Universal mechanisms of sound production and control in birds and mammals

C.P.H. Elemans, University of Southern Denmark
J.H. Rasmussen, University of Southern Denmark
C.T. Herbst, Palacky University
D.N. Düring, University of Southern Denmark
S.A. Zollinger, Max Planck Institute for Ornithology
H. Brumm, Max Planck Institute for Ornithology
K. Srivastava, Emory University
N. Svane, University of Southern Denmark
M. Ding, University of Southern Denmark
O.N. Larsen, University of Southern Denmark

Only first 10 authors above; see publication for full author list.

Journal Title: Nature Communications
Volume: Volume 6
Publisher: Nature Publishing Group: Nature Communications | 2015-11-27, Pages 8978-8978
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1038/ncomms9978
Permanent URL: https://pid.emory.edu/ark:/25593/rgk41

Final published version: http://dx.doi.org/10.1038/ncomms9978

Copyright information:
© 2015, Nature Publishing Group, a division of Macmillan Publishers Limited.

This is an Open Access work distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).
Universal mechanisms of sound production and control in birds and mammals

C.P.H Elemans1,2,3, J.H. Rasmussen1,*, C.T. Herbst4,*, D.N. Düring1,*, S.A. Zollinger5, H. Brumm5, K. Srivastava6,7, N. Svane1, M. Ding8, O.N. Larsen1, S.J. Sober6 & J.G. Švec4

As animals vocalize, their vocal organ transforms motor commands into vocalizations for social communication. In birds, the physical mechanisms by which vocalizations are produced and controlled remain unresolved because of the extreme difficulty in obtaining in vivo measurements. Here, we introduce an ex vivo preparation of the avian vocal organ that allows simultaneous high-speed imaging, muscle stimulation and kinematic and acoustic analyses to reveal the mechanisms of vocal production in birds across a wide range of taxa. Remarkably, we show that all species tested employ the myoelastic-aerodynamic (MEAD) mechanism, the same mechanism used to produce human speech. Furthermore, we show substantial redundancy in the control of key vocal parameters ex vivo, suggesting that in vivo vocalizations may also not be specified by unique motor commands. We propose that such motor redundancy can aid vocal learning and is common to MEAD sound production across birds and mammals, including humans.
In contrast to laryngeally vocalizing mammals, ~10,000 species of extant birds vocalize with a uniquely avian vocal organ, the syrinx, located at the tracheobronchial junction and suspended in an air sac of the respiratory system

The syrinx is structurally highly diverse across species, but how morphological diversity reflects functional diversity remains unexplored. In addition, while songbirds are a widely used experimental animal model for neural mechanisms underlying vocal imitation learning

we lack the empirical evidence to precisely map motor function onto neural circuitry. Addressing these questions requires empirical quantification of syringeal dynamics as a function of control parameters in different species under physiologically realistic, controlled conditions. However, imaging the syrinx in vivo remains a challenge

and we thus still lack this quantification of syringeal dynamics and control parameters.

Earlier endoscopic imaging identified syringeal vibratory tissues in songbirds and non-songbirds, arguing against purely aerodynamical whistle mechanisms in which sound is produced without periodic movement of the vocal apparatus. Many mathematical models for birdsong have thus far not established that syringeal sound production is based on a myoelastic-aerodynamic (MEAD) system. However, conclusive empirical evidence for MEAD is lacking.

The MEAD framework explains the physical mechanism underlying laryngeal sound production in mammals, preventing the need for muscle contractions at the rate of tissue vibration or other periodic input. Expiratory airflow is mechanically converted by vocal folds into pulse-like airflow, which causes air pressure disturbances constituting the acoustic excitation of the system. The mechanical properties and recruitment of different layers of vibrating tissues affect their resonance properties, which in combination with aerodynamic driving forces determine the frequency and mode of oscillation.

According to MEAD theory, the medio-lateral vibration of the inner vocal fold surface (vibrational component 1 (VC1)) that gates airflow can only be self-sustaining if moving around a stable equilibrium position and if no net energy loss occurs per oscillation cycle. The latter requires the presence of an aerodynamic force that changes magnitude with direction of vocal fold motion and thus is asymmetric over the oscillation cycle. One possibility is that during self-sustained vocal fold vibration the required asymmetric aerodynamic force is produced by time-varying supraglottal pressure due to inertia of air in the vocal tract. Such a mechanism would result in uniform medio-lateral vibration of the vocal folds. However, models suggest that this mechanism limits the range of fundamental frequencies (F0) produced. Another, more robust possibility is that the asymmetric aerodynamic force during self-sustained vocal fold vibration is produced by out-of-phase motion of the superior and inferior edge of the vibrating tissue (vibrational component 2 (VC2)). VC1 and VC2 are the respective medio-lateral and caudo-cranial components of a tissue surface wave, or mucosal wave, that travels on the inner vocal fold surface along the expiratory air stream and facilitates aerodynamic energy transfer into tissue. The wave phase changes cause the vocal folds to change shape from convergent during opening to divergent during closing parts of the cycle. Because the intraglottal pressure is higher for the convergent shape than for the divergent shape, the vocal folds are pushed apart during opening and pulled together during closing. VC2 presence thus indicates that intraglottal pressure forms the asymmetric forcing function over opening and closing phases of vibration essential to self-sustained oscillation. Sound excitation events in mammals occur mainly at glottal closure and/or opening, when airflow abruptly stops or starts.

Although isolated aspects of syringeal dynamics have been studied in birds, the asymmetric forcing function essential to maintain self-sustained oscillation has not yet been identified, and the caudo-cranial tissue-wave component VC2—a crucial underlying assumption in modelling studies—has not been demonstrated experimentally in the intact syrinx under appropriate physiological conditions. Furthermore, it is unknown how syringeal dynamics relate to sound generating events within a single oscillatory cycle. These essential features of syringeal dynamics required to confirm MEAD have yet to be established.

The translation of vocal motor commands into acoustic output depends on neural activity, musculature, morphology and physical mechanism of sound production. Many variables have been studied in vivo by correlating acoustical parameters, for example, F0, with physiological parameters, for example, lung pressure or muscle activity, taking advantage of the highly stereotyped vocal patterns employed by adult birds. This approach can inform us about a particular solution an individual uses to control its vocal output. However, if the control space is redundant (that is, there is more than one possible solution to achieve a specific vocal target, for example, F0), as commonly observed in motor control systems, studying stereotyped in vivo behaviour provides limited insight in the behaviour of the entire system as we only observe the final solution the individual uses that may not be unique. Furthermore if vocal control parameters covary it may be difficult to establish causal relationships. To understand how the brain controls vocal behaviour therefore requires systematic quantification of the system’s behaviour across its multi-dimensional parameter space. However, we currently lack an experimental paradigm to systematically study physiological control of the vocal system in birds.

Here, we present a novel ex vivo paradigm of the syrinx, which allows unprecedented experimental control and high-resolution imaging during sound production. First, to investigate if the MEAD physical mechanism of self-sustained oscillations as observed in mammals is applicable to birds, we test the hypotheses that a caudo-cranial travelling tissue surface wave is present and that this tissue wave is associated with sound production events. Second, to determine whether vocalizations are encoded by unique motor commands we test the hypothesis that the physiological control space of the syrinx ex vivo is redundant for key acoustic parameters. We show that birds employ the MEAD mechanism for sound production and that key vocal parameters exhibit redundant control ex vivo. We propose that motor redundancy may accelerate vocal learning and is common to MEAD sound production across birds and mammals.

**Results**

**Physical mechanism of sound production.** We developed an experimental paradigm that allows imaging of syringeal dynamics under controlled conditions ex vivo (Methods section; Supplementary Fig. 1). To test our first hypothesis, if the MEAD physical mechanism of self-sustained oscillations is applicable to birds (Fig. 1a), we took advantage of the diversity in syringeal morphology across species and first studied the domestic pigeon syrinx (Fig. 1b) because of its relatively simple...
morphology. We found self-sustained syringleal oscillations when both bronchial and air sac pressures were >0.5 kPa (N = 12). The lateral vibratory masses (LVMs) were visualized using transillumination of the syrinx (Fig. 1c). Within an oscillation cycle the LVM inner wall changed shape from divergent during closing, rectangular during full collision and to convergent during opening (Fig. 1c, Supplementary Movie 1). A consistent phase shift in LVM position, confirmed by
Figure 1 | The MEAD theory explains sound production in the domestic pigeon (Columba livia) syrinx. (a) Schematic illustration of vocal organs in mammals (larynx) and birds (syrinx) showing the two vibrational components essential to MEAD theory: medio-lateral oscillation (VC1), and caudo-cranial phase differences (VC2). The latter are related to a tissue surface wave moving cranially and cause the inner wall of the vibratory tissue to be divergent during closing (solid line, black arrows) and convergent during opening motion (dashed line, white arrows). The syrinx is suspended in the ICAS. Three pressures act during self-sustained oscillations over a range of F0 values (Fig. 5). In the pigeon, both opening and closing events were precisely accompanied by an acoustic excitation at very short delays of 170 and 90 μs, respectively (Fig. 5a), which are both below the 250 μs temporal accuracy of the opening and closing event timing (that is, one frame duration of high-speed video). The timing showed a strong acoustic excitation on syringeal opening, with a delay of 100 μs, also below the 250 μs temporal accuracy of the opening and closing event timing (Fig. 5b). A second, often weaker, excitation occurred 1.63 ms after closing. The zebra finch showed a very precisely timed strong acoustic excitation on syringeal closing at a delay of 40 μs (at a temporal accuracy of 33 μs; Fig. 5c).

In conclusion, these data confirm our first hypothesis that a caudo-cranial travelling tissue surface wave is present across a range of syringeal morphologies and sizes. Furthermore, our data confirm a close association between opening/closing event timing and sound generation events within single oscillatory cycles across a range of species. Collectively, our data thus provide the essential lacking demonstrations of syringeal dynamics and sound generation events required to conclusively show that MEAD theory is applicable to sound production in birds.

Vocal control redundancy ex vivo. To test our second hypothesis that the physiological control space of the vocal organ is redundant for key acoustic parameters, we systematically investigated the relationship between syringeal control parameters and acoustic output ex vivo. In mammals, the brain can control bronchial pressure and laryngeal muscle activity to initiate peripheral airflow and produce sound (Figs 6 and 7a–c). For each species the F0 ranges produced ex vivo corresponded well to the lower-end distribution of spontaneous vocalizations (Figs 6 and 7b,c). We further found that in each investigated species, multiple different combinations of bronchial and ICAS pressures could achieve the same target frequency (iso-F0 contours in Figs 6 and 7c), indicating that F0 control is redundant within the pressure control space. To investigate how much two other important vocal parameters, that is, sound pressure level (SPL) and sound quality, changed with F0, we quantified the SPL and Wiener entropy (WE) along evenly
The F0 range achieved by pressure increase of 300–750 Hz and 75–200 Hz for tinamou and zebra micro-stimulation (Fig. 7d). Muscle stimulation caused an F0 resulting in the same F0. These results thus demonstrate that muscular control of F0 is redundant in the ex vivo preparation. Taken together, the above data confirm our second hypothesis that the physiological control space of the vocal organ ex vivo is redundant for key acoustic parameters across a range of syringeal morphologies and sizes.

Discussion

Our data establish that birds use MEAD as the primary physical mechanism for sound production with strong similarities to mammalian MEAD systems: First, we demonstrate the presence of a tissue wave that travels from the caudal to cranial end of the syringeal vibratory tissue (Figs 1–4). This tissue wave thus causes the syringeal vibratory tissue shape to be convergent when opening, and divergent when closing during expiratory sound production, and can be considered analogous to the caudo-cranial mucosal wave observed in mammalian vocal folds. Our data therefore strongly suggest that in birds, just as in mammals, the dominant asymmetric forcing function over the opening and closing of the syrinx at two different phases during the oscillatory cycle. The region and location of these stills is indicated with a dotted box in a. For the ostrich the position of VKG analysis (dotted horizontal line) and DKG analysis (solid white line from point 1–2) presented in panel e is indicated. Transillumination provides sufficient detail to identify inner edges of right (green) and left (red) LVM in tinamou. (c) Evidence for medio-lateral vibrations (VC1) using tracheal endoscopic VKG. In the ostrich predominantly the right side vibrated. Full opening and closure of the cranial edge can be observed in tinamou. (d) Evidence for the caudo-cranial component (VC2) of the tissue wave as demonstrated by DKG for ostrich, and glottovibrogram for tinamou. (e) Synchronised acoustic waveforms.
closing phases of vibration essential to maintain self-sustained oscillation is not formed by the mass inertance of the air column in the vocal tract, but by the tissue-wave-induced intraglottal pressure changes. This mechanism reduces the dependency of the self-sustained syringeal oscillations on acoustic resonances of the vocal tract and allows for an expanded F0 range of vocalization, which in addition to labial morphology could aid birds in extending their F0 range. Second, the magnitude range of the tissue-wave speed measured (0.5–3.0 m s$^{-1}$) is in excellent agreement with values reported for the mammalian larynx (human, dog, calf) and an African elephant (0.4 m s$^{-1}$), and in humans the maximum acoustic excitation has been shown to occur mainly at the instant of glottal closure, a recent study reported that in an excised elephant larynx the maximum acoustic excitation occurred at the instant of glottal opening. While in humans the mammalian larynx thus exhibits more diversity than previously thought in what movements cause the predominant acoustic excitation. Our work demonstrates that this diversity is also present in birds and that more comparative studies are needed to explain the causal link between airflow, tissue vibration and acoustic excitation.

In conclusion, we find that despite the large diversity present in syringeal morphology of the bird species included in this study, all use MEAD as the primary physical mechanism of sound production, supporting MEAD-based approaches to modelling avian sound production. Moreover, our findings suggest that despite their different evolutionary origins, laryngeally vocalizing mammals and syringeally vocalizing birds have converged on the same physical mechanism for vocalization.

Furthermore, our data show that key acoustic parameters, such as particular F0 values, can be achieved by multiple distinct combinations of respiratory pressure and syringeal muscle vibration and generation of acoustic energy by modulating the glottal airflow as found in the mammalian larynx. While in humans the maximum acoustic excitation has been shown to occur mainly at the instant of glottal closure, a recent study reported that in an excised elephant larynx the maximum acoustic excitation occurred at the instant of glottal opening. The mammalian larynx thus exhibits more diversity than previously thought in what movements cause the predominant acoustic excitation. Our work demonstrates that this diversity is also present in birds and that more comparative studies are needed to explain the causal link between airflow, tissue vibration and acoustic excitation.

In conclusion, we find that despite the large diversity present in syringeal morphology of the bird species included in this study, all use MEAD as the primary physical mechanism of sound production, supporting MEAD-based approaches to modelling avian sound production. Moreover, our findings suggest that despite their different evolutionary origins, laryngeally vocalizing mammals and syringeally vocalizing birds have converged on the same physical mechanism for vocalization.

Furthermore, our data show that key acoustic parameters, such as particular F0 values, can be achieved by multiple distinct combinations of respiratory pressure and syringeal muscle vibration and generation of acoustic energy by modulating the glottal airflow as found in the mammalian larynx. While in humans the maximum acoustic excitation has been shown to occur mainly at the instant of glottal closure, a recent study reported that in an excised elephant larynx the maximum acoustic excitation occurred at the instant of glottal opening. The mammalian larynx thus exhibits more diversity than previously thought in what movements cause the predominant acoustic excitation. Our work demonstrates that this diversity is also present in birds and that more comparative studies are needed to explain the causal link between airflow, tissue vibration and acoustic excitation.
of pressure differentials, only bronchial pressure in combination with different levels of recruitment in a single vocal muscle leads to redundancy in F0 (Fig. 7e).

How acoustic redundancy of the vocal organ ex vivo reflects the functional motor redundancy in vivo remains to be explored. The speed at which consecutive motor commands can be executed in vivo (either on the same side of the syrinx or when vocal production switches rapidly from one side to the other within a complex syllable) may cause a significant reduction of the available functional redundancy, especially because at least some songbirds possess superfine syringeal muscles that produce peak force in <5 ms (refs 54,55). Here, we did not investigate the rapid time-varying pressure and muscle recruitment patterns that occur in vivo, and indeed for most species more detailed experimental in vivo data are required to meaningfully explore the high-dimensional control space ex vivo. On the other hand, we show that redundancy emerges when only varying the activity level of a single vocal muscle. Redundancy can thus be expected to increase when the brain can produce a given acoustic output by choosing from a large set of redundant motor commands when coordinating the ~16 syringeal muscles with several thousand motor units.

A recent study used zebra finch song acoustics to infer low-dimensional motor commands, which were then used to generate synthetic songs that closely resembled the acoustic output of natural vocal behaviour, and evoked auditory responses that closely resembled the pattern of premotor activity. These results led the authors to suggest that premotor area HVC encodes a low-dimensional forward model of gesture dynamics. In contrast, our data indicate that particular acoustic outputs can be specified by multiple motor states, suggesting that motor commands cannot be uniquely inferred from acoustics. It will therefore be important for future studies to investigate whether the method used in ref. 11 to infer motor commands from acoustic data yields the actual combination of control parameters employed by the bird to generate the sound, rather than a different set of control parameters that redundantly produces the same acoustic output. Similarly, future ex vivo studies employing time-varying stimulus patterns across multiple muscles will allow us to explore the limits of vocal redundancy during sophisticated motor trajectories similar to those in behaving animals. Despite the apparent conflict between our findings and elements of ref. 11, it is important to note that the instantiation of a low-dimensional forward model within HVC—which is multiple synapses upstream from the vocal and respiratory muscles—is not incompatible with a redundant vocal organ. Indeed, one strength of low-dimensional encoding in HVC is that it would allow HVC to represent control parameters (for example, vocal fold tension) without regard to the details of the redundant combinations of muscle contraction states required to achieve that tension.

In many fine motor control systems such as arm reaching, the brain must negotiate the so-called motor redundancy problem, in which a particular behavioural target can be achieved by vast numbers of motor commands. However, one potential advantage to a redundant control space is that it allows the brain to find subspaces of possible motor commands (for example, along the iso-F0 contours in Figs 6 and 7c in our ex vivo paradigm) rather than searching for unique motor commands to achieve the target. Crucially, redundant control spaces allow variability in task-irrelevant directions. In vocal production, the brain may select from within this motor command space to meet other demands, such as acoustic targets in future and/or past syllables. Consequently, we speculate that vocal motor redundancy may simplify trial-and-error learning during song acquisition by allowing the brain to rapidly discover one of many motor

---

**Figure 4** | The caudo-cranial tissue wave is present across a range of fundamental frequencies. Magnitude of the caudo-cranial component ($V_{cc}$) of the travelling tissue-wave’s velocity in (a) pigeon and (b) tinamou, (c) zebra finch and (d) cockatiel. The positive velocity values indicate that the wave travelled from caudal to cranial. Wave speed remained constant with F0 in pigeon and tinamou (linear regression, $P = 0.56$ ($n = 10$) and $P = 0.26$ ($n = 10$), respectively), but increased significantly with F0 in zebra finch and cockatiel (linear regression, $P < 0.001$ ($n = 11$) and $P < 0.001$ ($n = 118$), respectively).
solutions before learning how to further improve performance by exploiting motor redundancy. Vocal redundancy may therefore have aided, or even been necessary for, the evolutionary development of vocal learning. Future studies might evaluate this speculation by quantifying whether learned vocalizations exploit vocal redundancy. Furthermore, because many species do not learn their vocalizations and many vocal learners produce innate vocalizations in addition to learned ones, future work might also determine whether innate vocal motor programs similarly exploit redundancy to improve performance.

Our ex vivo results suggest that the redundant control of key vocal parameters represents a significant aspect of avian vocal control, and one that merits further investigation. Because redundancy in F0 control by pressure and muscle recruitment is also observed in various simplified computational models of the human larynx, and songbird syrinx as well as in ex vivo dog larynx preparations, we propose that vocal control redundancy is a typical feature of MEAD sound production systems and hence a common feature of vocal production and control in mammals and birds.
Figure 6 | Pressure control of F0 in sound production is redundant across avian taxa. Spectrograms and oscillograms of spontaneous vocalisation (left) and representative pressure control spaces ex vivo (right) of (a) pigeon (Columba livia), (b) Barbary dove (Streptopelia risoria), (c) cockatiel and (d) Bengalese finch (Lonchura striata domestica). The F0 ranges of sounds produced ex vivo (grey vertical bar right of spectrogram) correspond well to the lower range of spontaneous vocalizations. Pressure control may allow higher F0 values, as indicated by the grey bar fading out. Iso-F0 contours (values in Hz) are overlaid on the pressure control spaces and shown in white when redundant for all three acoustics parameters (F0, SPL and WE), or black when not. In the Bengalese finch, the left hemisyrinx produces higher F0 than the right hemisyrinx, corroborating earlier in vivo studies. (e) Spectrogram and oscillogram (left) of ostrich mating call, and ex vivo sound production including bronchial pressure (p_b) patterns underlying phonation (right). A complete pressure control space was not obtained for this species.
Figure 7 | Motor control of F0 in sound production is redundant. (a) Three-dimensional geometry of the tinamou and zebra finch syrinx with pressure control parameters: bronchial or subsyringeal ($p_{\text{sub}}$), and ICAS pressure ($p_{\text{ICAS}}$). Stimulated muscles are indicated. (b) Sound spectrogram (top) and oscillogram (bottom) of a tinamou honksqueal call and zebra finch song motif. (c) Representative pressure control spaces ex vivo of one individual. Iso-F0 contour as in previous figure (values in Hz). (d) Sound spectrogram (top) and muscle stimulus (bottom) during syringeal muscle stimulation. Black dots indicate measurements points. (e) Fundamental frequency (F0) is affected by muscle stimulation (blue dots, no stimulation; red dots, maximal stimulation) and pressure. The grey bar indicates the individuals’ frequency range achieved by varying only pressure. Pressure control may allow higher F0 values, as indicated by the grey bar fading out. Multiple different combinations result in the same frequency, for example, 500 and 650 Hz in tinamou and zebra finch, respectively, (green horizontal line).
Methods

Subjects. To study sound production mechanisms and the effect of muscle stimulation on sound generation, we used 12 adult domestic pigeons (*Columba livia*, order Columbiformes; nine males, three females), four domestic Barbary doves (*Streptopelia risoria*, order Columbiformes; three males, one female), two ostriches (*Struthio camelus*, order Struthioniformes; sex unknown), three cockatiels (*Nymphicus hollandicus*, order Psittaciformes; two males, one female), six Bengalese finches (*Lonchura striata domestica*, order Passeriformes; six males and 19 zebra finches (*Taenopygia guttata*, order Passeriformes; five adult males; two juvenile males (40–400 years); 12 females). Ostrich syrinxes were obtained from a local breeding farm. They were transferred to ice immediately after extraction, flash frozen, placed in nitrogen, kept at −80°C Bengalese finches were kept in indoor aviaries on a 14-h light:10-h dark light cycle with food and water ad libitum (Emory University, Atlanta, GA, USA). Zebra finches were kept in indoor aviaries on a 12-h light:dark photoperiod with food and water ad libitum (University of Southern Denmark (SDU), Odense, Denmark). All other animals were kept in 3 × 6 × 2 m outdoor aviary with food and water ad libitum (SDU, Odense, Denmark). Pigeons, doves, cockatiels and timanous were obtained from local breeders. Bengalese finch experiments were carried out at Emory University, and all other experiments at SDU, Denmark. All experiments were conducted in accordance with the Institutional Animal Care and Use Committee of Emory University and of SDU.

Surgical and mounting procedures. Animals were euthanized with isoflurane and cooled on ice or icepacks. The syrinx and associated blood vessels were dissected out using a stereoscope (M165-FC, Leica Microsystems) through a ventral incision to the sternum while regularly flushing with oxygenated Ringer solution (5°C, recipe cf. refs 54–56), and transferred to a Sylgard-covered petri dish on ice containing oxygenated Ringer solution. The syrinx was cleaned of fat and connective tissue to expose the thoracic and abdominal wall. Connective tissue was removed by using a stereoscope (M165-FC, Leica Microsystems) through a ventral incision (that is, intact perfused organ) under controlled conditions. This experimental protocol thoroughly describes above and connected to a modified two-channel electroglossotograph (model E2G, Glottal Enterprises Inc. NY, USA).

Data acquisition and synchronization. Sound was recorded with a ½ inch pressure microphone-pre-amplifier assembly (model 46AD with preamplifier type 26AH, G.R.A.S., Denmark), amplified and high-pass filtered (10 Hz, 3-pole Butterworth filter, model 12AQ, G.R.A.S.). The microphone sensitivity was measured before each experiment (sound calibrator model 42AB, G.R.A.S.). The microphone was placed at 2–3 cm from the tracheal connector outlet in the acoustic near field, and on a 45° angle to avoid the air jet from the tracheal outlet. The sound signal was time shifted for the travelling distance from vibratory membranes to microphone. Microphone, µEGG pressure and flow signals were low-pass filtered at 10 kHz (custom-built filter). These signals together with the synchronization signals from camera systems and muscle stimulators were digitized at 50 kHz (USB 6259, 16 bit, National Instruments, Austin, TX, USA). These signals were synchronized with all imaging systems with an accuracy of <21 µs before each experiment. All control and analysis software was written in Matlab (National Instruments) or Matlab.

Tissue-wave imaging protocol and analysis. Transillumination successfully visualized the inside outlines of vibratory tissues in the domestic pigeon, Barbary dove, elegant-crested timanou and zebra finch. High optical density of the tracheal and cockatiel syrinx did not allow transillumination. In the timanou, we dissected apart *M. syringealis* to allow LVM imaging using transillumination. In zebra finch transillumination was successful in adult females and juvenile males, but attempts in adult males were unsuccessful due to the high optical density of *M. syringealis ventralis* (VS) muscles. In the zebra finch sound production was induced in the right hemistriscus. Substantially different requirements in lighting conditions did not allow for simultaneous transillumination and tracheal endoscopy in zebra finch.

To image syringeal oscillatory behaviour we subjected the syrinx to a single bronchial pressure ramps at constant air sac pressure, while filming from 5,000 to 35,000 frames s−1. Our system could acquire and save a maximum of 2–10 s of data depending on frame rates, which comprised (tens of) thousands of frames. Further, the image complexity of the data made analysis by manual measurement impossible, and required frame-by-frame manual analysis. On selected sequences of high-speed images, the left and right LVM or labia were traced manually in Amira (Visage Imaging GmbH, Berlin, Germany) and processed in Matlab. We calculated displacement (for example, Fig. 1e top panel) as current position minus the most recent position during oscillation (indicated by the green line ‘min’ on the top left in Fig. 1e). With both left and right LVM shapes quantified we could then compute the syringeal opening as a function of position and time, a graph also known as the glottovibrogram29 as can be seen in Figs 1e,2d and 3d. To quantify syringeal dynamics over a range of F0 values, full glottovibrogram reconstructions over several cycles were not suitable due to the labour-intensive manual tracing required. Therefore we used an alternative approach to determine the presence of the VC2 component and speed of the tissue wave across a range of F0 values. For pigeon, timanou and zebra finch we used transillumination frontal views of the syrinx and calculated one DKG at the caudal edge and one DKG at the cranial edge of the LVM or labia in Fig. 1d at distance A1 (indicated in middle panel Fig. 1c). DKG’s contain many periodic traces as many structures move with each oscillation. By tracing selected LVM and labial inner
edges in video stills over five oscillation cycles, we could identify the wave representing the motion of the LVM or labial inner edge in the DKG’s and as such measure the time difference between maximal lateral position of LVM or labium on the two DKG’s (ΔLVM, Fig. 1d). The wave speed equaled \( v_{C2} = \Delta T / \Delta D_{LVM} \) and was averaged over five consecutive cycles. Because in the cockatoo syrinx transilluminated frontal views were not available, we quantified cross-correlation between two \( \mu \text{EGG} \) electrode pairs located at distance \( \Delta y \) to calculate time delay \( \Delta T_{\mu \text{EGG}} \) and wave speed \( v_{C2} \). The ostrich was not included in this analysis because both transillumination was not successful and \( \mu \text{EGG} \) did not resolve vibrations due to lack of syringeal closure. We did not estimate wave speed from the tracheal endoscopic view because the VC2 velocity component was almost normal to the imaging plane, leading to large inaccuracies.

In these five consecutive oscillations, we quantified the timing of syringeal opening and closing from the transilluminated frontal views as the first frame in which the entire syringeal passage was open or closed, respectively. The minimal precision for these measurements equals the duration of one frame as indicated by the black vertical line at the base of each column in Fig. 5. In addition, we quantified the timing of acoustic excitations in the sound pressure signal. The timing of the sound signals was corrected for the time needed for the sound to propagate from the oscillators through trachea and tubing to the microphone, assuming a sound speed of 340 m s\(^{-1}\). With these parameters we quantified the timing of acoustic excitation with respect to syringeal opening and closing events within single oscillatory cycles.

**Pressure control space protocol.** To explore syringeal oscillatory behaviour in the pressure control space we subjected the syrinx to a set of bronchial pressure ramps at randomized and variable air sac pressure, and one additional ramp where \( p_b \) equalled \( p_{\text{CAS}} \). Ramp speed was 1 kPa s\(^{-1}\). Depending on species, we used pressure ranges of 0–4–0.2 kPa to avoid regimes of high flow (that is, combinations of high \( p_b \) and low \( p_{\text{CAS}} \)) and/or mechanical failure of syringeal oscillatory structures. These sets were performed without transillumination imaging. SPL and WE were calculated on 100 ms segments by applying a sliding window with 50 ms steps. If the SPL of these segments were above the 60 dB threshold, segments were zero-padded to the next power of two, and a 256-point Fast Fourier Transform (FFT) was performed. The wave speed \( v \) was computed. A point along the iso-F0 contour was considered redundant for the 100 Pa start and end sections of the contour, only forward and backward values could be computed. A point along the iso-F0 contour was considered redundant for all the three parameters if the WE and SPL variation within the 100 Pa interval was below 0.2 units of WE and 1 dB SPL (ref. 66). This resulted in a logical (1/0) redundancy array for each F0 contour. The iso-F0 contour is plotted white when redundant and black when not in the \( p_b; p_{\text{CAS}} \) control space in Figs 6a–d and 7c.

We computed iso-F0 contours at fixed intervals (\( \text{Columbia, 10 Hz; Streptopelia, 50 Hz; Eudromia, 50 Hz; Nymphicus, 100 Hz; Lonchura, 100 Hz; Tucanopygia, 50 Hz;} \) within 30%. The range was varied in the preparation to derive the percentage of a pressure control space was expressed as the ratio of redundant points to the total amount of points on all iso-F0 contours computed in the \( p_b; p_{\text{CAS}} \) control space.

**Muscle stimulation protocol.** Muscles were stimulated with 50 μm diameter Teflon-coated twisted silver wire electrodes (A-M Systems) that were fixed with 10/0 suture. We used a stimulus isolator (model A395, WP1, Sarasota, FL, USA) to apply variable currents (0–10 mA) when limited compliance voltage was sufficient to contract muscles (tinamou). For zebra finches and Bengalese finches, we used two stimulators (model 14E11, DSA, Herlev, Denmark and model 2100, A-M Systems, respectively). To ensure muscle specificity, we placed carbon microfibres (20–40 μm diameter) at regular intervals on the muscles and surrounding tissues and filmed contractions at 1 kHz. We stimulated the single syringeal muscle present in tinamous, \( M. \text{syringalis} \). In songbirds, activity of the \( M. \text{syringalis ventralis} \) correlates to F0 (ref. 67). Therefore we focussed on this muscle to study F0 control in zebra finches (\( N = 5 \) adult males). F0 of the sound was determined 30 ms prior (control) and 30 ms after stimulation. We computed 312-point fast Fourier transforms and used parabolic interpolation around the peak power to determine F0.

**References**

1. King, A. S. Functional anatomy of the syrinx in Form and Function in Birds (eds King, A. S. & McClelland, J.) 4, 105–192 (Academic Press, 1989).
2. Brainard, M. S. & Doupe, A. J. Translating birdsong: songbirds as a model for basic and applied medical research. Annu. Rev. Neurosci. 36, 489–517 (2013).
3. Fee, M. S. & Scharff, C. The songbird as a model for the generation and learning of complex sequential behaviors. ILAR J. 51, 362–377 (2010).
4. Elemans, C. P. H. The single syringeal muscle in the morphology of avian sound production. Curr. Opin. Neurobiol. 28, 172–178 (2014).
5. Goller, F. & Larsen, O. N. A new mechanism of sound generation in songbirds. Proc. Natl Acad. Sci. USA 94, 1477–1479 (1997).
6. Jensen, K. K., Cooper, B. G., Larsen, O. N. & Goller, F. Songbird muscles use pulse tone register in two voices to generate low-frequency sound. Proc. Biol. Sci. 274, 2703–2710 (2007).
7. Elemans, C. P. H., Zaccarelli, R. & Herzel, H. Biomechanics and control of vocalization in a non-songbird. J. R. Soc. Interface 5, 691–703 (2008).
8. Zaccarelli, R., Elemans, C. P. H., Fitch, W. T. & Herzel, H. Modelling bird songs: voice onset, overtones and registers. Acta Acust. United Acust. 92, 741–748 (2006).
9. Mindlin, G. B. & Laje, R. The Physics of Birdsong (Springer, 2005).
10. Amador, A. & Margoliash, D. A mechanism for frequency modulation in songbirds shared with humans. J. Neurosci. 33, 11136–11144 (2013).
11. Amador, A., Perl, Y. S., Mindlin, G. B. & Margoliash, D. Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature 495, 59–64 (2013).
12. van den Berg, J. Myoelastic-aerodynamic theory of voice production. J. Speech Hear. Res. 1, 227–244 (1958).
13. Titze, I. R. Principles of Voice Production (National Center for Voice and Speech, 2000).
14. Titze, I. R. Comments on the myoelastic—aerodynamic theory of phonation. J. Speech Hear. Res. 23, 495–510 (1980).
15. Titze, I. R. & Allipur, F. The Myoelastic Aerodynamic Theory of Phonation (National Center for Voice and Speech, 2006).
16. Riede, T. & Goller, F. Peripheral mechanisms for vocal production in birds—differences and similarities to human speech and singing. Brain Lang. 115, 69–80 (2016).
17. Fitch, T. in Encyclopedia of language and linguistics 2nd ed., Vol. 10 (ed Brown, K) 115–121 (Oxford, Elsevier 2006).
18. Herbott, C. T. et al. How low can you go? Physical production mechanism of elephant infrasonic vocalizations. Science 337, 595–599 (2012).
19. Titze, I. R. & Strong, W. J. Normal modes in vocal cord tissues. J. Acoust. Soc. Am. 57, 736–749 (1975).
20. Titze, I. R. On the relation between subglottal pressure and fundamental frequency in phonation. J. Acoust. Soc. Am. 85, 901–906 (1989).
21. Svec, J. G., Horáček, J., Sram, F. & Veselý, J. Resonance properties of the vocal folds: in vivo laryngoscopic investigation of the externally excited laryngeal vibrations. J. Acoust. Soc. Am. 108, 1397–1407 (2000).
22. Berry, D. A. Mechanisms of modal and nonmodal phonation. J. Phon. 29, 431–450 (2001).
23. Flanagan, J. L. & Landgraf, L. L. Self-oscillating source for vocal-tract synthesizers. IEEE Trans. Audio Electroacoust. AU16, 57 (1968).
24. Titze, I. R. The physics of small-amplitude oscillation of the vocal folds. J. Acoust. Soc. Am. 83, 1029–1032 (1988).
25. Iwatsuka, K. & Flanagan, J. L. Synthesis of voiced sounds from a two-mass model of the vocal cords. Bell Syst. Tech. J. 51, 1233–1268 (1972).
26. Scherer, R. C., Torkaman, S., Kucinschi, B. R. & Aijfe, A. A. Infraglottal pressures in a three-dimensional model with a non-rectangular glottal shape. J. Acoust. Soc. Am. 128, 828–838 (2010).
27. Ficher, L. P. et al. Pressure distributions in a static physical model of the hemilarynx: measurements and computations. J. Voice 24, 2–20 (2010).
28. Miller, D. G. & Schutte, H. K. Characteristic patterns of sub- and supraglottal pressure variations within the glottal cycle. Transcr. XIIIth Symp. Care Prof. Care.
Voice (ed. Lawrence, V.) 70–75 (The Voice Foundation, New York, NY, USA, 1985).

29. Schutte, H. K. & Miller, D. G. Resonanzspiele der Gesangsstimme in ihren Beziehungen zu supra- und subglottalen Druckverlauf. Konsequenzen für die Stimmbildungstheorie. Folia. Phoniatric. (Basel) 40, 65–73 (1988).

30. Herbst, C. T. et al. Complex vibratory patterns in an elephant larynx. J. Exp. Biol. 216, 4054–4064 (2013).

31. Larsen, O. N. & Goller, F. Role of syringeal vibrations in bird vocalizations. Proc. R. Soc. Biol. Sci. 266, 1609–1615 (1999).

32. Fee, M. S., Shraiman, B., Pesaran, B. & Mitra, P. P. The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. Nature 395, 67–71 (1998).

33. Fee, M. S. Measurement of the linear and nonlinear mechanical properties of the oscine syrinx: implications for function. J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol. 188, 829–839 (2002).

34. Dureco, R. Funktionelle Biologie. Anatomie und Physiologie der Vogelstimme PhD thesis, University of Basel (Aku-Fotodruck, Basel, 1974).

35. Paulsen, K. Das Prinzip der Stimmbildung in der Wirbellosen und beim Menschen (Akademische Verlagsgesellschaft, 1967).

36. Amador, A. & Mindlin, G. B. Beyond harmonic sounds in a simple model for birdsong production. Chaos 18, 034123 (2008).

37. Goller, F. & Riede, T. Integrative physiology of fundamental frequency control in birds. J. Physiol. Paris 107, 230–242 (2013).

38. Suthers, R. A. & Zollinger, S. A. in Neuroscience of birdsong (eds Zeigler, H.P. et al. Elsevier, Amsterdam, 2006).

39. Goller, F. & Riede, T. Syrinx in songbirds - integrative physiology. J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol. 216, 485–486 (2012).

40. Kusuyama, T., Fukuda, H., Shiotani, A., Nakagawa, H. & Kanzaki, J. Analysis of vocal fold vibration by X-ray stroboscopy with multiple markers. Otolaryngol. Head Neck Surg. 124, 317–322 (2001).

41. Doreen, M., Chihara, M. & Berry, D. A. Empirical eigenfunctions and median surface dynamics of a human vocal fold. Methods Inf. Med. 44, 384–391 (2005).

42. Boesensnecker, A., Berry, D. A., Lohscheller, J., Eysholdt, U. & Dooling, L. Measurement of vocal fold properties of a human vocal fold. Acta Acust. United Acoust. 93, 815–823 (2007).

43. Kohler, J. B., Chang, E. W., Zeitsls, S. M. & Yun, S.-H. Dynamic imaging of vocal fold oscillation with four-dimensional optical coherence tomography. Laryngoscope 120, 1354–1362 (2010).

44. Dooling, M., Berry, D. A. & Berke, G. S. Maximal vocal variability characterized by X-ray stroboscopy with multiple markers. Otolaryngol. Head Neck Surg. 120, 58–63 (1993).

45. Maina, J. N. The Lung-Air Sac System of Birds (Springer Science & Business Media, 2006).

46. Maina, J. N., Singh, P. & Moss, E. A. Inspiratory aerodynamic valving occurs in the ostrich, Struthio camelus lung: a computational fluid dynamics study under resting unsteady state inhalation. Respir. Physiol. Neurobiol. 169, 262–270 (2010).

47. Mackelprang, R. & Goller, F. Ventilation patterns of the songbird lung/air sac system during different behaviors. J. Exp. Biol. 216, 3611–3619 (2013).

48. Brackenbury, J. H. Lung-air-sac anatomy and respiratory pressures in the bird. J. Exp. Biol. 57, 543–550 (1972).

49. Gaunt, A. S., Stein, R. C. & Gaunt, S. L. Pressure and air flow during distress calls of the starling, Sturnus vulgaris. Respir. Physiol. Neurobiol. 69, 259–270 (1987).

50. Sober, S. J., Wohlgemuth, M. J. & Brainard, M. S. Central contributions to acoustic variation in birdsong. J. Neurosci. 28, 10370–10379 (2008).

51. Scholz, J. P. & Schoner, G. The uncontrolled manifold concept: identifying control variables for a functional task. Exp. Brain Res. 126, 289–306 (1999.).