Controllable Biomimetic Birdsong

INTERFACE

rsif.royalsocietypublishing.org

Research

Cite this article: Mukherjee A, Mandre S, Mahadevan L. 2017 Controllable biomimetic birdsong. *J. R. Soc. Interface* **14**: 20170002. http://dx.doi.org/10.1098/rsif.2017.0002

Received: 1 January 2017
Accepted: 4 July 2017

Subject Category:
Life Sciences–Engineering interface

Subject Areas:
biomimetics, biophysics

Keywords:
birdsong, biomimetics

Author for correspondence:
L. Mahadevan
e-mail: lmahadev@g.harvard.edu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3825487.

THE ROYAL SOCIETY PUBLISHING

Controllable biomimetic birdsong

Aryesh Mukherjee[1], Shreyas Mandre[4] and L. Mahadevan[1,2,3,5,6]

[1]Paulson School of Engineering and Applied Sciences, [2]Department of Organismic and Evolutionary Biology, and [3]Department of Physics, Harvard University, Cambridge, MA 02138, USA
[4]School of Engineering, Brown University, Providence, RI, USA
[5]Kavli Institute for Nanobio Science and Technology, Harvard University, Cambridge, MA, USA
[6]Wyss Institute for Biologically Inspired Engineering, Boston, MA, USA

AM, 0000-0003-1165-8441; SM, 0000-0002-1525-8325; LM, 0000-0002-5114-0519

Birdsong is the product of the controlled generation of sound embodied in a neuromotor system. From a biophysical perspective, a natural question is that of the difficulty of producing birdsong. To address this, we built a biomimetic syrinx consisting of a stretched simple rubber tube through which air is blown, subject to localized mechanical squeezing with a linear actuator. A large static tension on the tube and small dynamic variations in the localized squeezing allow us to control transitions between three states: a quiescent state, a periodic state and a solitary wave state. The static load brings the system close to threshold for spontaneous oscillations, while small dynamic loads allow for rapid transitions between the states. We use this to mimic a variety of birdsongs via the slow–fast modulated nonlinear dynamics of the physical substrate, the syrinx, regulated by a simple controller. Finally, a minimal mathematical model of the system inspired by our observations allows us to address the problem of song mimicry in an excitable oscillator for tonal songs.

1. Introduction

The planning, execution, learning and evolution of birdsong has inspired scientists from a range of fields with an interest in understanding its origins, ethology, neurobiology and biophysics [1–6]. Anatomically, the syrinx, the vocal organ of the bird, contains the vibratory tissues that generate sound and lies at the junction of the bronchi and trachea, and is ensconced in the interclavicular airsac [4,7]. Morphologically, there is a diversity of syrinxes across species but calls and songs arise from a simple universal process: air flow generated by the air-sac couples to the vibrating structures (lateral vibratory masses (LVMs) and tympaniform membranes (MTMs) in non-songbrids and medial vibratory mass (MVM) and medial labium (ML) in songbirds) [4] under active control by the muscles that surround the syrinx [8,9]. Dynamically, the vibrating tissues are viscoelastic materials, composed largely of extracellular matrix materials such as collagen, and thought to determine the fundamental frequency of the song by changes in their tension due to active muscular contractions [10] and passive airsac pressure [11].

To understand how birds use their syrinx requires that we take it apart functionally and morphologically, understand the function of the parts individually, and then synthesize this into a complete explanation. Experiments with excised syrinxes from zebra finches show that they are nonlinear dynamical devices—air flow fed from a pressure source attached to its posterior end causes it to becomes acoustically active with a complex harmonic spectrum that eventually becomes chaotic above a critical driving pressure [12]. In addition to the nonlinearities embodied in the geometry and elastohydrodynamics of slender vibrating structures, biomechanical studies of the stress–strain relation of the vibrating structures show that they are also materially nonlinear [10,13]. However, the role of nonlinear material properties in sound production is debated, and it is thought efficacious to avoid these nonlinearities to achieve precise control of complex songs [14].

A complementary approach asks if it is possible to recreate song artificially by building a mechanical device with enough freedom to be able to reproduce birdsong. Mechanical models that take this route [15] and are capable of phonation above critical pressures. These anatomically inspired mechanical mimics of the syrinx use a combination of soft and stiff elastic tubes [15] and have the ability to phonate and produce complex sounds resembling calls. Recent work [16] shows that controlling bronichial and airsac pressure along with external muscle stimulation on an excised syrinx can be used to control the fundamental frequency, which can then be modulated. However, to produce birdsong in real time requires a fast actuation mechanism, something that has not yet been achieved. These experimental approaches serve as a minimal means to understanding the basic physical and engineering principles behind the production and control of complex birdsong, by focusing only the bare essentials of biological morphology and physiology.

Simultaneously, observations and experiments have been the inspiration for theoretical models to understand sound production. They are mathematically couched in terms of low-order dynamical systems that model the syrinx as an uncontrolled nonlinear self-excited oscillator [17,18]. Over time, by taking note of the myoelastic-aerodynamic (MEAD) mechanism used to explain human speech [19] and birdsong [16], these models have evolved [20] to incorporate the effects of muscle tension and lung pressure and can recreate certain aspects of complex songs [21]. However these lumped-parameter perspectives are unable to account for the fact that a 50-fold increase in membrane tension is required to explain the range of frequency shifts observed in a zebra finch, while mechanical measurements performed on an excised syrinx show that an applied strain as large as 30% changes the effective spring constant only by a factor of 2 [13] (and thus the frequency be just 40%). Furthermore, although enhanced mathematical models with parameters resembling measured ventral muscle activity and airsac pressure [22] from *in vivo* experiments have been used to create zebra finch songs [21], the physical parameters used correspond to a linear spring constant of $10^{-5}\,\mathrm{N\,m^{-1}}$ and a mass of 4×10^{-9} g [23] associated with a time scale $\approx 10\,\mu\mathrm{s}$ that is not consistent with the ability to control such songs [6]. All together, these studies suggest that much still needs to be done to understand the biophysical aspects of controllable birdsong.

Here we describe a biologically inspired physical model of the syrinx with real time control and demonstrate its ability to create and control complex birdsong. Our biomimetic syrinx is an elastic tube under biaxial tension with air flowing through it, and a mechanical probe that can change the tension in the tube dynamically. We show that when this artificial syrinx is loaded with a large static load, and then subject to a superposition of small dynamic loads, we can trigger and quench aeroelastic instabilities that create sound without evoking material nonlinearities, and mimic the complex vocal repertoire of a variety of songbirds. A minimal mathematical model of the system inspired by our observations allows us to address the problem of song mimicry in a linearized excitable oscillator for tonal songs. Our results suggest that the complexity of birdsong may be delegated to the dynamics of the physical substrate, the syrinx, using a large stationary load and a small dynamical one that is regulated by a simple controller.

2. Controllable biomimetic syrinx

The singing apparatus of songbirds like the zebra finch consists of the airsac, which acts as a pressure-controlled air supply [24], the bronchus, a thin tube that connects the airsacs to the syrinx, which contains the vibrating sound source (the ML and MVM) [4,13], and finally the acoustic filter formed by the tongue and beak [25]. In songbird species, thick muscles surround the syrinx and have a direct effect on the sound features as shown in muscle stimulation studies [16,26] but the exact mechanism of how muscle action and sound activity are related is still only partly understood. These observations serve as the backdrop for a minimal controllable artificial syrinx, which while inspired by earlier physical and mathematical models [15,27], differ in a fundamental way by incorporating a new method of scale separation of loading and control which allows for real-time birdsong production.

2.1. Physical model

To mimic the vibrating structures of the syrinx and the bronchus (figure 1*a*), we used a soft thin elastic tube (Young's modulus ≈ 1 MPa) of length 2 cm, outer diameter 2.5 mm and thickness 0.1 mm shown in figure 1*b* (also electronic supplementary materials, figures S1,S2). Our physical model deliberately seeks to avoid the complexity of real bronchi and the layered nature of the MTM, ML and MVM, but allows us to focus on understanding the controllable elastohydrodynamics of phonation. Furthermore, our device is scaled up relative to the zebra finch syrinx, which has a size of $O(1\,\mathrm{mm})$, and so its physical properties such as the bending modulus need to be scaled up accordingly to replicate the range of frequencies observed in nature. The dynamic action of the airsac was mimicked by a velocity-controlled air meter (flow rate $\approx 10^{-5}\,\mathrm{m^3\,s^{-1}}$) attached to the caudal end of the tube, while the cranial end was open to the atmosphere and free to vibrate. In reality, the airsac controls the pressure so that controlling flow velocity in the experiment is a simplification, but valid in the limit that the impedance of the bronchus is larger than that of the syrinx. Operating in the velocity-controlled mode ensures that the device does not behave like a pressure-controlled valve.

Song production requires the use of muscles around the syrinx to control the frequency content. For example, in the zebra finch, it has been observed that the major muscles syringealis ventralis, tracheobronchialis ventralis and tracheobronchialis dorsalis [28] directly control the position and tension of the MVM and LL (lateral labia) and reoreint the syrinx to a basal tensile state before phonation [8]. In our set-up, translation stages attached to the rear and front of the device (figure 1*b*) allowed us to adjust global longitudinal and lateral tensions to set the operation point of the device in a manner similar to actual songbirds. The translation stages are large and cannot be actuated quickly so can only provide the basal tensile state and slow control of global tension. The rapid changes in song frequency and amplitude, characteristic of finch songs, were induced by the action of a linear linear motor that presses down on the rubber tube and squeezes it at a point along its length (figure 1*b*; electronic supplementary material, figures S1,S2). The linear motor can be actuated rapidly at frequencies up to 30 Hz with amplitudes of up to 1 mm allowing us to

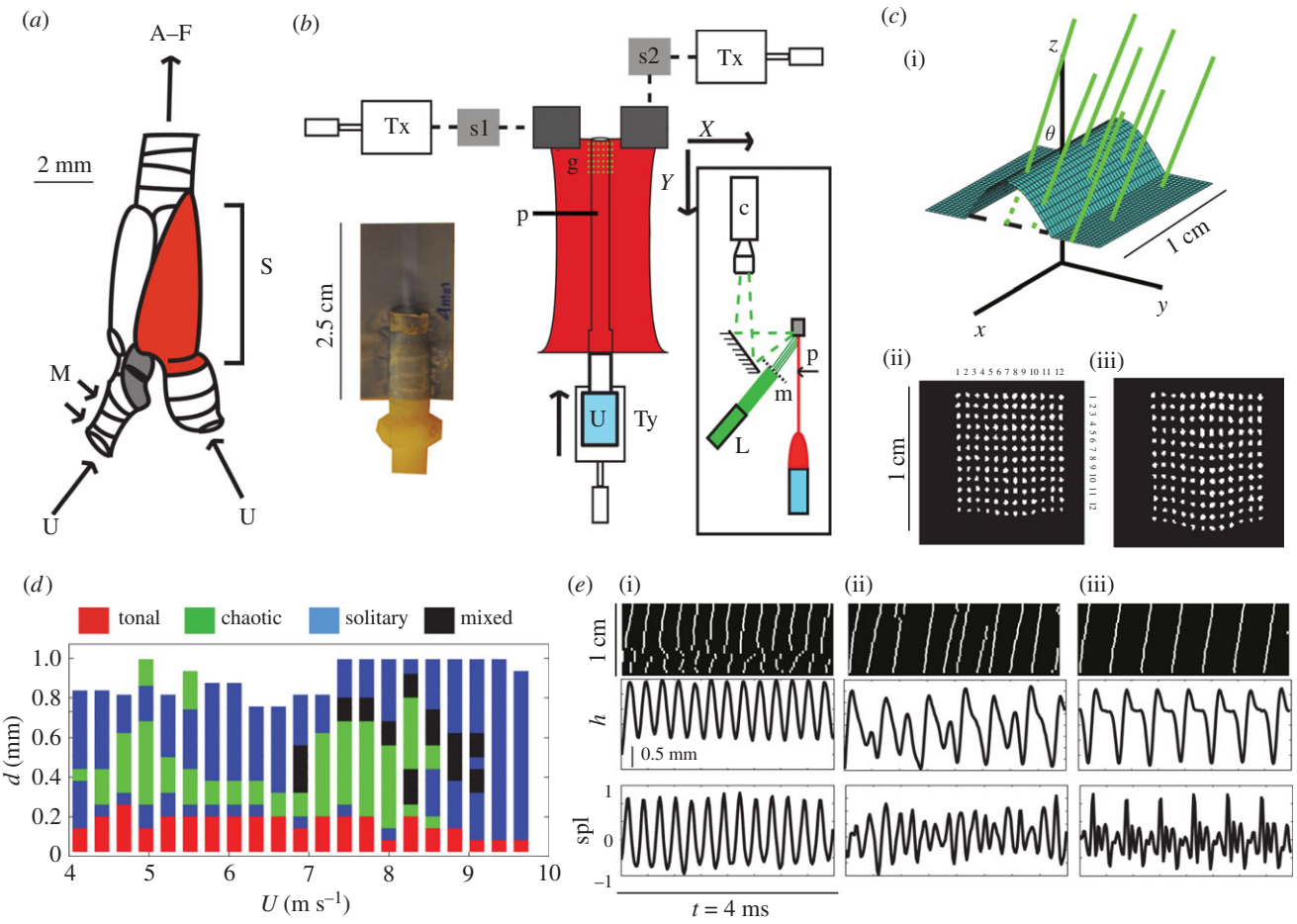

Figure 1. A bird syrinx, its biomimetic analogue, and the dynamics of the device. (*a*) Schematic of a zebra finch syrinx. S is the sound box (syrinx) of the bird, U is airflow coming from the lungs, A – F is the acoustic filter formed by the oropharyngeal – oesophageal cavity, beak and trachea [29], M denotes muscular action along the length of the bronchus. (*b*) Schematic of experimental set-up. s1, Tx, strain gauge and stage to control lateral tension; s2, Ty, gauge and stage to control longitudinal tension; U, flowmeter; g, grid projected onto device. Inset (on right) shows side view of the set-up. L, 10 mW Green laser pointer; m, mask used to produce grid; p, probe used to actuate device; c, high-speed video camera (up to 90 kHz frame rate). The probe p is a linear motor to excite the membrane by pushing anywhere along its length, and mimics muscular actuation along the length of the bronchus. An image of the actual device is shown on the left. (*c*)(i) – (iii) Demonstration of the principle of the optical reconstruction. (i) shows the top surface of the device, described by the function $z(x, y)$, being impinged by pencils of laser light described by the functions $z_n(x, y)$, where the index n characterizes the raster location. The light is incident on the device at an angle θ and the dashed lines show the intended path of the laser before it is intercepted by the surface. (ii),(iii) Images of the raw data with the flow turned off and on, respectively. The distance that each point in (*c*)(iii) has moved from its original location in (c)(ii) is proportional to the height at the observed point and this information can be used to reconstruct the entire surface in three dimensions. (*d*) The phase space of a device as a function of flow velocity U and probe depth d measured relative to the first depth (100 μm) at which sound is produced. Note that the solitary regime may have further subharmonic bifurcations A 'mixed' phase between chaotic and solitary also exists, marked in black in the phase space. (*e*) Digitized kymograph (positive y-direction corresponds to, height $h(t)$ of the membrane at a point $y = 0.5$ cm from the mouth, and sound pressure level (spl) for the three regimes: tonal, aperiodic and solitary. (Online version in colour.)

reproduce rapidly varying transitions in complex songs. In our flow-driven model, there is no acoustic filter; to the first approximation, this has negligible feedback on the vibrating surface and can be omitted. To capture the physical mechanism of sound production and the sound itself, we visualized the surface profile of the oscillating tube at 100 kHz using a three-dimensional direct imaging technique (figure 1*c*; electronic supplementary material), while simultaneously recording the sound with a microphone.

2.2. Uncontrolled passive modes

The onset of phonation requires a critical flow velocity in the range 0–8 m s^{-1}, that is a function of the global applied tension in the tube, which ranges from 0 to 100 mN. Even without any localized application of stress via the linear probe, the biomimetic syrinx produced sound when the global tension crosses a threshold value, even at low velocities (0.5 m s^{-1}). In this regime associated with global modes of the whole tube, the frequency of sound scaled weakly with the applied tension (between 2 and 3 kHz). Birdsong typically includes large and rapid swings in frequency, so that the set point of the device was adjusted to avoid exciting these global self-oscillation modes over the range of flow velocities used. The application of localized constriction via the linear actuator causes changes in the tension while also causing the local air velocities to increase. As a result, the transmural pressure drops. This couples with elastic modes of the tube to create self-sustained oscillations that produced sound with a large range in frequency (fundamental frequency 1–5 kHz, harmonics up to 20 kHz were observed); similar modes of oscillation have also been seen when air flows through un-tensioned soft cylindrical shells [30]. In the presence of air flow, the localized squeezing of the tube

yields a rich repertoire of sounds, including tonal, bursting and chaotic behaviour as a function of probe depth and flow velocity.

For small values of the displacement by the linear motor, the device produces sinusoidal modes as shown in figure 1*d*,*e* (see also electronic supplementary materials, figure S2*a*). These appear as travelling waves emanating from the probe location and propagate downstream (electronic supplementary material, Video S1). The static position of the probe controls the frequency of the sound produced; sweeping through frequency can be achieved via small dynamical displacements superposed on this static position (figure 2*e*; electronic supplementary material) as the frequency is very sensitive to probe position. These observations suggest that simple auditory gestures can be created by moving the linear actuator along a prescribed path as a function of time, at a rate (30 Hz) much slower than the song frequency (1 kHz) itself (electronic supplementary materials, Audio S1). When the probe presses further down into the device a new mode of oscillation appears, wherein a short section of the tube of length comparable with its radius forms a localized buckle that rapidly travels towards the frontal end of the tube at steady velocity of $35 \pm 5\,\mathrm{m\,s^{-1}}$ (electronic supplementary materials, figure S4, Video S3). The velocity of these localized pulses shows only a weak dependence on probe displacement or flow velocity. Following a very short pause (approx. 1 ms, see figure 1*e*(iii); electronic supplementary materials, figures S3,S4) during which the tube regains its original cylindrical shape, the cycle repeats itself. This is suggestive of the system being close to an excitability threshold, with a refractory period characteristic of an excitable mediaum [31]. In this regime of oscillation, the fundamental frequency and the harmonic structure of the sound are determined by the duration of the pulses which can be controlled by the position of the linear probe (electronic supplementary materials, figure S4). In this same regime, further complex sounds sometimes arise via period doubling bifurcations (see electronic supplementary materials, figure S3*b*, Video S2), with some qualitative similarities to the structure of certain zebra and bengalese finch song motifs [22]. As the probe indentation depth is further increased (figure 1*d*,*e*), we see a third regime wherein chaotic sounds are often heard, as in many zebra finch songs [32]. This phase is characterized by solitary pulses with stochastic refractory times between pulses as shown in figure 1*e*(ii). Indenting the probe into the tube still further eventually chokes the flow and causes the tube to burst.

Collectively, our observations bear a strong resemblance to self-excited hydroelastic modes observed in a class of devices called Starling resistors [33,34] that have been used to study bioinspired fluid–structure instabilities in the context of Korotkoff sounds [35]. However, there are also some critical differences that are linked to the ability to excite and extinguish sound production using a linear actuator that controls the appearance of dynamical signatures associated with the periodic, excitable and solitary modes of oscillation. Moving through these regimes of simple tonal oscillations and self-sustained localized oscillations is possible by varying the position of the probe by less than 10% and thus allows for a range of sounds via transitions between these phases in a manner that is qualitatively similar to fine muscle control over song. The tonal phase can be used to reproduce sounds from birds like a canary or vireo, while the bursting phase is required to mimic sounds produced by birds like the zebra or bengalese finch. Our observations suggest a mechanism for the real-time programming of a song sequence into the dynamics of the linear motor using the modular design and control of the biomimetic syrinx, whereby the global stretch of the tube brings the system close to threshold for spontaneous oscillations, while the rapidly controllable small strains (≈ 0.01) controls the onset of oscillation. This allows for the flexibility required in switching modes of phonation and thence song production.

We pause to emphasize that there are many detailed differences between an actual syrinx and our simple physical abstraction. For example, in a song bird the vibrating structure (MVM) is rather short (1 mm) and soft (1 kPa) compared with our long (1 cm) tube-like device made of a much stiffer elastomer. This leads to an absolute difference in the wave speeds observed in the two systems, but can be reconciled by the fact that the bending wave speed scales as $c \sim (1/r)\sqrt{Et^2/\rho}$, where r is the radius of the tube and t its thickness. Using the following parameter values as described in [13,28], $r \sim 10^{-3}$ m, $t \sim 10^{-4}$ m, $E \sim 1$ kPa, $\rho \sim 10^3\,\mathrm{kg\,m^{-3}}$, we find that $c_{\mathrm{model}}/c_{\mathrm{bird}} \sim 30$, and explaining the discrepancies in wave velocities observed. However, we note that within our experimentally observed paradigm of sound production via solitary pulses, the fundamental frequency is set by the time taken for the solitary pulse to reach the mouth from its point of origin, which would be the posterior end of the MVM for the bird and the location of the probe for the device, i.e. it scales as c/l, where l is the length of the vibrating structure. Hence although the device is much longer, the wave speed is also larger, which results in a similar fundamental frequencies for both real and mimicked structures. This implies that in spite of the morphological simplifications and relatively larger size of our device, we may still capture the essential biophysics of song production, and match what is observed in nature up to scaling factors.

2.3. Controllable biomimetic birdsong

Using a temporally programmed indentation profile $h(t)$ into the probe allows us to to sweep through frequency (electronic supplementary materials, figure S5), and figure 2*a* shows how sound with high harmonic content can be obtained with large successive frequency sweeps obtained by this method. We first position the motor so that the tube is poised at the onset of phonation and then use the motor as a switch that can be turned on or off rapidly. Thus, transitions between sounds can be triggered dynamically by choosing the bias points appropriately. In figure 2*b*,*c*, we show that it is possible to switch the acoustic behaviour of the device from generating solitary, pulsed to chaotic modes of vocalization as the motor is dynamically activated to move through the different phases that we delineated previously (electronic supplementary materials, Audio S1).

To demonstrate the versatility of the biomimetic syrinx, we chose to mimic three characteristic songs from three species (electronic supplementary material, Audio): vireo songs that are mostly tonal, bengalese finches that have songs with notes of both high and single harmonic content, and zebra finch songs that have notes with very high harmonic content and complicated transitions. To create the song of the vireo, the motorized probe was placed close to the mouth of the device to ensure operation in the tonal regime, and its

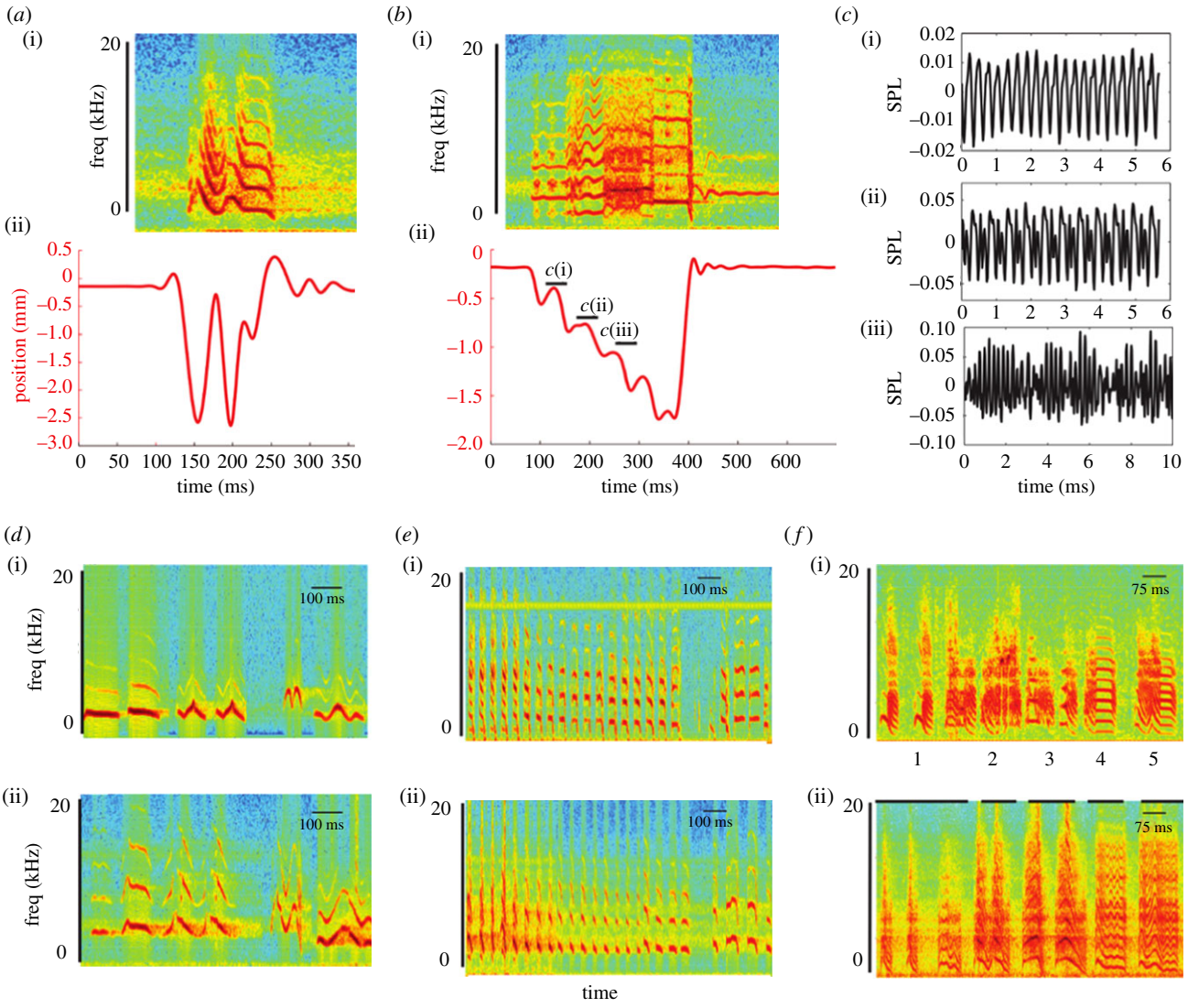

Figure 2. Controllable transitions in a biomimetic syrinx. (*a*–*c*) Dynamic action of probe. (*a*)(i) and (*b*)(i) The spectrograms of two complicated sounds created by dynamically controlling the probe. The probe is programmed with a pulse sequence to create the sounds (we specify the position and acceleration). (*a*)(ii) and (*b*)(ii) The position of the probe, relative to the spectrogram of the sound created. (*c*)(i)–(iii) The sound pressure levels for the sections indicated in (*b*)(ii). (*c*)(i) Sinusoidal, (ii) oscillations following a period-doubling transition, and (iii) chaotic sound. We see that by actuating the probe, we can change the qualitative nature of sound production. (*d*–*f*) Spectrograms of real and mimicked bird songs (using dynamic action of the probe). (*d*)(i) A song from a red-eyed vireo, and (*d*)(ii) songs mimicked by dynamical control of the device. The vireo has beautiful tonal songs with rapid and large variations in frequency. The device was operated in the tonal regime and reproduces songs well when the variation in frequency is not very large. (*e*)(i) A real and (*e*)(ii) mimicked song from a bengalese finch. These songs consisted of short harmonic pulses, which were well reproduced by the device. Even subtle frequency changes and transitions to a period doubled mode were well captured. (*f*)(i) A real and (*f*)(ii) mimicked song from a zebra finch. These songs show large spectral variation as well as high harmonic content and are difficult to reproduce. However, given the limitations of the linear motor, many features of the zebra finch song (large changes in frequency, transitions to chaos, high harmonic content) can be qualitatively reproduced. (Online version in colour.)

depth was programmed to change $\omega(t)$, the local fundamental frequency. The two spectrograms (electronic supplementary materials, Audio S2) show that this simple principle can be used to achieve a high degree of similarity of the mimicked song to the original; figure 2*d* confirms this visually be comparing the spectrogram of a recreated song of a vireo along with the original. To create the song of the bengalese finch, the motor was placed further behind the mouth to enable both tonal and pulsed sounds. The first section of the song shows pulses with large harmonic content created using the probe as a switch, rapidly moving to a depth commensurate with the fundamental frequency of the song and then rapidly retracting it. The first two syllables show a period doubled sound created by the device (electronic supplementary material, Audio S3). The second part of the song shows tonal structure much like that of a vireo song, and can be recreated in a similar fashion; a spectrogram of the characteristic pulsed song of a bengalese finch and its reproduction using our device is shown in figure 2*e*. Finally, we turn to the zebra finch song that combines the high-speed pulses of a bengalese finch with sweeps in tonal songs and rapid transitions from tonal to multiharmonic or from multiharmonic to chaotic. To create these notes the motor was placed even further behind the mouth; tonal to multiharmonic sweeps were generated using a tonal to solitary transition and from solitary to chaotic transition to generate multiharmonic to chaotic sweeps. The recreated song bears a qualitative resemblance to the original but is not exact because of limitations on the maximum acceleration of the linear probe. The song shows three distinct features; sweeping transitions from single to multiharmonic sounds (1 and 3 in figure 2*f*), a sweeping transition to chaotic sounds (2 in figure 2*f*) and

multiharmonic sweeps to harmonic stacks (4 and 5 in figure 2*f*). The recreated spectrogram shown in figure 2*f* (electronic supplementary material, Audio S4) shows oscillations in the frequency especially for the last notes due to ringing of the linear motor itself and could not be avoided as the motor was moving close to its maximum acceleration.

3. Mathematical model for tracking a tonal song

Having shown that our biomimetic device can effectively mimic a range of songs, a natural question that arises is that of quantifying the acoustics of sound and song production, and its control. The complex elastohydrodynamics of phonation and the ensuing nonlinear dynamics of fluid-structure interaction is a difficult problem in general. But is it possible to capture the dynamics of a simple song and consider its control? The modular nature of operating the syrinx is such that we can bring the system close to threshold slowly and then excite it rapidly. In the case of a tonal song, such as as that of the vireo, this is tantamount to exciting a single frequency that is changed, so that the system moves from a silent regime to a spontaneously oscillating one, and back. Here we couch this as a problem of determining how to move in between these regimes to track a given song, to mimic the vocal response of a finch son to that of his father's birdsong. When our system is close to the threshold for self-excited oscillations, it may be minimally modelled as an oscillator with a dynamically varying spring stiffness that allows for a controllable frequency $\omega(t)$ and a controllable damping coefficient $r(t)$ whose sign changes as the system transitions through the threshold for instability, a linearized variant of the nonlinear oscillator models used to mimic birdsong [20], but now couched in an optimization framework. The equation of motion for a harmonic oscillator with time varying stiffness and damping can be written as

$$\ddot{x} + r(t)\dot{x} + \omega^2(t)x = 0, \tag{3.1}$$

where $x(t)$ models the acoustic signal corresponding to the song, and $r(t)$ and $\omega(t)$ are two functions that the bird manipulates in order to generate the signal. In reality, the frequency is effectively controlled by the static and dynamic tube tensions, while the damping is controlled by the flow rate. Given a target bird song as a time series $u(t)$, we ask to determine slowly varying functions $r(t)$ and $\omega(t)$, for which the solution of (3.1) comes close to the target, formulated as the solution of the following problem: find $r(t)$, $\omega(t)$ subject to (3.1) that minimizes the global error

$$\int_0^T (x - u(t))^2 + W_1\left(\frac{\mathrm{d}\omega}{\mathrm{d}t}\right)^2 + W_2\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 \mathrm{d}t. \tag{3.2}$$

Here the first term penalizes deviations in the produced song, while large weights W_1 and W_2 suppress rapid variations in $\omega(t)$ and $r(t)$ (even though ω itself can be large, and r itself can be very small), and the harmonic oscillator equation serves as a minimal model near the onset of phonation. When these parameters vary slowly, we take advantage of the WKB approximation [36] for the solution of (3.1), which allows the approximate solution to (3.1) to be written as

$$x(t) = \frac{1}{\sqrt{\omega(t)}}\left(A\,\mathrm{e}^{(-\Psi(t)+\mathrm{i}\Phi(t))} + B\,\mathrm{e}^{(-\Psi(t)-\mathrm{i}\Phi(t))}\right), \tag{3.3}$$

where $\Psi(t) = \int_0^t (r(s)/2)\,\mathrm{d}s$ and $\Phi(t) = \int_0^t \omega(s)\,\mathrm{d}s$, and A and B are arbitrary constants determined by initial conditions. Substituting these results in (3.2), we convert the differential optimization algorithm to an algebraic one by dividing the duration of the bird song into N intervals, with r and ω approximated to be constants in each interval but varying from interval to interval, and A and B determined so that x and $\dot{x}$ are continuous across intervals. The global optimization problem is then converted into a sequence of N local optimization problems, where for the kth local optimization problem, the values of r_k and ω_k are determined by the solution of the following problem:

$$\underset{r_k,\omega_k}{\mathrm{Min}} \int_{t_{k-1}}^{t_k} (x - u(t))^2 + \tilde{W}_1(\omega_k - \omega_{k-1})^2 + \tilde{W}_2(r_k - r_{k-1})^2\,\mathrm{d}t,$$

where $x(t)$ is given by (3.3) by assuming $r(t) = r_k$ and $\omega(t) = \omega_k$, and $\tilde{W}_{1,2}$ are the rescaled weights $W_{1,2}$, respectively. Our simulations that were carried out using a steepest descent algorithm (Nelder-Mead) implemented in Matlab using a step size of 0.1; our results were insensitive to the choices of the weights in the regime $O(100-1000)$.

In figure 3*a*(i),(ii), we show the spectrograms of a real and simulated vireo song that shows good agreement, as confirmed in figure 3*b*(i), which shows the actual time trace $x(t)$ that sits on top of the target $u(t)$. In figure 3*b*(ii), we also show the time-dependent frequency $\omega(t)$ and the damping $r(t)$ derived from the solution of the optimization problem, noting that they vary very slowly, and are thus viable controls (electronic supplementary material, Audio S5,S6). We thus see that it is possible to capture the entire dynamics of the song using a linear oscillator model with slowly varying frequency and damping, consistent with our experimental study.

4. Discussion

Birdsong has all the hallmarks of complex behaviour in animals, as it exhibits the ability of the bird to exquisitely control a complex physical system, the syrinx, to realize the task of learning mimicry with gradual variations. The creation of song involves two separate processes. The first is the process of sound production itself via the onset of vibratory modes due to coupling between the membrane and aerodynamically generated forces. The second is the process of song control via motion through the phase space of parameters associated with frequency and amplitude modulation. Sound production via the MEAD mechanism, wherein dynamic and asymmetric membrane opening and closing leads to self sustained oscillations, has been verified in both oscine and non-oscine species [16]. However, the case for song control is still up for debate. Current *in vivo* studies measure correlations between fundamental frequency and muscle stimulation [37] but do not provide insight into details of the membrane tension or air sac pressure as a function of time. Our simple estimates suggest that controlling song production by varying these parameters requires large and rapid changes in membrane tension which have not been demonstrated experimentally. Furthermore, the effect of muscular forces on the membrane remains unclear, and, in particular, how it modulates tension on slow and fast time scales remains unknown.

Here, we have attempted to study the biomechanical aspects of the process by focusing on creating a controllable biomimetic syrinx to produce song. Following earlier work,

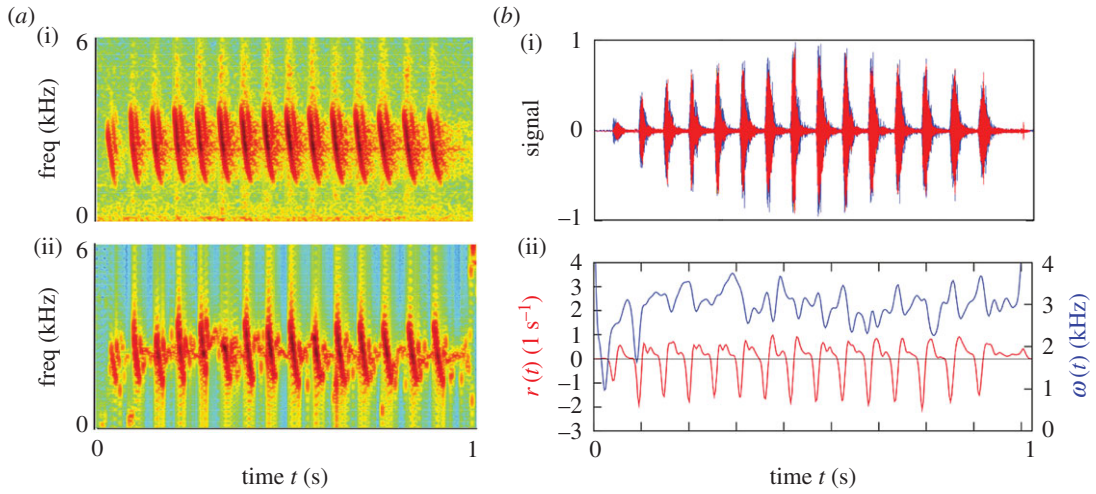

Figure 3. Mimicking a tonal song of the red-eyed vireo using a minimal model described in §3. (*a*) Spectrogram of a song of a red-eyed vireo: (i) original and (ii) computational mimic (see text). (*b*)(i) Time trace of the amplitude of the amplitude of the center frequency of the tonal song for the actual birdsong (blue) and its computational mimic (red) lie essentially on top of each other. (ii) The solution of the optimization problem (3.2) shows the temporal evolution of the damping $r(t)$ (left abscissa) and the frequency $\omega(t)$ (right abscissa). (Online version in colour.)

we have implemented this using an elastohydrodynamic instability associated with flow through a simple rubber tube, such that we can bring the system close to an excitability threshold using a combination of a static stretch and flow. A dynamic probe then allowed us to dynamically trigger phonation patterns to create the basic syllables of simple birdsongs. The advantage of this approach is that it allows for a slow, coarse actuator that brings the system close to an instability and then a using a fast, fine actuator to transition easily across the excitability threshold. The resulting behaviour of our spatially extended acoustic oscillator allows for a range of physiological outputs with a minimal controller. By placing the linear motor at different lengths along the tube, we further showed that songs from different species can be reproduced, suggesting the ease with which a simple avian vocal tract can mimic a range of songs and sounds. This strategy of localized control is qualitatively consistent with early observations made in [12], and later summarized in [13], that show that small medial rotations of the third bronchial ring that insert the lateral labium into the bronchial lumen can produce large changes in the fundamental frequency, without stretching the ML or MTM in a manner similar to the mechanical probe used in our study. Finally, we have shown how a minimal mathematical model that takes the form of an excitable harmonic oscillator with slowly varying damping and stiffness can be used to study how an optimization approach can mimic birdsong.

Overall, our study seeks to emphasize that the complexity of the overall task of producing birdsong can be broken down into a set of simpler biomechanical and control tasks that may not be so difficult to accomplish individually. This suggests that by harnessing the physical substrate of a soft tube susceptible to self-excited oscillations, developing organisms and perhaps evolution itself might have stumbled on using simple instabilities to produce and control birdsong. More broadly, our study adds to the growing realization that physical instabilities with rich nonlinear dynamics, when coupled to relatively simple control mechanisms, may provide a mechanism for organisms to begin to create complex behaviour by taking advantage of their physical, material nature.

Data accessibility. The datasets supporting this article have been uploaded as part of the supplementary material.

Competing interests. We declare we have no competing interests.

Funding. We thank the MacArthur Foundation for partial financial support.

Acknowledgments. We thank B. Olvecsky for suggestions on the mechanical probes and discussions. A.M. carried out the experiments and data analysis, participated in the design of the study and drafted the manuscript; S.M. designed and carried out the computations; L.M. conceived of the study, designed the study, coordinated the study, helped analyse the data and drafted the manuscript.

References

1. Brenowitz EA, Margoliash D, Nordeen KW. 1997 An introduction to birdsong and the avian song system. *J. Neurobiol.* **33**, 495–500. (doi:10.1002/(SICI)1097-4695(19971105)33:5<495::AID-NEU1>3.0.CO;2-#)
2. Fletcher N. 2014 Animal bioacoustics. In *Springer handbook of acoustics* (ed. TD Rossing), pp. 821–841. New York, NY: Springer.
3. Elemans CP. 2014 The singer and the song: the neuromechanics of avian sound production. *Curr. Opin. Neurobiol.* **28**, 172–178. (doi:10.1016/j.conb.2014.07.022)
4. Düring DN, Elemans CP. 2016 Embodied motor control of avian vocal production. In *Vertebrate sound production and acoustic communication*, pp. 119–157. Berlin, Germany: Springer.
5. Slabbekoorn H, Smith TB. 2002 Bird song, ecology and speciation. *Phil. Trans. R. Soc. Lond. B* **357**, 493–503. (doi:10.1098/rstb.2001.1056)
6. Suthers RA, Margoliash D. 2002 Motor control of birdsong. *Curr. Opin. Neurobiol.* **12**, 684–690. (doi:10.1016/S0959-4388(02)00386-0)
7. Riede T, Goller F. 2010 Functional morphology of the sound generating labia in the syrinx of two songbird species. *J. Anatomy* **216**, 23–36. (doi:10.1111/j.1469-7580.2009.01161.x)
8. Larsen ON, Goller F. 2002 Direct observation of syringeal muscle function in songbirds and a parrot. *J. Exp. Biol.* **205**, 25–35.
9. Laje R, Gardner TJ, Mindlin GB. 2002 Neuromuscular control of vocalizations in birdsong: a model. *Phys. Rev. E* **65**, 051921. (doi:10.1103/PhysRevE.65.051921)

10. Riede T, Goller F. 2014 Morphological basis for the evolution of acoustic diversity in oscine songbirds. *Proc. R. Soc. B* **281**, 20132306. (doi:10.1098/rspb.2013.2306)
11. Amador A, Goller F, Mindlin GB. 2008 Frequency modulation during song in a suboscine does not require vocal muscles. *J. Neurophys.* **99**, 2383–2389. (doi:10.1152/jn.01002.2007)
12. Fee MS, Shraiman B, Pesaran B, Mitra PP. 1998 The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature* **395**, 67–71. (doi:10.1038/25725)
13. Fee MS. 2002 Measurement of the linear and nonlinear mechanical properties of the oscine syrinx: implications for function. *J. Comp. Phys. A* **188**, 829–839. (doi:10.1007/s00359-002-0349-z)
14. Elemans CP, Laje R, Mindlin GB, Goller F. 2010 Smooth operator: avoidance of subharmonic bifurcations through mechanical mechanisms simplifies song motor control in adult zebra finches. *J. Neurosci.* **30**, 13 246–13 253. (doi:10.1523/JNEUROSCI.1130-10.2010)
15. Elemans CPH, Muller M, Larsen ON, VanLeeuwen JL. 2009 Amplitude and frequency modulation control of sound production in a mechanical model of the avian syrinx. *J. Exp. Biol.* **212**, 1212–1224. (doi:10.1242/jeb.026872)
16. Elemans CPH *et al*. 2015 Universal mechanisms of sound production and control in birds and mammals. *Nat. Commun.* **6**, 8978. (doi:10.1038/ncomms9978)
17. Fletcher N. 1988 Birdsong—a quantitative acoustic model. *J. Acoust. Soc. Am.* **84**, S178–S178. (doi:10.1121/1.2025990)
18. Fletcher N, Tarnopolsky A. 1999 Acoustics of the avian vocal tract. *J. Acoust. Soc. Am.* **105**, 35–49. (doi:10.1121/1.424620)
19. Titze IR. 1988 The physics of small-amplitude oscillation of the vocal folds. *J. Acoust. Soc. Am.* **83**, 1536–1552. (doi:10.1121/1.395910)
20. Mindlin GB, Laje R. 2006 *The physics of birdsong*. Berlin, Germany: Springer.
21. Amador A, Perl YS, Mindlin GB, Margoliash D. 2013 Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature* **495**, 59–64. (doi:10.1038/nature11967)
22. Perl YS, Arneodo EM, Amador A, Goller F, Mindlin GB. 2011 Reconstruction of physiological instructions from zebra finch song. *Phys. Rev. E* **84**, 051909. (doi:10.1103/PhysRevE.84.051909)
23. Amador A, Mindlin GB. 2008 Beyond harmonic sounds in a simple model for birdsong production. *Chaos* **18**, 043123. (doi:10.1063/1.3041023)
24. Goller F, Suthers RA. 1996 Role of syringeal muscles in controlling the phonology of bird song. *J. Neurophys.* **76**, 287–300.
25. Riede T, Suthers RA, Fletcher NH, Blevins WE. 2006 Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc. Natl Acad. Sci. USA* **103**, 5543–5548. (doi:10.1073/pnas.0601262103)
26. Srivastava KH, Elemans CP, Sober SJ. 2015 Multifunctional and context-dependent control of vocal acoustics by individual muscles. *J. Neurosci.* **35**, 14 183–14 194. (doi:10.1523/JNEUROSCI.3610-14.2015)
27. Sitt JD, Amador A, Goller F, Mindlin GB. 2008 Dynamical origin of spectrally rich vocalizations in birdsong. *Phys. Rev. E* **78**, 011905. (doi:10.1103/PhysRevE.78.011905)
28. Düring DN, Ziegler A, Thompson CK, Faber C, Müller J, Scharff C, Elemans CP. 2013 The songbird syrinx morphome: a three-dimensional, high-resolution, interactive morphological map of the zebra finch vocal organ. *BMC Biol.* **11**, 1. (doi:10.1186/1741-7007-11-1)
29. Ohms VR, Snelderwaard PC, Ten Cate C., Beckers GJ. 2010 Vocal tract articulation in zebra finches. *PLoS ONE* **5**, e11923. (doi:10.1371/journal.pone.0011923)
30. Zimoch P, Tixier E, Hsu J, Winter A, Hosoi A. 2012 Fast, large amplitude vibrations of compliant cylindrical shells carrying a fluid. arXiv:1210.3861.
31. Izhikevich EM. 2007 *Dynamical systems in neuroscience*. Cambridge, MA: MIT Press.
32. Marler PR, Slabbekoorn H. 2004 *Nature's music: the science of birdsong*. New York, NY: Academic Press.
33. Olsen JH, Shapiro AH. 1967 Large-amplitude unsteady flow in liquid-filled elastic tubes. *J. Fluid Mech.* **29**, 513–538. (doi:10.1017/S0022112067001004)
34. Cowley SJ. 1982 Elastic jumps on fluid-filled elastic tubes. *J. Fluid Mech.* **116**, 459–473. (doi:10.1017/S002211208200055X)
35. Drzewiecki GM, Melbin J, Noordergraaf A. 1989 The Korotkoff sound. *Ann. Biomed. Eng.* **17**, 325–359. (doi:10.1007/BF02368055)
36. Jeffreys H. 1924 On certain approximate solutions of linear differential equations of the second order. *Proc. Lond. Math. Soc.* **23**, 428–436.
37. Suthers RA, Zollinger SA. 2008 Neuroscience of birdsong (eds HP Zeigler, P Marler). Cambridge, UK: Cambridge University Press.

,

,

Supplementary Information for "Controllable Biomimetic Birdsong"

A. Mukherjee,[1] S. Mandre,[2] and L. Mahadevan[1,3]

[1]*Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

[2]*School of Engineering and Applied Sciences, Brown University, Providence, Rhode Island, USA*

[3]*Department of Organismic and Evolutionary Biology and Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA**

(Dated: June 18, 2017)

EXPERIMENTAL SETUP

Our biomimetic syrinx was built by attaching two pieces of nitrile rubber ($4cm \times 4cm$) along its lateral edges with an adhesive, leaving an inflatable flat rectangular section ($0.25cm \times 2cm$) in the middle to form the vocal tract. The device was then attached to a flow meter at the rear entry port, and to two translation stages with strain gauges at the exit port. The flow meter was also mounted on a third translation stage. The three stages were used to apply pre-tensions on the device (Fig S1,S2).

Typically, stereoscopic imaging techniques are used to capture three dimensional aspects of vibrating surfaces. However the small scale and high speeds required in our syrinx ($100\mu m$,$100kHz$) rendered stereoscopic imaging too cumbersome to implement. Instead, we used a 12 x 12 optical grid created by passing an expanded collimated beam of laser light through an appropriate stencil that was projected on the device.

The optical source was a 10mW laser pointer (from Dragon Lasers), which was beam expanded to 4 times using a 4:1 telescopic setup. The expanded beam was then passed through a grid ($7mmx7mm$; 12 holes on each side), made out of paper, in which holes ($d = 0.25mm$; spacing $0.625mm$) were cut out using a Versalaser laser cutter. The grid was placed no more than 5cm away from the device to prevent any diffraction rings from the holes, that might interfere with the imaging. The images of the grid projected onto the surface of the membrane were reflected of a mirror (M1 in Fig S2a) and captured with a Phantom V7.3 high speed camera at 88kHz frame rate (see Fig S1 and S2 for photographs of the actual setup).

The laser impinged on the rubber surface at 30^o; any out of plane displacement of the device caused the projected grid to deform in relation to a flat state (Fig S1,S2). The displacement of each grid point from a reference flat state was proportional to the local height, allowing us to reconstitute the entire surface by splining over the heights at the 144 sampled points. The resolution is set by the total distance covered by the grid divided by number of pixels used to sample the image which in our case corresponds to 35 microns. A substantial advantage of this technique was that it only required a single camera for three dimensional imaging.

.

FUNDAMENTAL MODES OF VIBRATION

As mentioned in the main text, when the system was allowed to spontaneously oscillate, three different states were observed: sinusoidal, solitary and chaotic. The sinusoidal regime was excited when the linear motor used to gently impinge on the device close to the mouth. This mode occupies half the length of the tube and leads to traveling sinusoidal waves as shown in Fig S3a, where one period of oscillation is shown (see Video S1 showing this tonal mode of vibration). When the linear probe was pushed further into the syrinx, we observed a transition from the sinusoidal state to an excitable solitary pulse phase. The probe locally stretches the membrane, and the applied tension allows the device to undergo oscillations. However when the amplitude becomes large, the device locally buckles and couples to solitary bending waves that travel to the mouth and the cycle repeats after a rest period. These waves are deemed solitary since they appear as localized depressions that travel with their own wave velocity different from the air flow speed. Only when one wave has passed is another depression created. Fig S3b depicts a sequence of 3d images associated with the formation and propagation of such a pulse. Fig S3c complements this with a view of the height of the membrane at different points along the length of the tube as a function of time. The cycle starts with a traveling sinusoidal wave of the kind described above ($0-0.17ms$, denoted by the bump A in Fig S3c) followed by a larger depression that travels to the mouth($0.24ms - 0.38ms$, denoted by the bump B in Fig S3c), and then a refractory rest period before the entire cycle repeats. Owing to the short duration of these temporal pulses, the sound produced is spectrally rich.

The wave speed showed only weak dependence on parameters like applied pre tension, flow speed, radius (see Fig S4a) or z-position of the probe. We found within the accuracy of our measurement that the solitary pulse traveled at the same speed of $\approx 25 - 40m/s$. Fig S4b shows that the fundamental frequency also depends on the z-position or depth of the probe; it is possible to get a large change in frequency by applying a small change in strain values. There are two time scales associated with this mode. The first is the wave speed of the solitary bump and the second is the refractory period of the solitary pulse. The observed changes in the frequency

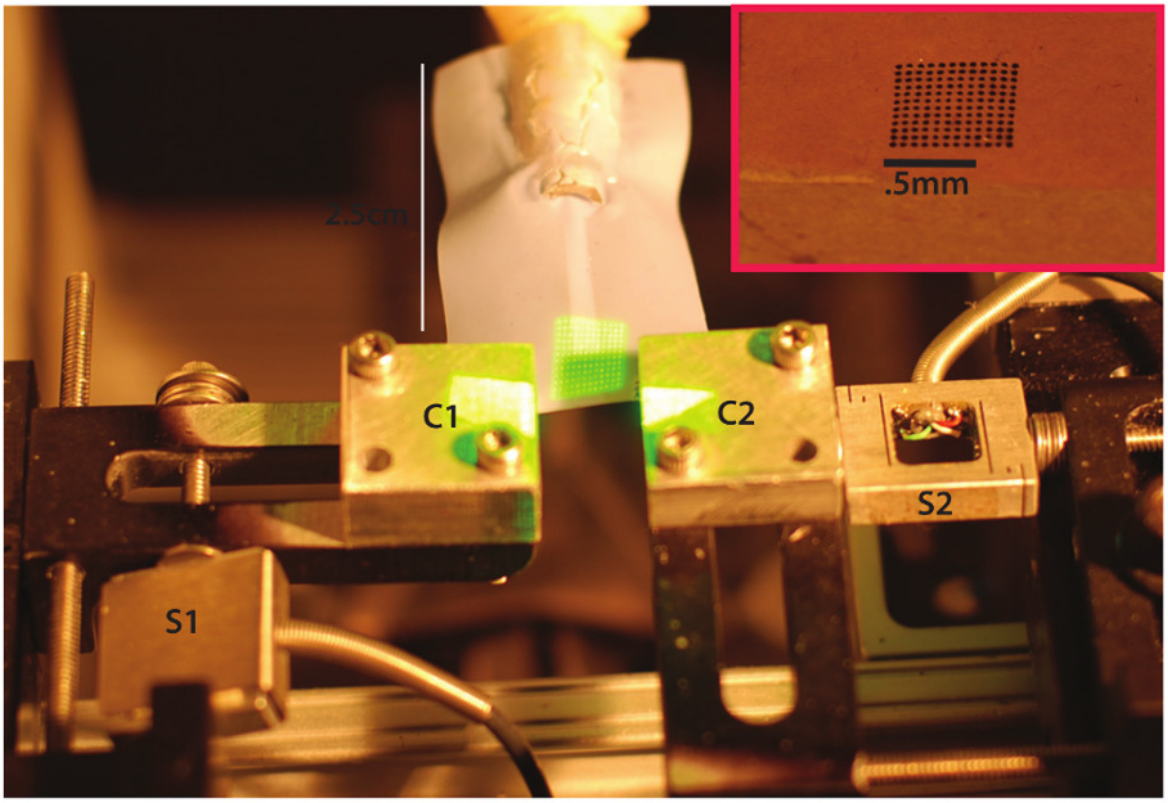

Figure S 1. Close up of the biomimetic syrinx shows a grid projected onto the device. C1 and C2 are clamps that are connected to strain gauges S1 and S2 in order to measure the (static) longitudinal and lateral strains on the syrinx. The inset shows the pattern that was used to generate the grid which was itself created by laser cutting holes of diameter $0.25mm$ using a Versalaser laser cutter.

of the emitted sound were determined primarily by the refractory period, which in turn were controlled by the position of the probe. If the strain at the point of contact is $\epsilon \sim (z/r)^2$, where z is probe depth and r is the radius of the tube. The typical frequency then scales as $\sim \sqrt{E\epsilon/\rho}$ where E is the Young's Modulus and shows that $\omega \sim z$, and allows us to explain the observed linear dependence of the frequency on probe depth.

In Fig S4c, we show a reconstructed solitary pulse that proceeds through three distinct stages. First,$(0.04 - 0.1ms)$ a depression starts to form at the point of probe contact and starts to move away from its point of origin. Second,$(0.1 - 0.3ms)$ the depression travels to the mouth of the device and finally$(0.3-0.8ms)$ there is a rest period after which the cycle repeats.

A final mode of oscillation that is particularly relevant to mimicking songs of a zebra finch is associated with the coupling of the solitary pulses to localized modes at the mouth of the device in response to the airflow without any probe. Fig S4d shows an example of this induced by depicting the height of the device $0.5cm$ from the mouth for various probe indentations. As the probe depth was increased (position b in Fig S4d), a solitary pulse traveled to the mouth and exited the system, and was then followed by oscillations at the lip of the device. The existence of this mode was visualized by simultaneous imaging of both the mouth of the cylindrical and the surface (Video S5). Fig S4e is a spectrogram of the sound produced by this mode and shows the dense harmonic structure that is characteristic of a zebra finch song motif.

DYNAMIC PRODUCTION AND CONTROL OF SONG

We discovered that by dynamically positioning the depth of the probe, it was possible to turn on the different modes described above while also controlling the frequency. This allowed us to mimic simple songs across many song bird species. To implement this protocol, we programmed the linear motor to produces "pulses" of motion, and the shape of these pulses then controlled the spectrum and duration of sound produced. Typically, we found that it was sufficient

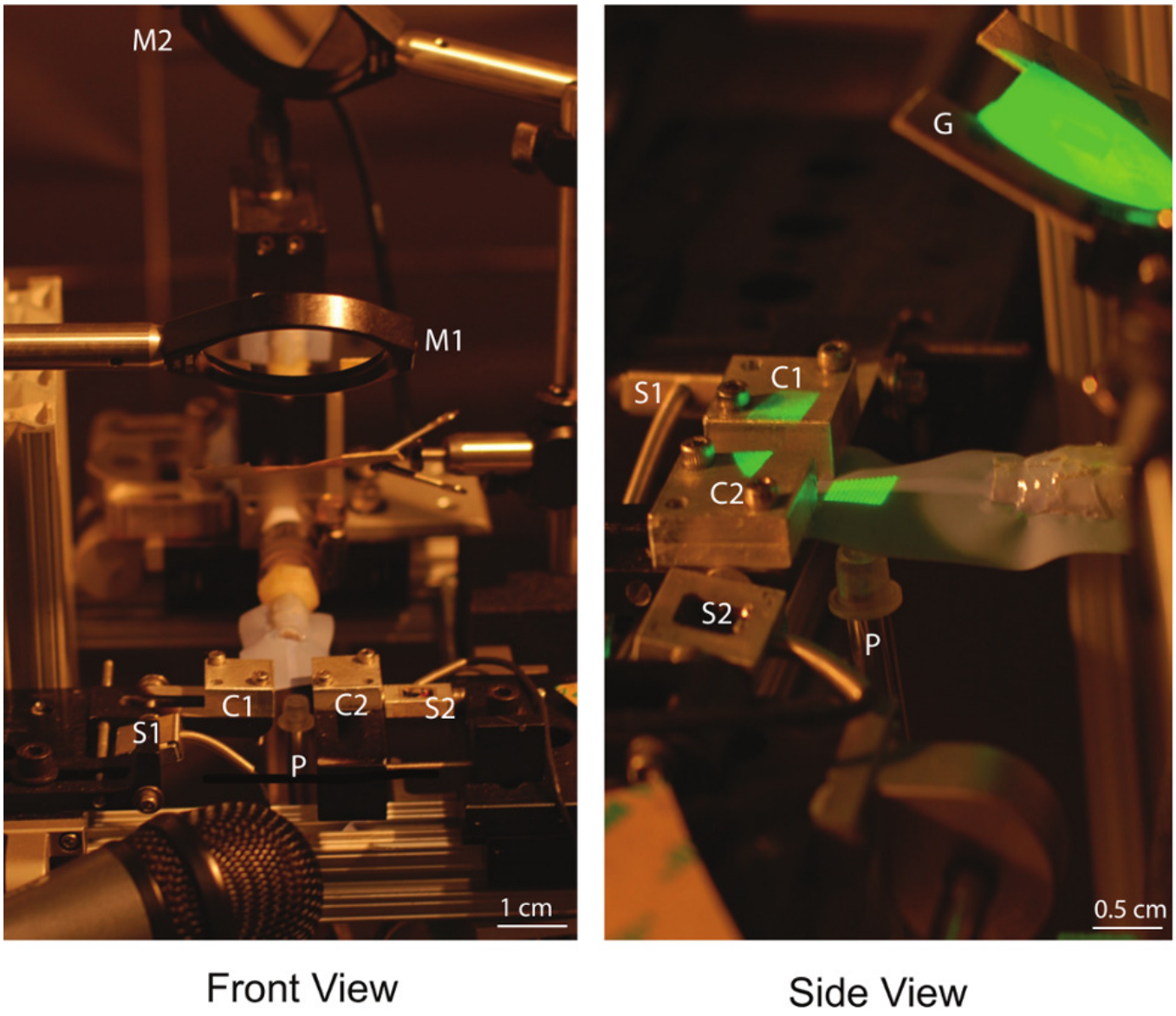

Figure S 2. Front and Side view of the measurement set up. C1 and C2 are clamps that are connected to strain gauges S1 and S2 in order to measure the (static) longitudinal and lateral strains incurred by the device. P shows the linear motor that acts as an "artificial muscle" and is used to statically and dynamically probe the device. M1 and M2 are mirrors that are used to project the image of the grid on the device to a high speed camera. At the bottom is a microphone used to record the sounds generated.

to keep the flow rate, pretensions, and x-position of the probe constant and only pulse the linear motor to produce a complete song. Fig S5 shows two simple songs along with the pulses used to program the linear motor. In the first song, the "notes" were rapidly pressed and held down for $350ms$ while in the second, a rapid sequence of sound pulses were produced. Not only can the fundamental frequency of the sound produced be controlled by the depth of the linear motor, dynamical transitions like transitions to period doubled sounds can also be induced.

VIDEO CAPTIONS

Video S1: Tonal mode of vibration. The linear probe was placed 1cm from the mouth and pressed 0.5mm into the it, the flow velocity was set at 5m/s. The movie shows a traveling wave of a sinusoidal nature that starts at the point of the linear probe and travels to the mouth. The color legend shows the height of the device in steps of $100\mu m$.

Video S2: Period doubled mode. The linear probe was placed 1cm from the mouth and the pressed 1mm into the it, the flow velocity was set at 5m/s. The movie shows a traveling wave of a sinusoidal nature that starts at the point of the lin-

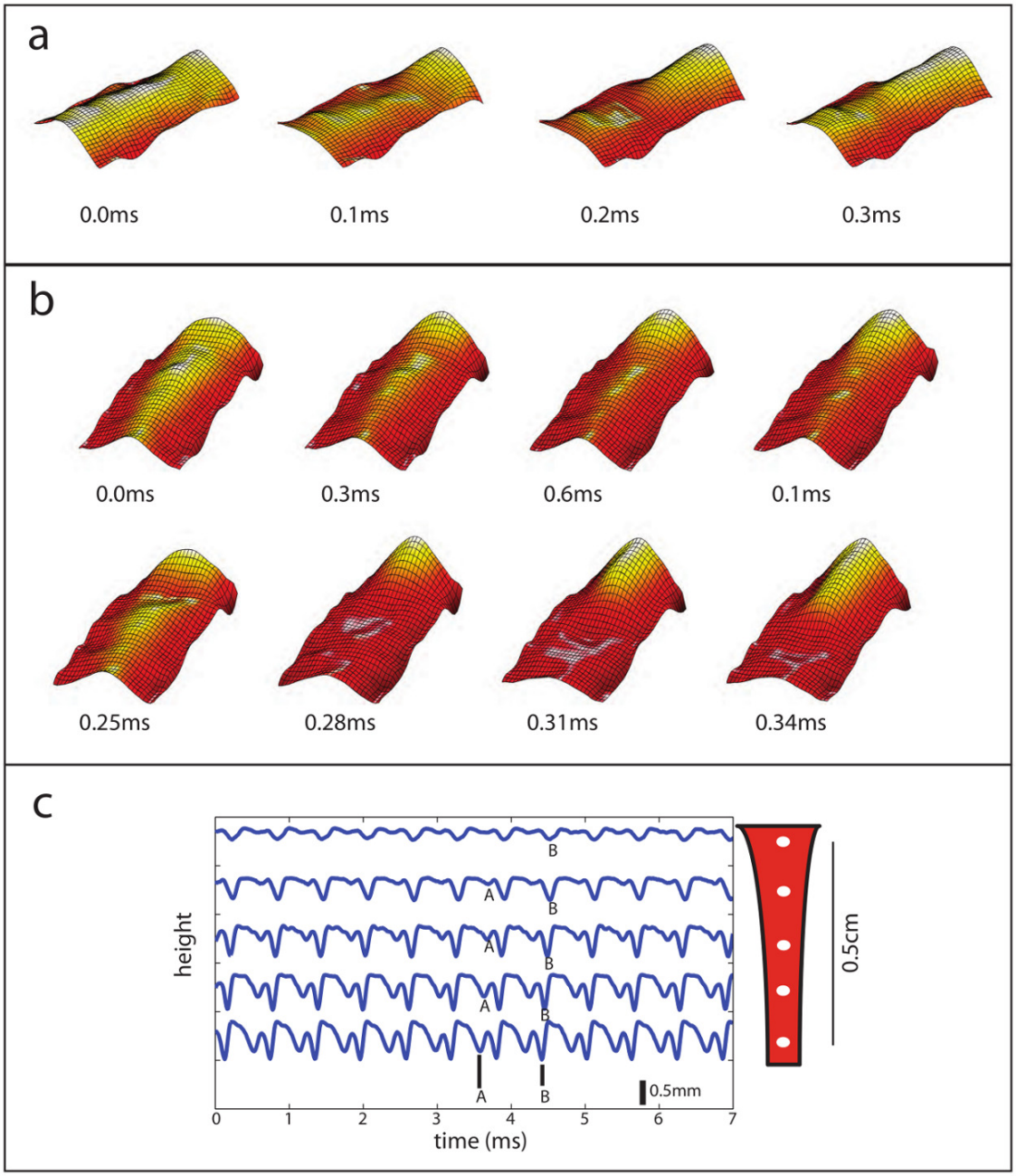

Figure S 3. (a) shows the 3d reconstructed image of the tonal mode of vibration. In this case the probe was placed 0.5cm away from the mouth, and the tonal mode appeared as gentle undulations of the surface of the tube. (b) Period doubled mode- This mode shows a lower fundamental frequency that than the corresponding tonal mode (not shown, the original tonal mode had time period $0.2ms$), that arises because a shallow solitary solitary wave ($0.25 - 0.34ms$) is created right after a sinusoidal variation has passed, thereby extending the total period of the mode. (c) Shows a time trace of the height of the device at different points along its length, delineated by the white dots of the sketch of the device (to the right of the plot). The sinusoidal variation is shown by the bump at A and the solitary pulse is shown in B. Note that in order to see this mode one had to press the motor 1cm from the mouth. As a result the tonal mode did not make it all the way to the mouth.

ear probe but does not travel to the mouth, after which a wave of a solitary kind is created from the position of the probe. The bump created in this case is not as deep as a true solitary mode. This cycle repeats after this.The color legend shows the height of the device in steps of $100\mu m$.

Video S3: Solitary pulse mode. The linear probe was placed 1cm from the mouth and the pressed 2mm into the it, the flow velocity was set at 5m/s. The movie shows a de-

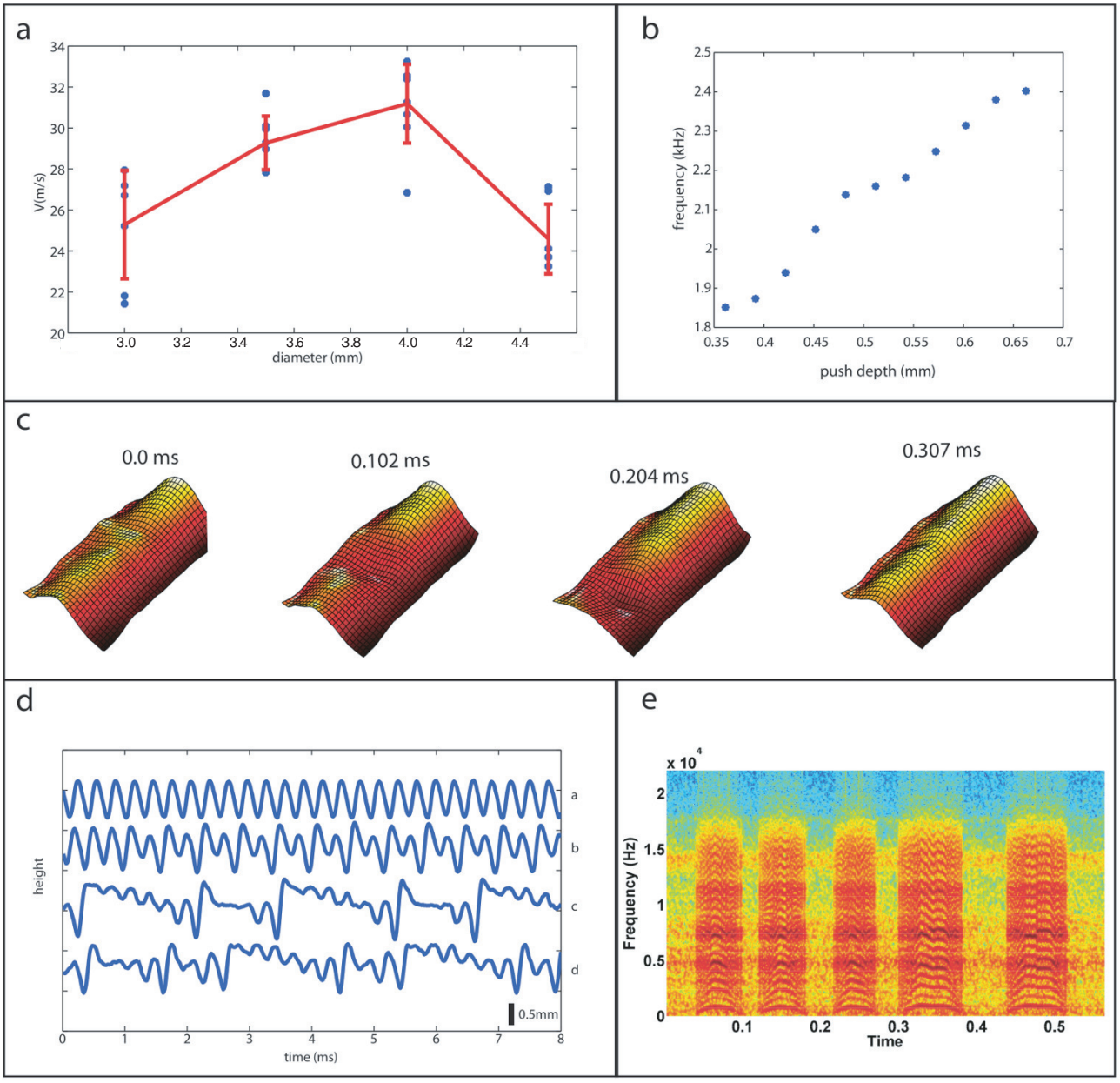

Figure S 4. (a) Wave speed of the solitary mode as a function of radius of device. For each radius flow velocity and push depth was changed to generate the data points. The speed is measured by first extracting the midline from reconstructed 3d image from each time frame and place next to each other to create a kymograph. The kymograph is the digitized and the solitary wave shows up as a straight line. The slope of this line gives the wave speed. (b) Fundamental frequency of the solitary mode as a function of push depth. The frequency changes by a factor of 1.5 when push depth is changed by a factor of 2. This range is further increased when the sound is dynamically created. (c) 3d reconstructed images of the solitary mode. A solitary bump is created close to the probe position (below the surface being imaged)which then travels to the mouth. After the wave passes there is waiting period after which another pulse is created. The frequency is controlled by changing the time interval of the waiting period. (d)Shows height at one point on the surface as a function of time for different push depths (increasing from a-d). In a) sinusoidal mode is observed, b) a period doubled mode c)-d) higher period mode created by interaction between the solitary mode and tension modes at the mouth of the device. The sound created by this mode is rich in harmonic content and some sample sounds are shown in the spectrogram in (e).

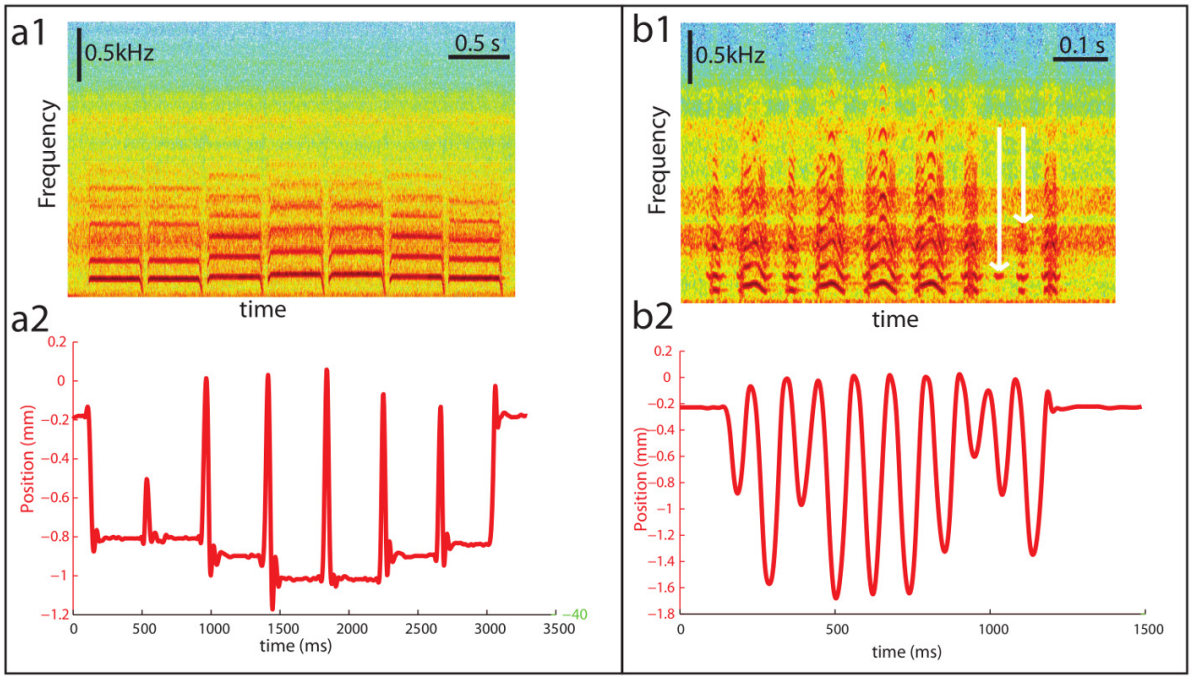

Figure S 5. Simple Songs (a1) shows the the spectrogram of a simple song that was created by pressing down and holding a note for 0.35s and (a2) shows the corresponding pulse sequence (position of the probe) used to program the song. (b1-2) shows the spectrogram and programmed pulse sequence for a series of rapid pulses with high harmonic content. A transition to non-period doubled sound occurs when the probe position is less than 0.5mm (shown by the white arrows)

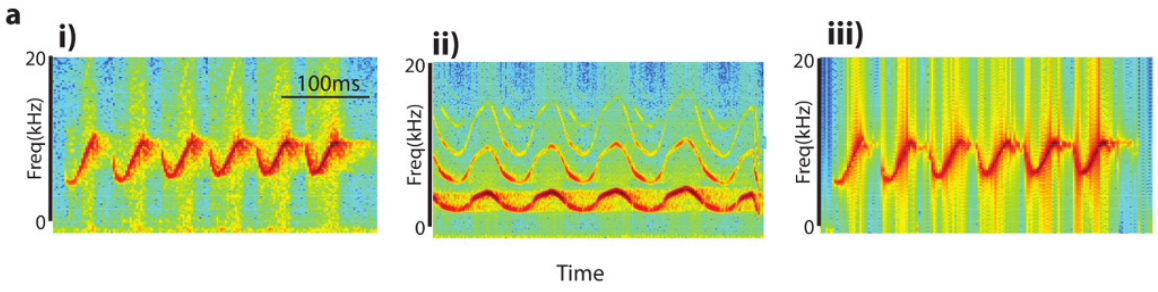

Figure S 6. Spectrograms of i)Original, ii)mechanical and iii) computational songs of a representative song of a Red Eyed Vireo.

pression being created at the location of the linear probe that consequently travels to the mouth. This is followed by a short rest period after which the cycle repeats. The color legend shows the height of the device in steps of $100\mu m$.

Video S4: Projected and deformed grid. This movie shows a raw image of the deformable grid projected onto the vibrating device.

Video S5: Close up of solitary pulse mode. We see the surface and mouth of the device imaged by placing a small mirror in front of the mouth. The mode shown is similar to the one described in the text where the solitary wave passes followed by subsequent oscillations at the mouth, before the next cycle repeats.

AUDIO CAPTIONS

Audio S1: This audio segment demonstrates a rapid transition from periodic to chaotic sound (see Main Text Fig 2b) driven by a large change in the probe indentation depth.

Audio S2: (a) Vireo song. (b) Biomimetic Vireo song (Main Text fig 2d) was created simply; in the first part the notes were just pressed and held, while in the second part the frequency was moderately swept. It was observed from the sound spectrogram that the fundamental frequency roughly followed the position of the probe. The probe was placed 0.5cm from the mouth, with $s_x = .1$ and $s_y = 0.05$ and U =5m/s.

Audio S3: (a) Bengalese finch song. (b) Biomimetic Bengalese Finch song (Main Text fig 2e). To recreate this song the device was first prepared deep in the solitary pulse phase. Then the individual notes were created by probing the motor to different depths. The song started with a short section of sound followed by a series of short pulses slowly descending in frequency, and ended with a short section where the fundamental frequency of the multiharmonic song varied a little. To produce this song the linear motor was placed $1.2cm$ from the mouth and the airflow set to $5m/s$. The exact values of pre-strains was not critical as long as it was finite ($s_x \sim 0.1$, $s_y \sim 0$).

Audio S4: (a) Zebra finch song. (b) Biomimetic Zebra Finch song (Main Text Fig 2f). It combines the high acceleration sweeps associated with the tonal song of the vireo along with sudden changes in harmonic mode structure seen in the Bengalese Finch. To capture this, we broke the song into 5 segments based on their features. Segment 1 starts with a transition from an almost pure tone to a multi harmonic downsweep by accelerating the probe rapidly from rest in the presence of a minimal pretension $s_x < .1$, and a flow speed of $\approx 5m/s$. Segment 2 shows transitions to chaos with some structure of a rapid downsweep. To mimic this section, the flow velocity was increased to $7m/s$ to find a region of coexistence of the solitary pulse and chaotic modes, and then apply a probe pulse sequence to approximate the downward sweep in frequency. Segment 3 is similar to segment 2, except there is a short wait before the downsweep. Segment 4 starts with a rapid downsweep and then suddenly holds at a particular frequency. This sudden acceleration and deceleration of the motor driving the probe created many undamped transient responses both from the device and the linear probe itself; as result of which the spectrogram had small undamped oscillations. Segment 5 is similar except there is a down-up-down sweep at the beginning. Again the oscillations in the motor did not allow us to wait steadily at one particular frequency.

Audio S5: (a) Vireo song. (b) Computational mimic of Vireo song (See Main Text Fig. 3).

Audio S6: (a) Vireo song. (b) Biomimetic Vireo Song. (c) Computational mimic of Vireo song (See Fig.S 6).

Audio S7: Trill sound. Produced by rapidly pulsing the linear motor, that acts as a switch, turning on and off a periodic song at a high rate ($30Hz$).