High-fidelity continuum modeling predicts avian voiced sound production

Authors: W. Jiang[1,§], J.H. Rasmussen[2,§,¶], Q. Xue[1], M. Ding[3,4], X. Zheng[1*], C.P.H. Elemans[2*]

§ Equal contribution

*To whom correspondence should be addressed: Xudong Zheng (xudong.zheng@maine.edu) and Coen Elemans (coen@biology.sdu.dk)

Affiliations:

[1] Department of Mechanical Engineering, University of Maine, Orono, ME, USA.

[2] Department of Biology, University of Southern Denmark, Odense M, Denmark.

[3] Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Denmark

[4] Department of Clinical Research, University of Southern Denmark

[¶] Current address: Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX, USA.

Abstract: Voiced sound production is the primary form of acoustic communication in terrestrial vertebrates, particularly birds and mammals, including humans. Developing a causal physics-based model that links descending vocal motor control to tissue vibration and sound requires embodied approaches that include realistic representations of voice physiology. Here we first implement and then experimentally test a high-fidelity three-dimensional continuum model for voiced sound production in birds. Driven by individual-based physiologically quantifiable inputs, combined with non-invasive inverse methods for tissue material parameterization, our model accurately predicts observed key vibratory and acoustic performance traits. These results demonstrate that realistic models lead to accurate predictions and support the continuum model approach as a critical tool towards a causal model of motor control of voiced sound production.

Main text:

During voiced sound production laryngeal vocal folds (VFs) or analogous syringeal structures exhibit self-sustained vibrations through fluid-tissue interactions and elastic restoring forces. VF vibrations mechanically convert expiratory airflow into pulsatile airflow, which causes air pressure oscillations constituting the acoustic excitation of the system (*1–4*). This physical framework is known as the myoelastic-aerodynamic theory (MEAD) (*2*, *3*) and shared between mammals and birds (*5*). In the human larynx, which has been most intensively studied, VF tissue properties, aerodynamic forces and neuromuscular control of laryngeal and VF geometry all interact to ultimately control VF kinematics and glottal flow that set crucial acoustic source parameters such as fundamental frequency, registers, and spectral slope (*1-4*, *6*). In contrast to humans, birds are not as well-studied but have the significant advantage that the entire neuromechanical motor control loop of their voice production is experimentally accessible (*7*, *8*). Furthermore the process of song motor sequence learning in songbirds can be monitored individually and shares many parallels with human speech acquisition (*9–12*). Thus birds present a tractable model system for *(i)* studying the physics and learned motor control of voiced sound production and *(ii)* to inform on basic principles of MEAD systems control.

Causally linking descending motor control to voiced sound production requires computational biophysical models to explore the high dimensional control space (*13*). High-fidelity continuum models include the full fluid-structure-acoustics interaction (FSAI) complexity of voiced sound production in anatomically realistic geometries of vocal fold and tract (*4*, *14–21*). These models are critically needed when realistic representations of voice physiology and biomechanics are essential, such as in the clinical management of voice disorders (*1, 4*), or understanding motor control of human (*22-24*) or avian (*25*) voiced sound production.

However, the continuum model approach has not yet been applied to voiced sound production in avian model systems. Additionally, the approach in general currently faces three key challenges (*1, 4*). *(i)* The implementation itself of complex FSAI models capable of resolving the large range of spatial and time scales remains challenging (*14-21*). On the experimental side, although partial physiological datasets have been collected in mammals (*1*, *2, 23*, *24*), we *(ii)* currently lack quantification of all critical physiological parameters in single individuals. This would include parameterization of *(a)* anatomical boundary conditions and (*b*) VF tissue properties, and time-resolved quantification of *(c)* flow, pressure and VF posturing changes due to motor control, and *(d)* VF shape within single oscillations. Obtaining these complete individual-based datasets would lift the last critical roadblock and *(ii)* allow experimental validation of FSAI model predictions to strengthen confidence in the continuum model approach.

Here we resolve the above challenges in an avian model system and implemented and experimentally test predictions of a physics-based, high-fidelity three-dimensional (3D) continuum model of voiced sound production. First, we implemented a 3D immersed-boundary (IBM), finite-element method based fluid-structure interaction solver (See methods), where airflow was governed by 3D, unsteady, viscous, incompressible Navier–Stokes equations and dynamics of the avian VF analog were governed by the Navier equation. The fluid and structure solvers were explicitly coupled through a Lagrangian interface where airway and vibrating structures contacted. To allow rapid simulations when resolving the large range of spatial and time scales in simulating voice production, we used the Message-Passing-Interface parallelization method (*26*), combined with IBM as an advanced numerical method designed especially for simulating moving boundaries. By carrying out simulations on Cartesian grids,

IBM circumvents complicated re-meshing algorithms in conventional body-fitted mesh methods and can deal with complex moving boundaries.

To address the second challenge, i.e., quantification of all critical physiological parameters in single individuals, we focused on rock pigeons, where we previously (*5*, *27*) achieved high-quality kinematic data of the avian VF analog, the lateral vibratory masses (LVMs, Fig. 1A). We quantified syringeal anatomy for five individuals using DiceCT scans and implemented high-resolution finite-element meshes of the syrinx (Fig. 1A). To mimic the anatomical constraint of the trachea rings on the motion of the LVMs, zero displacement boundary conditions were applied at mesh nodes where tracheal rings were located. Unilateral LVMs consisted of 1,988-4,620 and 2,159-4,693 nodes and 8,101-18,657 and 9,838-19,131 elements in the solid domain for left and right LVM respectively (Table S1). To parameterize LVM tissue elasticity in each individual, we developed a novel non-invasive, combined experimental/simulation approach. We experimentally induced static LVM displacement by stepwise increase of the transmural pressure (Fig. 1B). We subsequently used the LVM finite-element model combined with genetic-algorithm based optimization (*28*) to simulate LVM displacement as a function of transmural pressure (Fig 1C) and find the elastic modulus (EM) value with the smallest difference between experiment and simulation (Fig 1C). These values were used for further dynamic simulation and ranged from 1.8 to 4.0 kPa for the five individuals (Table S1).

To quantify flow, pressure and LVM posturing changes due to motor control, we applied boundary conditions where LVM vibration was obtained reliably previously (*5*, *27*) to both the experimental preparation and FSAI model (p_b = 1.0 kPa, p_{icas} = 0.5 kPa). Both the *in vitro* syrinx and corresponding individual FSAI models demonstrated self-sustained stable oscillations in all

five cases (Fig 2; Movie S1). To quantifying the time-resolved VF shape within oscillations, we took advantage of both the unique coronal view offered by the pigeon syrinx (*5*) and the lack of a dorsoventral vibrational component (*27*), to quantify the time-resolved syringeal or glottal opening as a function of caudo-cranial position, i.e. a coronal glottovibrogram (cGVG, Fig 2C).

To rigorously test the predictions of the FSAI model during dynamic simulation (the third challenge), we used a blinded procedure where the modeling team (WJ, QX, XZ) had access only to geometry, static loading test and pressure boundary conditions, but was blind to all other data of the experimental procedures (performed by JHR, CPHE). We compared key vibratory and acoustic predictions by the simulations (SIM) to the behavior in the experiment (EXP). The cGVG allowed for comparing four key parameters describing vibratory kinematics: fundamental frequency of the vibration (f_o), the speed of the caudocranial (CC) wave, peak of the minimal glottal opening, and the open-closed quotient (OCQ). In EXP, f_o of the LVM vibration measured 123±25 Hz (median: 114 Hz; range: 95-154 Hz, N=5), CC speed was 0.82±0.34 m/s (median: 0.88 m/s; range: 0.48-1.27 m/s, N=5), peak minimal opening was 1.80±0.24 mm (median: 1.80 mm; range: 1.49-2.11 mm; N=5) and OCQ was 1.3±0.7 (median: 1.1, range: 0.7-2.5; N=5). In SIM, f_o of the LVM vibration measured 143±18 Hz (median: 152 Hz; range: 115-158 Hz, N=5), CC speed was 1.02±0.26 m/s (median: 1.03 m/s; range: 0.68-1.38 m/s, N=5), peak minimal opening was 1.06±0.43 mm (median: 0.90 mm; range: 0.62-1.64 mm; N=5) and OCQ was 1.6±0.3 (median: 1.6, range: 1.1-2.1; N=5). At group level, all predicted values were not significantly different from the experimental values with p=0.18, 0.50, 0.06 and 0.11 respectively (Heteroscedastic two-tailed paired t-test, N=5) (Fig 2D). We furthermore compared the time-resolved LVM shape between EXP and SIM at fixed phases within an oscillatory cycle (Fig 3). During the closed phase (0-120°) the simulated LVM shape matched EXP very well and was

never significantly different from the bootstrapped shape ($p<0.01$, see methods). The only observed significant ($p>0.05$, 2-sample Kolmogorov-Smirnov test) discrepancy occurred during
125 late closed/early opening in three subjects on one side, where the LVM mass tended to move ~0.5 mm more cranial (superior in human anatomical terminology) in EXP compared to SIM (Fig 3BC). Taken together, our FSAI model accurately predicted key kinematic parameters of LVM motion.

Lastly, we compared SIM prediction and EXP observation of two additional key acoustic
130 parameters specifying a sound source – in addition to f_o (Fig 2D) -, namely source level and spectral slope and these did not differ significantly (Fig 4A). Because the simulated LVM vibratory kinematics matched our observations, we exploited our FSAI models to calculate parameters that could not be quantified in the current experiments, such as spatiotemporal pressure and flow velocity distributions over the vibratory cycle (Fig 4B-E, Movie S2). The
135 convergent LVM shape during opening causes high glottal pressures (~0.9 kPa) that transfer 23.4±15.7 μJ of positive energy (N=5) from flow into LVM (Fig 4F), facilitating opening. When maximum opening is reached at 300° phase, the LVMs are straight (0° angle in Fig 4E) reducing glottal pressure. During early closing the inferior LVM edge is moving inwards causing an energy transfer of 7.2±2.2 μJ back into the flow. Consecutively, the LVMs close the glottis by
140 moving together in a divergent shape causing rapid pressure reduction driven by elastic forces (Arrow in Fig 4E). Furthermore, flow inertia in the trachea causes negative pressures (-0.52±0.23 kPa) near the glottis exit prior to full closure (blue region in Fig 4D), facilitating closing. Interestingly, another positive energy transfer to LVMs is observed nearly the end of the cycle. Our avian data thus show that two primary factors contribute to the pressure asymmetry during
145 VF opening and closing that drive self-sustained oscillation: *i)* an alternating

convergent/divergent medial surface profile and *ii)* airflow inertia, corroborating earlier model predictions (*29*) and measurements on (hemi)larynges (*30*) in mammalian voice production.

Our FSAI model accurately predicted key performance traits of tissue motion and acoustics driven solely by physiological parameters (static geometry, tissue elasticity and boundary conditions) and without optimization of either geometry or material properties parameterization on dynamic performance (*20*). While the FSAI model implementation itself is complex (*14–21*), the inputs are simple and have, most importantly, directly measurable physiological, material and geometrical properties. Measurements of VF material properties (*31*, *32*), initial configuration, initial stress and detailed flow-induced 3D VF motion (*33*) have been achieved separately in human and mammalian model systems, but complete physiological data sets have not been obtained in these clades, nor in any birds, in single individuals nor consecutively used to thoroughly test FASI model predictions in a blinded approach. Recent studies encouragingly suggested that realistic continuum 3D models lead to more robust VF dynamics compared to 1D and 2D models (*34*). Our data shows that they also lead to accurate predictions performance. Our results therefore strongly support the continuum model approach as a critical step in integrating *in vitro, ex vivo and in vivo* experimental data with brute force computational approaches towards a causal model of motor control of voiced sound production applicable to birds and mammals.

Continuum models currently allow only limited systematic parameters exploration because they are computationally expensive (*21*). In contrast, phenomenological reduced order (RO) models - that simplify VFs to one or two masses (*35*, *36*)- allow broad exploration of a simplified control space with less computational power and as such have been used widely to capture the dynamics of mammalian (*37*) and avian voiced sound production (*13*, *38–40*).

However, because RO models are not first-principle physics models they essentially lack physiological representation of tissue properties and geometry (*1*, *4*, *37*). Reducing computational costs and time is thus crucial for making continuum models useful in experimental or clinical settings (*1*). Rapid developments in machine learning tools combined with supercomputing provide promising avenues to predict complex flow solvers (*41*, *42*) that could significantly reduce computational power and allow elaborate parameter exploration.

All vocal behaviors result from system-wide interactions between the nervous system, body, and surrounding environment (*43, 44*). Understanding the activity of neural circuitry controlling voice can therefore only be understood by embodied models that include realistic biomechanics of sound generation and musculoskeletal actuation of the vocal organ (*8*, *45, 46*). During postnatal vocal development, gradual body changes, such as lung growth or VF stiffness in marmosets (*46, 47*) or increased muscle speed in songbirds (*48*), can drive changes in vocal behavior. Embodied human VF models have direct clinical relevance to pathological physical behaviors with abnormal vocal output (*1*) and patient specific model-assisted phonosurgery (*49*, *50*), because most laryngeal human voice disorders are caused by changes in VF geometry, structural integrity, or kinematics (*4*). The embodied approach to voice production as presented here thus improves our understanding how neural mechanisms and biomechanics interact to drive vocal behavior in vertebrates.

References and notes:

1. Z. Zhang, Mechanics of human voice production and control. *J. Acoust. Soc. Am.* **140**, 2614–2635 (2016).

2. I. R. Titze, *Principles of Voice Production* (National Center for Voice and Speech, Denver, 2000).

3. J. van den Berg, Myoelastic-aerodynamic theory of voice production. *J. Speech Hear. Res.* **1**, 227–244 (1958).

4. R. Mittal, B. D. Erath, M. W. Plesniak, Fluid Dynamics of Human Phonation and Speech. *Annu. Rev. Fluid Mech.* **45**, 437–467 (2013).

5. C. P. H. Elemans, J. H. Rasmussen, C. T. Herbst, D. N. Düring, S. A. Zollinger, H. Brumm, K. Srivastava, N. Svane, M. Ding, O. N. Larsen, S. J. Sober, J. G. Švec, Universal mechanisms of sound production and control in birds and mammals. *Nat. Commun.* **6** (2015), doi:10.1038/ncomms9978.

6. E. B. Holmberg, R. E. Hillman, J. S. Perkell, P. C. Guiod, S. L. Goldman, Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *J. Speech Hear. Res.* **38**, 1212–1223 (1995).

7. T. Riede, F. Goller, Peripheral mechanisms for vocal production in birds - differences and similarities to human speech and singing. *Brain Lang.* **115**, 69–80 (2010).

8. C. P. H. Elemans, The singer and the song: The neuromechanics of avian sound production. *Curr. Opin. Neurobiol.* **28**, 172–178 (2014).

9. M. S. Brainard, A. J. Doupe, Translating Birdsong: Songbirds as a Model for Basic and Applied Medical Research. *Annu. Rev. Neurosci.* **36**, 489–517 (2013).

10. O. Tchernichovski, P. P. Mitra, T. Lints, F. Nottebohm, Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science.* **291**, 2564–2569 (2001).

11. M. A. Long, M. S. Fee, Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*. **456**, 189–194 (2008).

12. T. S. Okubo, E. L. Mackevicius, H. L. Payne, G. F. Lynch, M. S. Fee, Growth and splitting of neural sequences in songbird vocal development. *Nature*. **528**, 352–357 (2015).

13. A. Amador, Y. S. Perl, G. B. Mindlin, D. Margoliash, Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature*. **495**, 59–64 (2013).

14. Q. Xue, X. Zheng, R. Mittal, S. Bielamowicz, Subject-specific computational modeling of human phonation. *J. Acoust. Soc. Am.* **135**, 1445–1456 (2014).

15. W. Jiang, X. Zheng, Q. Xue, Computational modeling of fluid-structure-acoustics interaction during voice production. *Front. Bioeng. Biotechnol.* **5** (2017), doi:10.3389/fbioe.2017.00007.

16. X. Zheng, Q. Xue, R. Mittal, S. Beilamowicz, A coupled sharp-interface immersed boundary-finite-element method for flow-structure interaction with application to human phonation. *J. Biomech. Eng.* **132** (2010), doi:10.1115/1.4002587.

17. R. Mittal, H. Dong, M. Bozkurttas, F. M. Najjar, A. Vargas, A. von Loebbecke, A versatile sharp interface immersed boundary method for incompressible flows with complex boundaries. *J. Comput. Phys.* **227**, 4825–4852 (2008).

18. F. Alipour, D. A. Berry, I. R. Titze, A finite-element model of vocal-fold vibration. *J. Acoust. Soc. Am.* **108**, 3003–3012 (2000).

19. D.J. Daily, S. L. Thomson, Acoustically-coupled flow-induced vibration of a computational vocal fold model. *Computers & structures* 116 (2013): 50-58 (2013)

20. S. Chang, C. K. Novaleski, T. Kojima, M. Mizuta, H. Luo, B. Rousseau, Subject-Specific Computational Modeling of Evoked Rabbit Phonation. *J. Biomech. Eng.* **138** (2016), doi:10.1115/1.4032057.

21. Z. Zhang, Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *J. Acoust. Soc. Am.* **139**, 1493–1507 (2016).

22. E. J. Hunter, I. R. Titze, F. Alipour, A three-dimensional model of vocal fold abduction/adduction. *J. Acoust. Soc. Am.* **115**, 1747–1759 (2004).

23. D. K. Chhetri, J. Neubauer, J. L. Bergeron, E. Sofer, K. A. Peng, N. Jamal, Effects of asymmetric superior laryngeal nerve stimulation on glottic posture, acoustics, vibration. *Laryngoscope*. **123**, 3110–3116 (2013).

24. D. K. Chhetri, J. Neubauer, D. A. Berry, Neuromuscular control of fundamental frequency and glottal posture at phonation onset. *J. Acoust. Soc. Am.* **131**, 1401–1412 (2012).

25. E. P. Kingsley, C. M. Eliason, T. Riede, Z. Li, T. W. Hiscock, M. Farnsworth, S. L. Thomson, F. Goller, C. J. Tabin, J. A. Clarke, Identity and novelty in the avian syrinx. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 10209–10217 (2018).

26. W. Gropp, E. Lusk, N. Doss, A. Skjellum, A high-performance, portable implementation of the MPI message passing interface standard. *Parallel Comput.* **22**, 789–828 (1996).

27. J. H. Rasmussen, C. T. Herbst, C. P. H. Elemans, Quantifying syringeal dynamics in vitro using electroglottography. *J. Exp. Biol.* **221** (2018), doi:10.1242/jeb.172247.

28. G. Liu, B. Geng, X. Zheng, Q. Xue, H. Dong, G. V. Lauder, An image-guided computational approach to inversely determine in vivo material properties and model flow-structure interactions of fish fins. *J. Comput. Phys.* **392**, 578–593 (2019).

29. I. R. Titze, The physics of small-amplitude oscillation of the vocal folds. *J. Acoust. Soc. Am.* **83**, 1536–1552 (1988).

30. F. Alipour, S. Jaiswal, Phonatory characteristics of excised pig, sheep, and cow larynges. *J. Acoust. Soc. Am.* **123**, 4572–4581 (2008).

31. G. R. Dion, S. Jeswani, S. Roof, M. Fritz, P. G. Coelho, M. Sobieraj, M. R. Amin, R. C. Branski, Functional assessment of the ex vivo vocal folds through biomechanical testing: A review. *Mater. Sci. Eng. C*. **64**, 444–453 (2016).

32. A. K. Miri, Mechanical characterization of vocal fold tissue: A review study. *J. Voice*. **28**, 657–667 (2014).

33. A. M. Vahabzadeh-Hagh, Z. Zhang, D. K. Chhetri, Three-dimensional posture changes of the vocal fold from paired intrinsic laryngeal muscles. *Laryngoscope*. **127**, 656–664 (2017).

34. X. Zheng, R. Mittal, Q. Xue, S. Bielamowicz, Direct-numerical simulation of the glottal jet and vocal-fold dynamics in a three-dimensional laryngeal model. *J. Acoust. Soc. Am.* **130**, 404–415 (2011).

35. K. Ishizaka, J. L. Flanagan, Synthesis of Voiced Sounds From a Two-Mass Model of the Vocal Cords. *Bell Syst. Tech. J.* **51**, 1233–1268 (1972).

36. B. H. Story, I. R. Titze, Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Am.* **97**, 1249–1260 (1995).

37. B. D. Erath, M. Zañartu, K. C. Stewart, M. W. Plesniak, D. E. Sommer, S. D. Peterson, A review of lumped-element models of voiced speech. *Speech Commun.* **55**, 667–690 (2013).

38. G. B. Mindlin, The physics of birdsong production. *Contemp. Phys.* **54**, 91–96 (2013).

39. R. Zaccarelli, C. P. H. Elemans, W. T. Fitch, H. Herzel, Modelling bird songs: Voice onset, overtones and registers. *Acta Acust. united with Acust.* **92**, 741–748 (2006).

40. C. P. H. Elemans, R. Zaccarelli, H. Herzel, Biomechanics and control of vocalization in a non-songbird. *J. R. Soc. Interface*. **5**, 691–703 (2008).

41. B. Kim, V. C. Azevedo, N. Thuerey, T. Kim, M. Gross, B. Solenthaler, Deep Fluids: A Generative Network for Parameterized Fluid Simulations. *Comput. Graph. Forum*. **38**, 59–70 (2019).

42. B. N. Hanna, N. T. Dinh, R. W. Youngblood, I. A. Bolotnov, Machine-learning based error prediction approach for coarse-grid Computational Fluid Dynamics (CG-CFD). *Prog. Nucl. Energy*. **118**, 103140 (2020).

43. E. D. Tytell, P. Holmes, A. H. Cohen, Spikes alone do not behavior make: Why neuroscience needs biomechanics. *Curr. Opin. Neurobiol.* **21**, 816–822 (2011).

44. K. Nishikawa, A. A. Biewener, P. Aerts, A. N. Ahn, H. J. Chiel, M. A. Daley, T. L. Daniel, R. J. Full, M. E. Hale, T. L. Hedrick, A. K. Lappin, T. R. Nichols, R. D. Quinn, R. A. Satterlie, B. Szymik, Neuromechanics: An integrative approach for understanding motor control. *Integr. Comp. Biol.* **47**, 16–54 (2007).

45. D. N. Düring, C. P. H. Elemans, in *Vertebrate Sound Production and Acoustic Communication* (2016), pp. 119–157.

46. Y. S. Zhang, A. A. Ghazanfar, Vocal development through morphological computation. *PLoS Biol.* **16** (2018), doi:10.1371/journal.pbio.2003933.

47. Y.S. Zhang, D.Y. Takahashi, D.A. Liao, A.A. Ghazanfar, C.P.H. Elemans, Vocal state change through laryngeal development. *Nat. Commun.* **10***,* 4592*.* (2019)

48. I. Adam, C. P. H. Elemans, Vocal Motor Performance in Birdsong Requires Brain-Body Interaction. *eNeuro*. **6** (2019), doi:10.1523/ENEURO.0053-19.2019.

49. R. Mittal, X. Zheng, R. Bhardwaj, J. H. Seo, Q. Xue, S. Bielamowicz, Toward a simulation-based tool for the treatment of vocal fold paralysis. *Front. Physiol.* (2011), doi:10.3389/fphys.2011.00019.

50. H. Sadeghi, S. Kniesburges, S. Falk, M. Kaltenbacher, A. Schützenberger, M. Döllinger, Towards a clinically applicable computational larynx model. *Appl. Sci.* **9** (2019), doi:10.3390/app9112288.

51. B. D. Metscher, Micro CT for comparative morphology: Simple staining methods allow high-contrast 3D imaging of diverse non-mineralized animal tissues. *BMC Physiol.* **9** (2009), doi:10.1186/1472-6793-9-11.

52. D. N. Düring, A. Ziegler, C. K. Thompson, A. Ziegler, C. Faber, J. Müller, C. Scharff, C. P. H. Elemans, The songbird syrinx morphome: A three-dimensional, high-resolution, interactive morphological map of the zebra finch vocal organ. *BMC Biol.* **11** (2013), doi:10.1186/1741-7007-11-1.

53. J. Seo, R. Mittal, A high-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries. *J. Comput. Phys.* **230**, 1000–1019 (2011).

54. J. H. Seo, Y. J. Moon, Linearized perturbed compressible equations for low Mach number aeroacoustics. *J. Comput. Phys.* **218**, 702–719 (2006).

55. C. P. H. Elemans, *How do birds sing? Sound analysis-mechanical modelling-muscular control* (Wageningen University, 2004).

Acknowledgements: The authors with to thank T Christensen, S Jakobsen, P Martensen and F Mortensen for technical support, I Adam for discussion, and C Herbst and J Rattcliffe for comments on the manuscript. **Funding**: This research was supported by the Danish Research Council and NovoNordisk Foundation (CPHE). **Author contributions**: WJ, JHR, QX, XZ, CPHE designed research; WJ, JHR, MD, QX, XZ, CPHE performed research, with the modeling team consisting of WJ, QX and XZ, and experimental team of JHR and CPHE; CT scans were performed by JHR and MD; QX, XZ, CE contributed new reagents or analytic tools. WJ, JHR, QX, XZ, CPHE analyzed data; CPHE wrote the first draft; All authors read and approved the final manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All data and code used for analysis are available upon request.

FIGURES

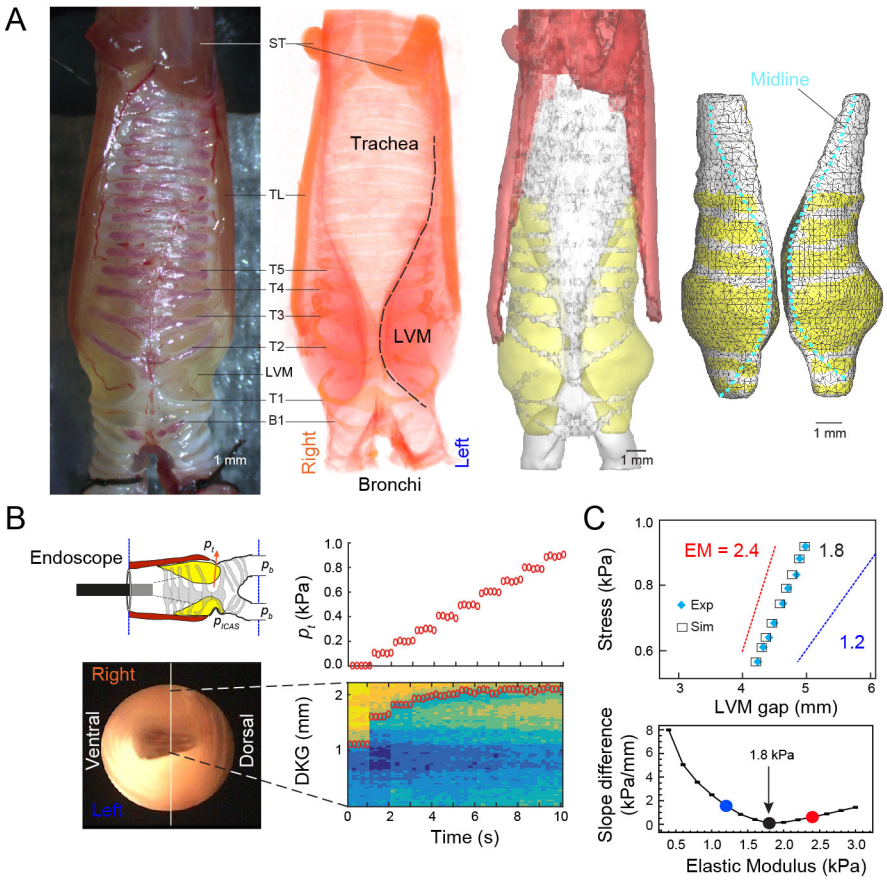

Figure 1. Parameterization of vocal organ geometry and tissue properties. (A) Workflow to parameterize FSAI model geometry (Subject P1), with from left to right: 1) photo of syrinx mounted in *in vitro* experimental setup, 2) voltex rendering of iodine contrasted µCT scan showing the bilateral Lateral Vibratory Masses (LVM), 3) 3D anatomy, and 4) Finite Element Mesh of LVM solid domain. Yellow and white mesh elements on LVM outer surface have free and zero displacement boundary conditions, respectively. (B) Tissue properties were determined by a combined experimental and modeling approach. A stepwise increase of (top) transmural pressure p_t caused a stepwise sideways displacement of the LVMs (bottom) as viewed by digital kymogram (DKG) along the white vertical line in the endoscopic image (left). **(C)** We used a reverse engineering approach to determine the LVM elastic modulus (EM) for each individual. Top: stress-gap width curve of experimental (blue diamonds) and modeled displacement for LVMs with EM=1.8 kPa (black squares). Also indicated are simulated aligned gap widths of EM=1.2 (blue line) and 2.4 kPa (red line). Bottom: Minimal difference between experimental and simulated data indicates occurs at EM = 1.8 kPa for this individual. B1, bronchial ring; T1,…, T5, tracheal rings; LVM, Lateral Vibratory Membrane; p_t, transmural pressure TL, tracheolateralis muscle; ST, sternotrachealis muscle.

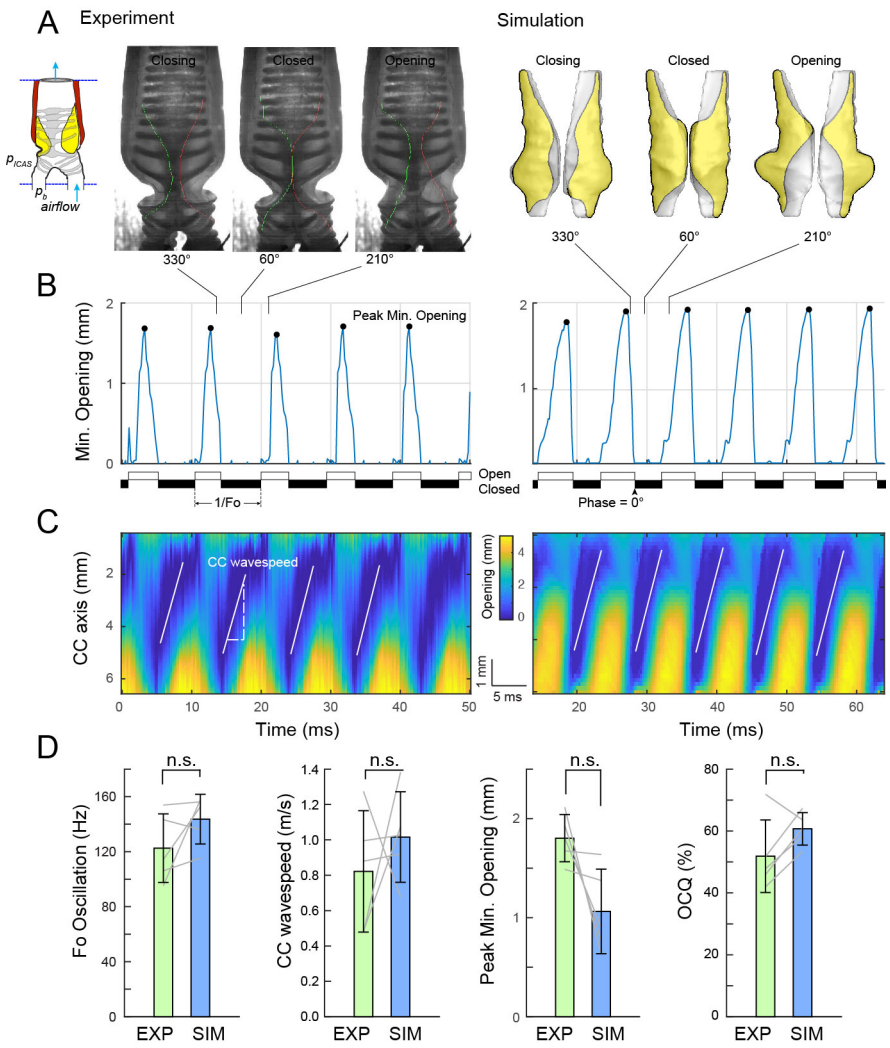

Figure 2. FSAI simulation accurately predicts key features of LVM kinematics. (A) Trans-illuminated syrinx in experiment (EXP, left) and simulation (SIM, right) showing the LVM shape in coronal plane at 330, 60 and 210° of the vibratory cycle, where 0° is defined as the first frame of full closure. (B) The minimal glottal opening shows when the glottis in open and closed (white and black horizontal bars) and vibratory frequency *f*o. The open-closed quotient (OCQ) is the ratio between duration of open and closed glottis. (C) Coronal glottovibrogram (cGVG) showing time-resolved glottal opening along caudocranial axis of 5 oscillations. The white lines indicate the CC wavespeed regression slopes based on the closed glottis. (D) Observed EXP data and SIM predictions are not significantly different for the above indicated four key kinematic parameters.

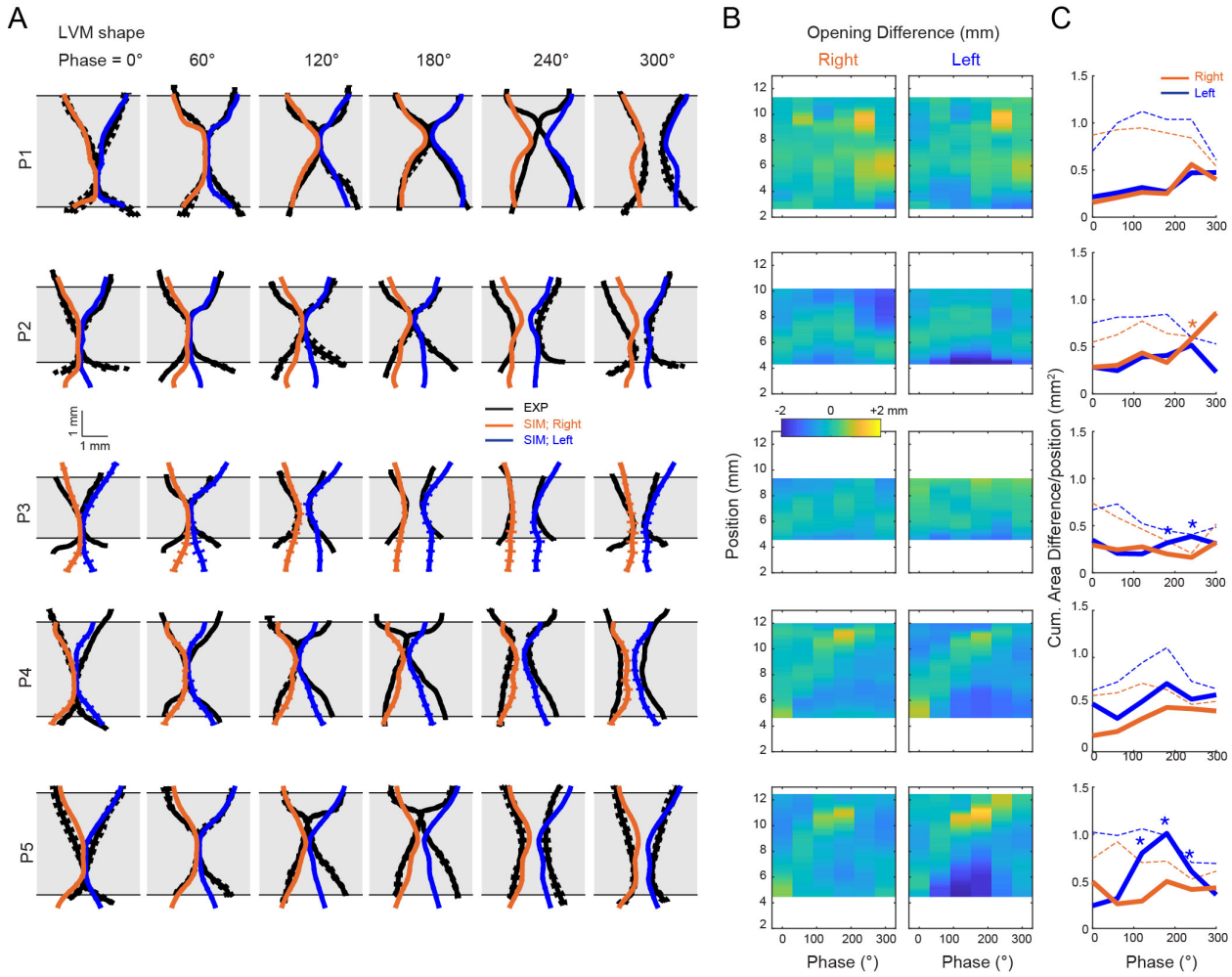

Figure 3. FSAI simulation accurately predicts within-cycle LVM shape. (A) Bilateral LVM shape of EXP and SIM in coronal plane at 0, 60,120, 180, 240, 300 degrees averaged over 5 cycles (mean (solid lines) ± S.D (dotted lines). (B) Lateral position difference between EXP and SIM left and right LVMs and (C) unsigned cumulative area difference between EXP and SIM shows that the predicted shapes are very similar to the observed shapes. The most pronounced discrepancies occur during late closed/early opening. Dotted lines are mean. Asterisks indicate significant difference (p>0.01) between SIM data and 1000x bootstrapped EXP shapes for 2-sample Kolmogorov-Smirnov tests.

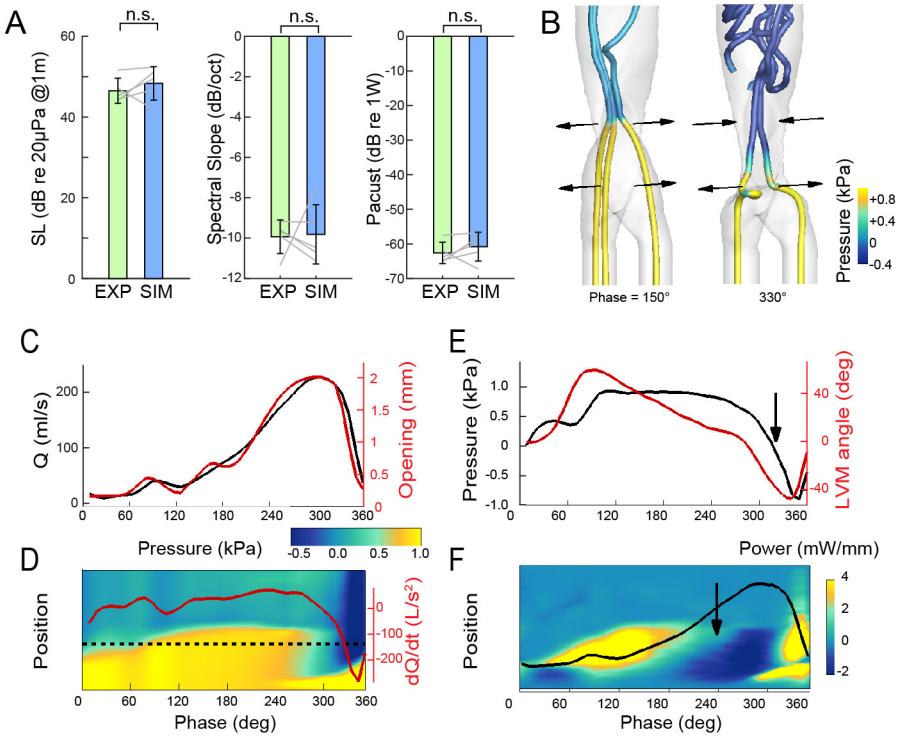

Figure 4. FSAI predictions of acoustics and spatiotemporal pressure and power profiles. (A) The SIM accurately predicted key acoustic parameters; source level (SL), spectral slope and acoustic power. (B) 3D flow inside the syrinx as indicated by flow streamlines (pressure color-contoured). Horizontal arrows indicate the motion direction of the LVMs. (C) Flow rate Q (black line) correlates strongly with glottal opening (red). (D) Spatiotemporal pressure distribution along the airway centerline over a cycle with superimposed dQ/dt (red solid line) as a proxy for sound pressure. (E) Glottal pressure (black line) evaluated at the horizontal dotted line in (D) with LVM angle (red line). (F) Spatiotemporal power transfer distribution from flow to LVM evaluated along the airway centerline with superimposed flow Q (black solid line).