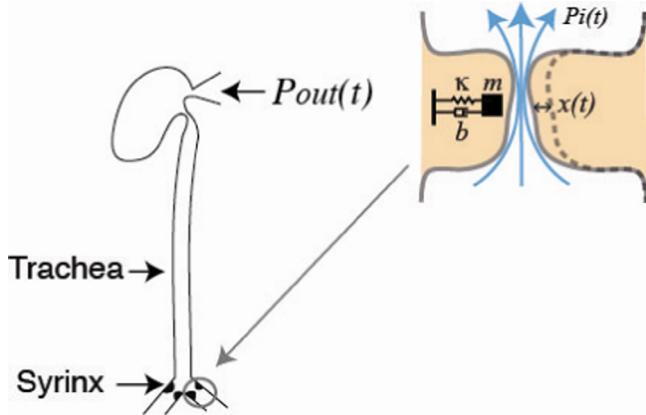


Figure 1. A schematic of the avian vocal organ is shown. At the junctures of the bronchi and the trachea, two pairs of labia modulate the airflow injected into the trachea. The two sound sources can be independently controlled. The main parameters in the model are the labial tension and the air sac pressure.

its attempt to stop an oscillation. In the framework of our problem, the responsibility of carrying out this task lies on the inter-glottal pressure. In this way, one of the first attempts to build a computational model for phonation was based on designing a configuration presenting an inter-glottal pressure that was higher when the masses were departing from each other than when they were approaching.

2.3. The one mass model

In the one mass model (Titze 1994), a simple block of mass m attached to a spring, represents the oscillating tissue (the actual phonating device consists of two mirror copies of this configuration). The airflow, after passing through the space between the opposite masses, enters the trachea. In this model, the tube plays a key role, through the inertia of the air in it. As the masses depart from each other, the volume flow increases, and due to the inertia of the air in the tube, so does the inter-glottal pressure. A similar argument can be used to explain suction when the masses approach each other. In this way, the delayed response of the vocal tract is the key element for generating force acting in the same direction as the velocity.

This model was historically important. The first conceptual models for phonation were mostly based on the action of the Bernoulli pressure, which by itself is not sensitive to the direction of motion of the tissue (and therefore not capable of overcoming energy losses). On the other hand, the one mass model assumes restrictive conditions on the plausible oscillating frequencies of the device, since the mass of the tube has to behave as a single mass element. That is only a good approximation as long as the masses oscillate at low frequencies. Our next model involves a mechanism that is independent of a vocal tract. As it is usually the case, abandoning a model implies giving up some simplifying hypothesis. We will need to relax the hypothesis that the oscillating tissue behaves as a simple mass, whose unique degree of freedom is a lateral displacement.

2.4. The two mass model

Direct observations have shown that vocal folds in humans, or labia in birds, do not behave as rigid masses. They are made of tissue capable of sustaining both lateral displacements and wavelike motions of the cover layers. Ishizaka and Flanagan were the first to model this internal structure of the oscillating element in terms of two coupled masses [4]. Many versions of this conceptual model have been proposed in the literature, but they all share a search of balance between mathematical simplicity and a sensible description of the diversity of physical phenomena shaping the force on the oscillating tissue [5, 6]. In terms of the language presented in the previous section, a computational implementation of this model starts by writing Newton's equations for the displacements from the equilibrium positions x_1 and x_2 of two masses m_1 and m_2 :

$$\begin{aligned}\frac{dx_i}{dt} &= y_i \\ \frac{dy_i}{dt} &= \frac{1}{m_i}(f_i - K(x_i) - B(x_i, y_i) - k_c(x_i - x_j)),\end{aligned}$$

for $i, j = 1$ or 2 for the lower and upper masses respectively. The functions K and B represent the restitution and damping of the tissue, and f stands for the hydrodynamic force on the masses, and k_c is a constant describing the coupling strength between the masses. In the particular functional forms for these functions is where the art of the modeler resides. These functions are typically defined through different functional forms depending on the values of the variables. This allows accounting for the qualitative changes in stiffness and dissipation that occur when the masses collide, or the change of the hydrodynamic force depending on whether the masses present a convergent or divergent profile. A number of researchers have, over the years, iterated this model that constitutes a very sound equilibrium between a realistic description of the physics involved and the mathematical simplicity that comes when a problem is formulated in terms of a few ordinary differential equations [12].

The numerical integration of these models has allowed us to gain intuition on the mechanisms behind sound production. In order to be able to oscillate, the lower and upper masses must present a convergent profile when moving away from their mirror companions, and a divergent profile while approaching them. The physical reason is the following: when the labia present a convergent profile, the average pressure between them is close to the pressure at the bottom of the system. When the masses present a divergent profile, the average pressure between the set of opposite masses is similar to the atmospheric pressure. In this way, the pressure is high when the opposite masses are moving away from each other, and low when they are approaching. This allows to overcome dissipative forces and to transfer energy from the airflow to the masses.

2.5. The flapping simplification

The main achievement of the two mass model is to capture the simultaneous existence of two modes: a lateral displacement and the flapping. In a seminal work, Titze managed to design a low dimensional model whose only dynamical variable

was the labial midpoint position (x), and yet it could capture the necessary interplay between the two modes [7]. He proposed a kinematic restriction on the labia, such that the area between the labia along the vertical direction would be compatible with the existence of a flapping. More specifically, calling a_1 and a_2 half the separations between the lower and upper labial edges respectively, he proposed the following kinematics:

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

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

where a_{10} , a_{20} describe the configuration of the labia at rest, and τ the travel time of across the labia of the upwards wave. In this way, it is guaranteed that the upper half separation decreases as the lower one increases. This kinematics also implies that the labia present a convergent profile as they separate from each other, and a divergent one as they approach. Once this dynamics is imposed, it is possible to compute the average pressure between the labia:

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

where p_{sub} stands for the sub-labial pressure. This enables us to write Newton's equations in terms for the variable describing the mid point position of a labium:

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

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

where it is assumed a restitution due to labial elasticity, a linear dissipation, and a nonlinear one accounting for energy loses that occur either when the labia reach the walls or at labial collision.

This model does not address why or under which conditions these two modes would lose stability; it simply assumes that kinematics and finds it compatible with a self-sustained oscillating dynamics. Interesting enough, the highly nonlinear nature of this model allows finding a rich set of dynamical responses. The plot displays a bi-dimensional diagram, where the axes correspond to two of the parameters in the problem (the restitution constant and the sub labial pressure). The lines separate regions of the parameter space where different dynamics can be found. The kind of dynamics is indicated by the insets, which correspond to sets of trajectories to be found for different initial conditions, if the dynamical system is numerically integrated for a pair of parameter values within the region where the inset is displayed. Being a 2D model, the dynamics is never more complex than oscillatory. A two mass model is capable of displaying more complex asymptotic behavior, even chaotic one.

2.6. The dynamical simplification

Given the tremendous simplifications assumed by the different models, it is legitimate to wonder why they work at all. It is difficult to accept that no other forces are present, that the restitution can be approximated by a linear function, or that the nature of the nonlinear dissipation is really cubic, just to name a randomly selected, brief list of questions. The explanation is within the lessons learned in the last decades from nonlinear dynamics. There are algorithmic and systematic procedures that allow finding nonlinear changes of coordinates that convert the nonlinear system under study into a simpler one. Those procedures constitute what is known as ‘normal form reduction’. In this way, many different nonlinear systems get mapped into the same simpler system, which acts as a ‘model for models’. This is at the core of the success of many simple models: they might be mapped into the same simpler system that other sophisticated and detailed ones.

As an example, the equations describing the dynamics of the flapping model can be converted into

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

where α and β represent unfolding parameters, and γ a time scale constant. This dynamical system presents simpler nonlinear terms than the ratio of polynomials that was necessary to describe the average pressure between the labia [8]. As displayed in figure 2, the simpler system presents similar partitions of the parameter space, where qualitatively similar dynamics occurs.

Notice that both in the parameter space portrait of the complete physical model, and in the one corresponding to the normal form, one can transition from a region of the parameter space where the attracting solution is a fixed point (labia at rest), to another one in which the solution is a limit cycle (the labia oscillate, and therefore modulate the airflow, thus generating sound). In fact, there are different ways in which oscillations can be born. Transitioning from region 1 to region 2, an oscillation is born with a well-defined frequency and zero amplitude. This bifurcation is known as a Hopf bifurcation. On the other hand, transitioning from region 5 to region 2, the oscillation is born with a well-defined finite amplitude and zero frequency. This bifurcation is known as a ‘saddle node in a limit cycle’. In this process, low frequency solutions are spectrally rich, and in fact, as the parameters are moved, there is a specific relationship between the spectral content of a solution and its fundamental frequency. That relationship was verified in zebra finches, where the spectral content and fundamental frequencies of segments from different syllables uttered by different birds were computed. The relationship between these acoustic indexes was compared against the one expected if the oscillation was assumed to be born in a saddle node in a limit cycle [8]. Any model including this generic bifurcation would have been able to reproduce this acoustic feature: it is deeply rooted into the underlying dynamics, and does not depend on the details of the model.

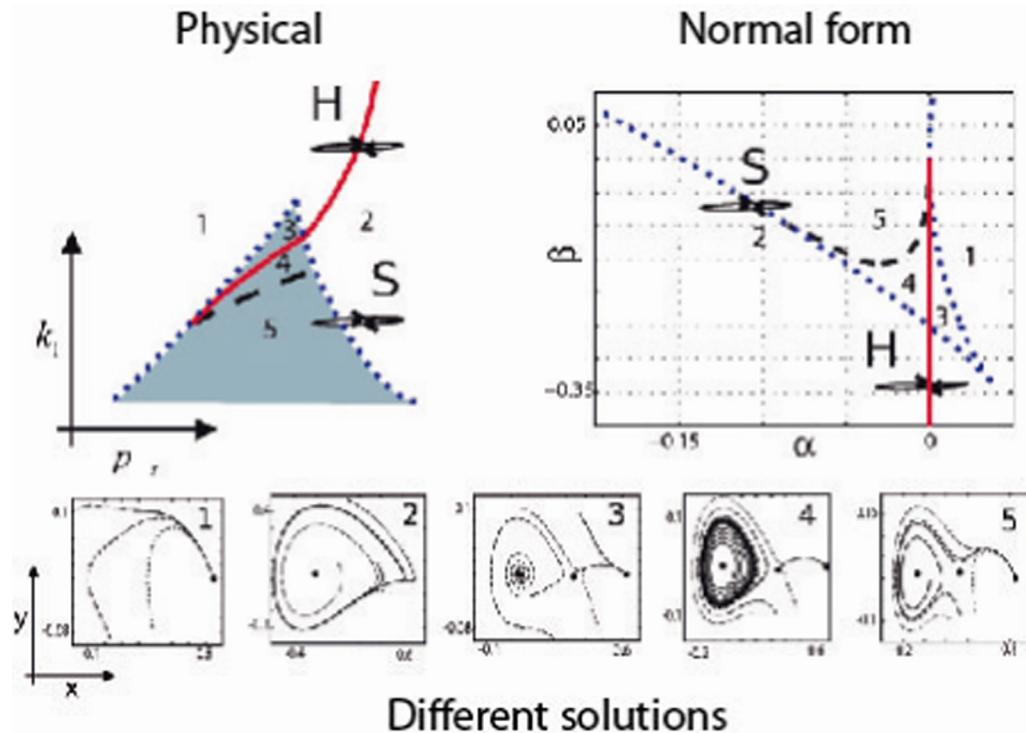


Figure 2. The lines separate the parameter space of a model into regions. Within each region we find parameters that lead to solutions that are qualitatively similar. Crossing the lines implies that the solutions undergo a qualitative change that is called a bifurcation. The physical model (top, left) and the normal form model (top, right) present similar regions. The solutions to be found in the regions are displayed in the bottom panels.

2.7. Birdsong: measuring the parameters and testing for salience

The first suggestion that with minimally complex gestures of pressure and tension one could generate a variety of different sounds was proposed in [16]. The first validation involving experimental data required the simultaneous measurement of song, air sac pressure, and tension of different syringeal muscles. Those measurements were plausible, and in fact, Suthers, Goller and others had already carried out a long-term program to unveil the parameters in charge of controlling the peripheral avian vocal organ in songbirds [9, 10]. The first test consisted in synthesizing song by feeding a low dimensional model for labial dynamics with physiological variables that were time depending parameters of the model [11]. The synthetic songs generated with the model were similar to the recorded songs, as it displayed in figure 3. That conceptual exercise was repeated for different species, including non-songbirds [12], and for models with the capacity to account for different sounds [13].

Beyond the similarity between the song and the synthetic sound at the level of the sonogram, pertinence of the model was explored in electrophysiological experiments. There are neurons in a nucleus of the song system of songbirds (HVC) that respond selectively to the bird's own song (BOS), i.e. they do not respond to the song of conspecific birds, reverse song; just to BOS [14]. Yet, these neurons would

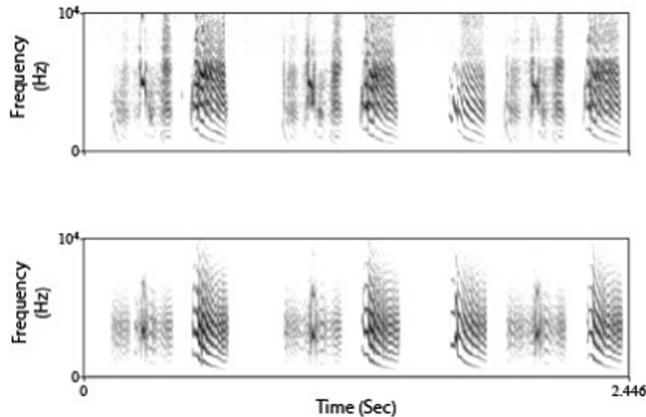


Figure 3. Sonograms of a natural song by a Zebra finch (top) and of a synthetic song obtained integrating a normal form with time varying parameters.

respond similarly to BOS and to surrogate songs generated with a low dimensional model [15].

2.8. Filters

So far, we have analyzed the airflow modulation as the source of sound. But sound is not a delayed copy of that signal. As the pressure fluctuations travel throughout the rest of the vocal apparatus in its way to the atmosphere, a series of partial reflections add up to the originally injected signal and the result is a time trace whose spectrum reflects the geometry of the path travelled by the sound. If the tract can be approximated by a series of tubes, and the frequency is such that in each of the tubes the sound wave can be approximated as a plane wave, we can analyze the process as follows. From the dynamics of the source, we compute the flow and its time derivative. A linear combination of those (with coefficients that depend on the tract and the way it couples with the source) represents the pressure fluctuations injected into the first tube. Let us call this function $s(t)$. Then, if the pressure at the input of the tract is denoted by $p_i(t)$, the pressure reflected back at the end of the tube by $p_{\text{back}}(t)$, the sound velocity by c , the reflection coefficient by γ and the tube length is L , then

$$p_i(t) = s(t) + p_{\text{back}}(t - L/c)$$

$$p_{\text{back}}(t) = -\gamma p_i(t - L/c).$$

A similar procedure can be carried out with subsequent tubes, where the injected signal is the transmitted pressure wave [16].

A remarkable feature in birdsong is the rapid modulation of the fundamental frequency. Yet, the nature of the upper vocal tract does play a role in songbirds [17]. In the electrophysiological tests on selective neurons in HVC, where the responses elicited by surrogates generated with synthetic models were compared with those elicited by the BOS, the presence of a filter including the oroesophageal cavity was essential in order to obtain significant responses [15].

3. Sound sources II: vortex sound

Human speech alternates voiced sounds with unvoiced ones. In the first family we include the vowels, and the physics involved in their production resembles the mechanisms described in the previous sections, with the vocal folds playing the role of the labia in birds. Unvoiced sounds (those not produced by the modulation of airflow by self sustained tissue vibrations) involve a different physics, studied in an area known as aeroacoustics. Some bird species, like the Zebra finch, also alternate sounds with regular oscillations with extremely noisy ones. In this section we present for the first time a discussion on the contribution of aeroacoustic sounds to birdsong production.

When the flow crosses the lumen and is injected into the trachea, it becomes separated from the walls, forming a jet. This is a focused, high momentum region surrounded by stagnant air. Between these regions, the air particles undergo a rotation, what is characterized quantitatively by the vorticity of the velocity field. In this way, the shear layers between the jet and the walls coalesce into irregular coherent structures traveling along the trachea, until they reach the glottis. At this point, a new acoustic mechanism has to be taken into account: *sound production by vorticity past an obstacle* [19, 21].

Howe (1998) derives an expression for the acoustic pressure fluctuation p generated when a vortex passes through a constriction:

$$\frac{D^2 p}{Dt^2} - c_0^2 \nabla^2 p = \rho_0 \nabla \cdot (\vec{\omega} \times \vec{v}),$$

where D stands for the material derivative, ρ_0 for the ambient undisturbed density, and $\vec{\omega}$ and \vec{v} for the vorticity and local velocity respectively. Notice that the vector product involves quadratic contributions of the velocity. Therefore, this equation can be interpreted as the second order correction to the homogeneous wave equation studied in linear acoustics. In this way, the effects discussed here should be added to the pressure fluctuations eventually generated by mass injection.

The solution of this equation can be found in an integral form:

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

where A is the cross section at the observation point, M the Mach number, and \vec{U}^* stands for the ideal flow velocity field that would exist if the duct contained a uniform steady flow [21]. The brackets denote that the integration is carried out at the retarded time $t_{\text{ret}} = t - |x - y|/c_0(1 + M)$.

The previous integral can be very difficult to compute, but a qualitative picture can be sketched. The vector product in the integral $\vec{\omega} \times \vec{v}$, computed in a regime such that a vortex travels in the axial direction, points in the radial direction. On the other hand, the unperturbed flow \vec{U} , in a set up with a constriction, has a negative radial component before the vortex passes the constriction, and a positive one once the vortex has passed it. In our problem, the glottis is a constriction at the upper end of the trachea (see figure 1).

We can write a phenomenological expression for the temporal dependence of the pressure fluctuation originated as a vortex passes through the constriction. In what follows, $t = 0$ corresponds to the precise instant at which the vortex passes through the constriction. For times outside the range $[-H/U_c, H/U_c]$ (with U_c the convection speed of the vortex through the constriction), the scalar product defining the value of the pressure fluctuation will be zero (outside this time interval, the vortex is far away from the constriction and therefore $(\vec{\omega} \times \vec{v}) \cdot \vec{U}^* \approx 0$). Only as t enters the interval above it is that there will be non zero pressure values. For negative times within this time interval, there will be a depressurization. As the vortex leaves the constriction behind, there will be a positive pressurization. As t grows away from the interval $[-H/U_c, H/U_c]$, again $(\vec{\omega} \times \vec{v}) \cdot \vec{U}^* \approx 0$. In this way, if α is the angle between the pipe axis (the direction of \vec{v}) and the field \vec{U}^* , we model the source S for an acoustic pressure pulse due to the convection of *one* vortex as

$$W(t) = \zeta U_c \sin(\alpha(t))$$

with

$$\alpha = \begin{cases} 0, & t \in (-\infty, -H/(2U_c)) \\ 2\pi(U_c/H)t, & t \in (-H/(2U_c), H/(2U_c)) \\ 0, & t \in (H/(2U_c), \infty) \end{cases}$$

where ζ depends on the geometry of the system, U_c is the flow speed at the constriction and H is its characteristic length. The strength of this vortex (measured in terms of the volume integral of its vorticity, i.e. its circulation Γ) is assumed to be unitary. The actual computation of the amplitude ζ can be found in [21].

To account for the effect of a train of N vortex rings, we have to compute the convolution of the source term generated by one vortex with an arrival function $I(t)$, consisting of a series of delta functions at the arrival times T_{arrival} , with amplitudes that reflect the strength of the arriving vortex, i.e. their circulation. In this way,

$$S(y, t) = \int_{-\infty}^{\infty} W(\tau) I(t - \tau) d\tau$$

where

$$I(t) = \sum_{n=0}^{n=N} \delta(t - nT_{\text{arrival}})\Gamma_n.$$

Notice that the circulation of the n th vortex Γ_n is proportional to the difference between the arrival times of the n th and $(n - 1)$ th vortex. This is because a large time between two consecutive arrivals means that the second vortex was generated far from the constriction, having gained circulation in its travel towards it. In our simulations, vortexes arrive whenever a uniform random variable (taking values between zero and one) exceeds a threshold (threshold = 0.999) but it should not be ruled out that in a confined jet as the one involved in this problem, the backflow in the recirculation might play a role destabilizing the jet. Therefore, a lower dimensional description might be plausible [22].

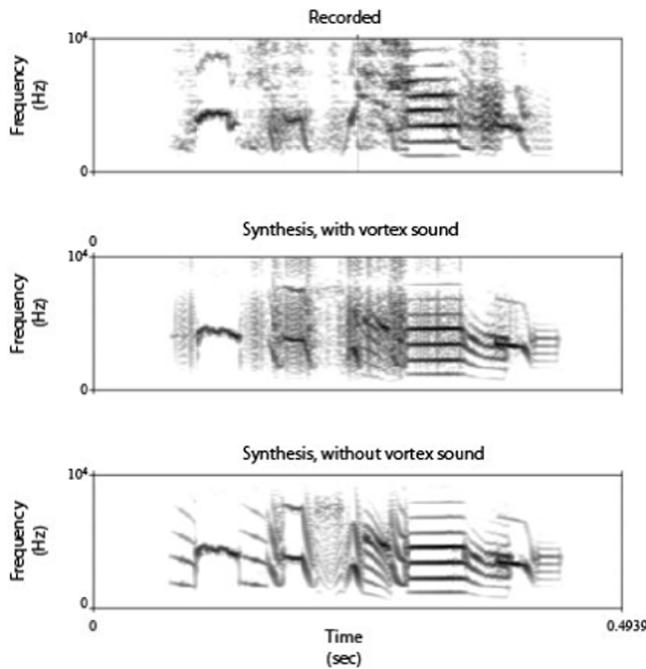


Figure 4. As a vortex traveling along the trachea (left to right) passes through the constriction, the scalar product $\vec{\omega} \times \vec{v} \cdot \vec{U}$ changes sign (top). Therefore, the pressure pulse detected at a distant position x is a sine-like pattern with a time scale defined by $H/2U_c$. A sequence of vortex sound pulses.

3.1. The effect of vortex sound in timbre

Zebra finch song presents complex features at two levels. The methods described and developed in the previous sections allow explaining the relationship between fundamental frequencies and higher harmonics in terms of the bifurcations involved in the onset of labial oscillations. But we are also presenting in section 3 a method for generating noise in a way which is consistent with the known anatomy of birdsong. In this section we illustrate the effect of taking this sound source into account.

This noisy component adds a characteristic roughness to the timbre of the Zebra finch song. In fact, some Zebra finches incorporate at least one syllable in their repertoire where it is difficult to track a fundamental frequency, as the first sound in the song displayed in the first panel of figure 4. In that figure we also display a simulation where the sound is generated with the two mechanisms described in the text (second panel), as well as using only the mass injection mechanism (third panel). In some time intervals, there is phonation achieved exclusively with vortex sound (see the first sound in the first panel, which corresponds to a recorded song, and compare it with its synthesis in the second panel). Figure 5 shows the spectra of the three songs. Notice that the vortex sound, responsible for the spectral features at high frequencies, is capable of reproducing features as the lack of spectral components within specific ranges (see the arrows in figure 5). Those features require specific modeling of the basic pulse, and carry information on the shape of the constriction. In our simulations, we used a sine function as described above, and the duration of the pulse was estimated in 0.2 ms. The selection of this value followed the inspection of the spectrum of the song to be modeled, which is shown in figure 5(a). The duration of the vortex sound pulse was chosen

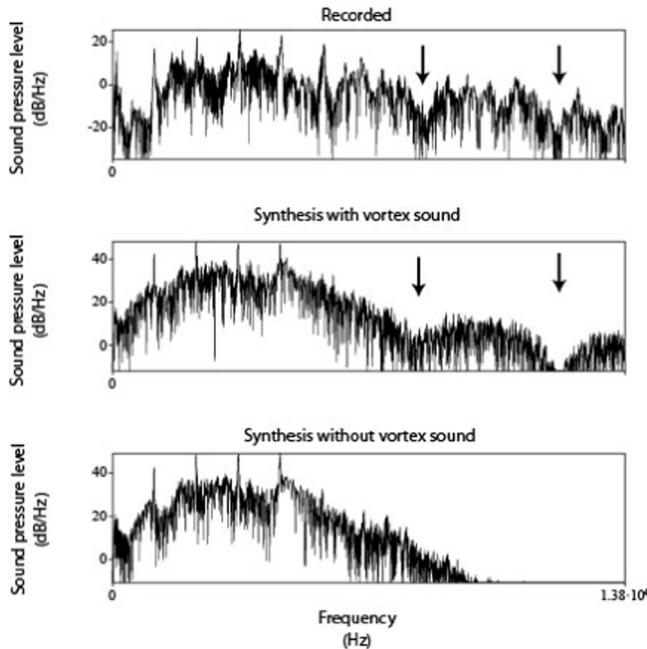


Figure 5. Sonograms of a natural song by a Zebra finch (top) and of a synthetic song obtained integrating a normal form with time varying parameters, as in figure 3, but now vortex sound is added to the output.

so that its spectrum would have the first two minima in those frequency values. The values are reasonable since they assume a constriction length of the order of 1 mm, and velocity $U_c \approx 5 \text{ m s}^{-1}$.

Songs incorporating both mechanisms resemble the alternation of fricatives and vowels in human speech. The analogy is actually quite precise: vortex sound is at the heart of fricative sounds, just as those generated through mass injection involve the same mechanisms than vowels.

4. Conclusions

Significant advances have been made in the elaboration of computational models for birdsong. The elaboration of a computational model using as parameters the actual physiological instructions allows integrating our understanding of the problem. In fact, a computational model can not only summarize a series of experiments but also turn them into an operational and predictive theory. In this work we both present a review of previous results as well as, for the first time, a discussion on vortex sound in the framework of birdsong production. This allows reproducing the characteristics of the timbre associated with high spectral components.

Behavior emerges from the interaction between a nervous system, a biomechanical device, and its environment. How much of the complexity observed in a particular behavior depends on each of these factors? In the case of birdsong, a mathematical model describing the complex periphery as a nonlinear dynamical system leads to the conclusion that nontrivial behavior emerges even when the organ is commanded by

simple motor instructions: smooth paths in a low dimensional parameter space. An analysis of the model provides insight into which parameters are responsible for generating a rich variety of diverse vocalizations, and what the physiological meaning of these parameters is. One of the most significant contributions of this line of work is to warn about which features are independently controlled and which are bounded to appear highly correlated due to biomechanical restrictions.

Beyond their role in our understanding, computational models of phonation will play an important role in the development of bio-prosthetic devices for human phonation. The feasibility of real time integration of low dimensional models driven by physiological instructions has been demonstrated in birds [20]. It is the use of low-dimensional nonlinear mathematical models of the peripheral effector what allows the emulation to be computed with very small computational effort. This is an example of the plausibility of a kind of interface between the central motor pattern generator and the synthetic, bio-mimetic behavior. An advance towards models in which certain complex features of the motor behavior are understood in terms of the underlying non-linear mechanisms of the peripheral effectors has the potential to enhance solutions of brain-bio-mimetic effector interfaces in many ways. If met by theoretical advances in the modeling of vortex sounds, this will open interesting possibilities of electronic bio-prosthetic applications in impaired humans.

Acknowledgments

This work was partially supported by CONICET, ANCyT, UBA, and NIH through R01-DC-012859 and R01-DC-006876. Suggestions by Ana Amador, Franz Goller and Dan Margoliash are greatly appreciated.

References

- [1] Chiel H J and Beer R D 1997 The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment *Trends Neurosci.* **20** 553–7
- [2] Laje R and Mindlin G B 2005 *The Physics of Birdsong* (Berlin: Springer)
- [3] Goller F and Riede T 2013 Integrative physiology of fundamental frequency control in birds *J. Physiol.* **107** 230–42
- [4] Ishizaka K and Flanagan J L 1972 Synthesis of voiced sounds from a two-mass model of the vocal cords *Bell Syst. Tech. J.* **51** 1233–68
- [5] Lucero J C and Koenig L L 2005 Phonation thresholds as a function of laryngeal size in a two-mass model of the vocal folds *J. Acoust. Soc. Am.* **118** 2798–801
- [6] Lucero J C, Koenig L L, Lourenço K G, Ruty N and Pelorson X 2011 A lumped mucosal wave model of the vocal folds revisited: recent extensions and oscillation hysteresis *J. Acoust. Soc. Am.* **129** 1568–79
- [7] Titze I R 1988 The physics of small-amplitude oscillation of the vocal folds, 1988 *J. Acoust. Soc. Am.* **83** 1536–52
- [8] Sitt J D, Amador A, Goller F and Mindlin G B 2008 Dynamical origin of spectrally rich vocalizations in birdsong *Phys. Rev. E* **78** 011905
- [9] Goller F and Suthers R A 1996 Role of syringeal muscles in controlling the phonology of bird song *J. Neurophysiol.* **76** 287–300
- [10] Goller F and Suthers R A 1996 Role of syringeal muscles in gating airflow and sound production in singing brown thrashers *J. Neurophysiol.* **75** 867–76
- [11] Mindlin G B, Gardner T J, Goller F and Suthers R 2003 Experimental support for a model of birdsong production *Phys. Rev. E* **68** 041908
- [12] Amador A, Goller F and Mindlin G B 2008 Frequency modulation during song in a suboscine does not require vocal muscles *J. Neurophysiol.* **99** 2383–9

Avian vocal production beyond low dimensional models

- [13] Perl Y S, Arneodo E M, Amador A, Goller F and Mindlin G B 2008 Reconstruction of physiological instructions from Zebra finch song *Phys. Rev. E* **84** 051909
- [14] Margoliash D 1983 Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow *J. Neurosci.* **3** 1039–57
- [15] Amador A, Perl Y S, Mindlin G B and Margoliash D 2013 Elemental gesture dynamics are encoded by song premotor cortical neurons *Nature* **495** 59–64
- [16] Gardner T, Cecchi G, Magnasco M, Laje R and Mindlin G B 2001 Simple motor gestures for birdsongs *Phys. Rev. Lett.* **87** 208101
- [17] Riede T, Suthers R A, Fletcher N H and Blevins W E 2006 Songbirds tune their vocal tract to the fundamental frequency of their song *Proc. Natl Acad. Sci.* **103** 5543–8
- [18] Blake W K 1986 *Mechanics of Flow Induced Sound and Vibrations* (San Diego, CA: Academic)
- [19] Howe M S 2003 *Theory of Vortex Sound* (Cambridge: Cambridge University Press) vol 33
- [20] Arneodo E M, Perl Y S, Goller F and Mindlin G B 2012 Prosthetic avian vocal organ controlled by a freely behaving bird based on a low dimensional model of the biomechanical periphery *PLoS Comput. Biol.* **8** e1002546
- [21] Krane M H 2005 Aeroacoustic production of low frequency unvoiced speech sounds *J. Acoust. Soc. Am.* **118** 410–27
- [22] Villermaux E and Hopfinger E J 1994 Self-sustained oscillations of a confined jet: a case study for the non-linear delayed saturation model *Physica D* **72** 230–43