Triangular body-cover model of the vocal folds with coordinated activation of the five intrinsic laryngeal muscles

Gabriel A. Alzamendi,[1,a)] Sean D. Peterson,[2,b)] Byron D. Erath,[3,c)] Robert E. Hillman,[4,d)] and Matías Zañartu[5,e)]

[1]*Institute for Research and Development on Bioengineering and Bioinformatics (IBB), CONICET-UNER, Oro Verde, Entre Ríos 3100, Argentina*

[2]*Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada*

[3]*Department of Mechanical and Aerospace Engineering, Clarkson University, Potsdam, New York 13699, USA*

[4]*Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, Massachusetts 02114, USA*

[5]*Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso 2390123, Chile*

ABSTRACT:

Poor laryngeal muscle coordination that results in abnormal glottal posturing is believed to be a primary etiologic factor in common voice disorders such as non-phonotraumatic vocal hyperfunction. Abnormal activity of antagonistic laryngeal muscles is hypothesized to play a key role in the alteration of normal vocal fold biomechanics that results in the dysphonia associated with such disorders. Current low-order models of the vocal folds are unsatisfactory to test this hypothesis since they do not capture the co-contraction of antagonist laryngeal muscle pairs. To address this limitation, a self-sustained triangular body-cover model with full intrinsic muscle control is introduced. The proposed scheme shows good agreement with prior studies using finite element models, excised larynges, and clinical studies in sustained and time-varying vocal gestures. Simulations of vocal fold posturing obtained with distinct antagonistic muscle activation yield clear differences in kinematic, aerodynamic, and acoustic measures. The proposed tool is deemed sufficiently accurate and flexible for future comprehensive investigations of non-phonotraumatic vocal hyperfunction and other laryngeal motor control disorders.

© 2022 Acoustical Society of America. https://doi.org/10.1121/10.0009169

(Received 14 May 2021; revised 24 November 2021; accepted 6 December 2021; published online 3 January 2022)

[Editor: Zhaoyan Zhang]   Pages: 17–30

I. INTRODUCTION

Non-phonotraumatic vocal hyperfunction (NPVH) is a common voice disorder associated with excessive and poorly regulated activity of the intrinsic and extrinsic laryngeal muscles (Hillman *et al.*, 2020), causing a range of different types of disordered voice quality (Van Stan *et al.*, 2021) but without trauma to the vocal fold (VF) tissue. A common manifestation of NPVH during phonation is high levels of stiffness and tension in the VFs accompanied by incomplete glottal closure causing the voice to be rough, strained, and breathy with increased subglottal pressure and slightly higher, less periodic, and less variable pitch (Espinoza *et al.*, 2020; Espinoza *et al.*, 2017; Hillman *et al.*, 2020; Van Stan *et al.*, 2021). Very little is known about the specific physical mechanisms that underlie this or other manifestations of NPVH. However, the apparent increase in muscle activity and vocal fold stiffness, abducted glottal

configuration, and relatively restricted pitch variability suggest that poor coordination of antagonist laryngeal muscle pairs plays a key role in the altered biomechanics and resulting dysphonia associated with NPVH.

Pitch, loudness, and quality of the voice are primarily controlled by laryngeal muscles in a complex process that is still not fully understood. The process involves the coordinated action of intrinsic and extrinsic muscles (Huber *et al.*, 2004), neural muscle effects (Titze *et al.*, 2002), and auditory and somatosensory feedback and feedforward mechanisms (Lester *et al.*, 2020). Though no mathematical model currently captures all of these factors, efforts have been made to describe components of laryngeal muscle control by either simplified biomechanical representations (Farley, 1996; Titze and Hunter, 2007; Titze and Story, 2002) or high-fidelity three-dimensional finite element models (Geng *et al.*, 2020; Gömmel *et al.*, 2007; Hunter *et al.*, 2004; Movahhedi *et al.*, 2021; Yin and Zhang, 2014). These high order models can capture the complex biomechanical and geometrical changes due to muscle activation but are likely too computationally demanding to account for neural motor control effects. Low-order models can better handle the latter task (Manriquez *et al.*, 2019) and are more suitable for comprehensive parametric simulations that would be needed in the context of laryngeal motor control. However, there is

a)Also at: Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Oro Verde, Entre Ríos 3100, Argentina. Electronic mail: galzamendi@ingenieria.uner.edu.ar

b)ORCID: 0000-0001-8746-2491.

c)ORCID: 0000-0003-0057-6731.

d)Also at: Department of Surgery, Massachusetts General Hospital–Harvard Medical School, Boston, MA 02115, USA.

e)ORCID: 0000-0001-5581-4392.

currently no muscle activation framework for low-order lumped-element models of the vocal folds that incorporates the possibility of independent co-contraction of all agonist/antagonist intrinsic muscle pairs.

The present study introduces a physiologically-based scheme for controlling the mechanical properties of a self-sustained, low-order model of the vocal folds through both independent and coordinated activation of all five intrinsic laryngeal muscles. The approach builds upon prior efforts that describe rules for controlling low-order models (Titze and Story, 2002), vocal fold posturing (Titze and Hunter, 2007), and a triangular body-cover vocal fold model (Galindo et al., 2017). The scheme provides a flexible and physiologically relevant way to control the self-sustained, fully interactive voice production model for both sustained vowels and time-varying glottal gestures that enables the exploration of the role of antagonistic muscle pairs in phonation. In this study, we illustrate the capabilities of the proposed scheme and contrast the simulations against prior numerical and experimental studies. Future studies will be devoted to comprehensively exploring the range of vocal disturbances associated with NPVH applying the proposed scheme.

The paper is organized as follows. In Sec. II, the different components of the proposed model are introduced. In Sec. III, results are presented for posturing, steady and dynamic phonatory gestures, and antagonistic muscle behavior. In Sec. IV, the discussion of the results is provided, along with the proposes guidelines for future work. Finally, Sec. V summarizes the work and provides the main conclusions of the study.

II. PHYSIOLOGICALLY-CONTROLLED VOICE PRODUCTION MODEL

Herein, we aim to simulate the muscle control exerted by the coactivation of the five intrinsic laryngeal muscles during phonation: cricothyroid (CT), thyroarytenoid (TA), lateral cricoarytenoid (LCA), interarytenoid (IA), and posterior cricoarytenoid (PCA); along with the passive contribution of the vocal ligament (LIG) and VF mucosa (MUC). Laryngeal function is described herein in terms of the glottal

posture, resulting from the accommodation of the laryngeal cartilages, and of the VF oscillations around the phonatory configuration.

The main components of the proposed model are comprised of a VF posturing scheme, the triangular body-cover VF model, an extended set of physiological rules for the VF model, and a vocal tract model. Each component is introduced in this section and their interrelationships are described.

A. Vocal fold posturing model

Phonatory posturing refers to the large, slowly-varying changes (relative to VF oscillations) in the glottal geometry due to the accommodation of the surrounding structures and the mechanics of the laryngeal tissues, which can be described in terms of the relative movement of the main laryngeal cartilages (see Fig. 1). The cricoid cartilage is a ring-shaped structure that delimits the larynx inferiorly. It provides support anteriorly to the thyroid cartilage, a large shield-like structure forming the main anterior framework of the larynx, and posteriorly to the paired small pyramidal-shaped arytenoid cartilages. The anterior angle in each arytenoid, referred to as the vocal process (VP), serves as the posterior points of attachment for the VFs. Adduction/abduction comes from the movement of the arytenoid cartilages around the cricoarytenoid joint (CAJ), the anatomical link with the cricoid cartilage [see Fig. 1(a)]. Whereas adduction is controlled by the LCA, IA, and TA muscles, abduction is mainly controlled by the PCA muscle with a minor contribution from the CT muscle (Chhetri et al., 2012, 2014; Geng et al., 2020; Hunter et al., 2004). On the other hand, the relative movement of the thyroid and cricoid cartilages around the cricothyroid joint (CTJ), as shown in Fig. 1(b), is driven by the co-activation of the CT and TA muscles and plays a key role in VF elongation (Chhetri et al., 2014; Geng et al., 2020).

The biomechanical model introduced by Titze and Hunter (2007) was applied to capture the main effects of the laryngeal musculature on the laryngeal configuration. A detailed description of the model development and the measurement of the biomechanical parameters can be found in
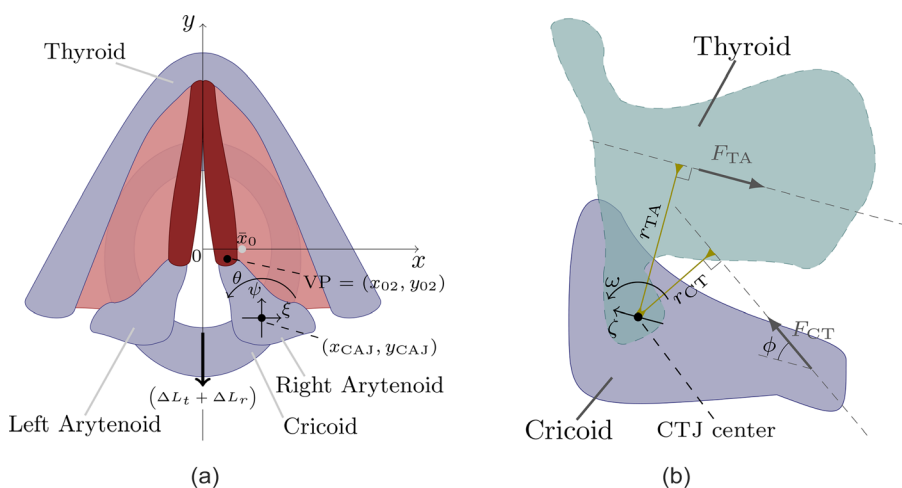

FIG. 1. (Color online) Main laryngeal structures involved in the prephonatory posturing. (a) Cricoarytenoid accommodation at the glottal plane; and (b) Projection of effective cricothyroid accommodation on the glottal plane. Glottal geometry and vocal fold adjustment are controlled via the relative accommodation of major laryngeal cartilages. Figures adapted from Titze (2006). CAJ: cricoarytenoid joint, CTJ: cricothyroid joint.

Chapter 3 of Titze (2006). The model is built around the relative configuration of laryngeal cartilages; namely, that of the arytenoid cartilages with respect to the cricoid cartilage at the CAJ, and of the thyroid cartilage with respect to the cricoid cartilage at the CTJ. Phonatory configuration is described in a transverse plane at the glottal level [see Fig. 1(a)]. The global coordinate system is set with the y axis along the midplane in the glottis, pointing anteriorly, and the x axis passing through the vocal process in the cadaveric position, pointing to the right. Hence, the Cartesian coordinates of the right VP in the cadaveric state is $(\bar{x}_0, 0)$, and geometric symmetry across the midplane is henceforth assumed.

The complex movements involved in arytenoid accommodation are modeled as coordinated "effective" movements of translation and rotation relative to the fixed cadaveric CAJ center, (x_{CAJ}, y_{CAJ}). Let (ξ, ψ) and θ be, respectively, the relative Cartesian displacements and the angle of rotation of the right arytenoid cartilage with respect to the CAJ center, as described in Fig. 1(a). The equations of motion are as follows (Titze, 2006; Titze and Hunter, 2007):

$$M_{ac}\ddot{\xi} + d_x\dot{\xi} + k_x\xi = \sum_{i\in\mathcal{I}} \alpha_i F_i, \qquad (1a)$$

$$M_{ac}\ddot{\psi} + d_y\dot{\psi} + k_y\psi = \sum_{i\in\mathcal{I}} \beta_i F_i, \qquad (1b)$$

$$I_{ac}\ddot{\theta} + \delta\dot{\theta} + \kappa\theta = \sum_{i\in\mathcal{I}} \gamma_i F_i, \qquad (1c)$$

where a dot over a variable indicates differentiation with respect to time, t. F_i denotes the point force magnitude for each element in the set of laryngeal tissues $\mathcal{I} = \{$LCA, IA, PCA, CT, TA, LIG, MUC$\}$, and α_i, β_i, and γ_i are the associated directional cosines and directional moment, as reported by Titze and Hunter (2007). Biomechanical parameters M_{ac} and I_{ac} are the mass and moment of inertia of the arytenoid cartilage; k_x, k_y, and κ are the translational and rotational stiffnesses (computed according to Chapter 3 in Titze, 2006); and d_x, d_y, and δ are the translational and rotational damping coefficients, respectively.

Following Titze (2006), the relative cricoid accommodation with respect to the thyroid cartilage is in turn simulated in terms of translational, ζ, and rotational, ω, movements around the fixed CTJ center, as illustrated in Fig. 1(b). A simple scheme for projecting the effects of CT muscle activation on the horizontal glottal plane is considered, which allows assessment of the concomitant changes on the glottal configuration and vocal fold elongation. It is based on the following second-order equations:

$$M_{cc}\ddot{\zeta} + k_t(t_t\dot{\zeta} + \zeta) = [\cos\phi\, F_{CT} - (F_{TA} + F_{LIG} + F_{MUC})],$$
$$(2a)$$

$$I_{cc}\ddot{\omega} + k_r(t_r\dot{\omega} + \omega) = [r_{CT}F_{CT} - r_{TA}(F_{TA} + F_{LIG} + F_{MUC})],$$
$$(2b)$$

where M_{cc} and I_{cc} are the mass and moment of inertia associated with translation and rotation of the cricoid-arytenoid complex around the CTJ, k_t and k_r are the translational and rotational stiffnesses (computed according to Chapter 3 in Titze, 2006), t_t and t_r are the time constants for the translational and rotational viscous damping, r_{TA} and r_{CT} are the moment arms for the TA and CT torques relative to CTJ center, and ϕ is the angle between the line of action of F_{CT} and the translation axis of the cricoid cartilage, respectively.

Cricoid accommodation gives rise to a concomitant posterior displacement, $(\Delta L_t + \Delta L_r)$, in the arytenoid cartilage, as illustrated in Fig. 1(a), where $\Delta L_t = \zeta$ and $\Delta L_r = r_{TA}\omega$ are the resulting posture effects due to the translation and rotation around the CTJ, respectively. As the VP is structurally linked to the arytenoid cartilage, its Cartesian coordinates are obtained as follows (Titze and Hunter, 2007):

$$x_{02} = x_{CAJ} - (x_{CAJ} - \bar{x}_0)\cos\theta + y_{CAJ}\sin\theta + \xi, \qquad (3a)$$

$$y_{02} = y_{CAJ}(1 - \cos\theta) - (x_{CAJ} - \bar{x}_0)\sin\theta + \psi$$
$$- (\Delta L_t + \Delta L_r). \qquad (3b)$$

The contribution to vocal fold elongation due to arytenoid adduction is obtained by (Titze and Hunter, 2007)

$$\Delta L_a = -[y_{CAJ}(1 - \cos\theta) - (x_{CAJ} - \bar{x}_{02})\sin\theta + \psi].$$
$$(4)$$

Then, the total VF elongation is modeled as the sum of the translational, rotational, and adductory components, and thus the vocal fold strain is given by

$$\epsilon = \frac{1}{L_0}(\Delta L_a + \Delta L_t + \Delta L_r), \qquad (5)$$

where L_0 is the cadaveric VF length. Equation (5) dynamically couples the adduction and elongation procedures (Titze, 2006; Titze and Hunter, 2007). The elongated VF length is $L_g = (1 + \epsilon)L_0$.

In order to simulate the internal stress-strain response in the laryngeal tissues and to obtain the forces F_i, $i \in \mathcal{I}$, the modified one-dimensional Kelvin model of elongation is applied (Hunter and Titze, 2007; Hunter et al., 2004). The modeled tissue force accounts for the active component due to the internal muscle contractile forces, and the passive response force in connective tissue. The activation level for each intrinsic muscle is then controlled through a normalized coefficient, i.e., $0.0 \le a_i \le 1.0$, for $i \in \{$LCA, IA, PCA, CT, TA$\}$. Because the vocal ligament and mucosa are non-contractile connective tissues, the active component is set to zero for $i \in\{$LIG, MUC$\}$. In the Appendix, the muscle model and the biomechanical parameters for the laryngeal tissues are described in more detail.

B. Triangular body-cover model of the vocal folds

The triangular body-cover model (TBCM) introduced by Galindo et al. (2017) is revised for numerically modeling

the VF oscillations. Figure 2 provides a schematic of the TBCM model, with the associated tissue parameters and coordinates, where the subscripts u, l, and b henceforth denote the upper, lower, and body masses, respectively. The VF model is based on the layered approximation proposed by Story and Titze (1995) for capturing the body-cover structure of the VFs, where two masses stacked in the inferior-superior direction model the cover and a single large mass situated laterally encapsulates the body dynamics. These lumped elements are coupled *via* non-linear mechanical elements to account for tissue viscoelasticity. In addition, the TBCM incorporates a gradual tilt of the VF edges in the anterior-posterior direction as a function of the degree of abduction, as is shown in Fig. 2, giving rise to a triangular glottis accounting for the anatomical dorsal-ventral gradient in the VFs, and the zipper-like closure commonly observed in female phonation (Birkholz *et al.*, 2011a,b). The TBCM formulation considers the structural link between the membranous area and the posterior glottal opening extended onto the cartilaginous glottis (Erath *et al.*, 2013; Zañartu *et al.*, 2014).

Originally, the TBCM was built around the simplification that muscle control only affects the VF adjustment, so the arytenoid accommodation was independently parameterized through *ad hoc* rotation and displacement parameters. In this study, on the contrary, it is assumed that arytenoid accommodation is solved within the laryngeal posturing framework described in Sec. II A. Hence, the adductory displacement in the posterior border of the upper mass, relative to the anterior commissure, coincides with the VP horizontal coordinates x_{02} in Eq. (3a) (see Fig. 2). For the adductory displacement of the lower mass, x_{01}, the net glottal convergence is also considered following the ideas in Titze and Story (2002).

The coupled equations for simulating the TBCM dynamics are

$$F_u = m_u \ddot{x}_u = F_{k,u} + F_{d,u} - F_{kc} + F_{e,u} + F_{Col,u}, \quad (6a)$$

FIG. 2. (Color online) Schematic of the triangular body-cover model of the vocal folds.

$$F_l = m_l \ddot{x}_l = F_{k,l} + F_{d,l} + F_{kc} + F_{e,l} + F_{Col,l}, \quad (6b)$$

$$F_b = m_b \ddot{x}_b = F_{k,b} + F_{d,b} - [F_{k,u} + F_{d,u} + F_{k,l} + F_{d,l}], \quad (6c)$$

where m is mass, x is the medial-lateral displacement over time, and F is the force component for each block. Rest positions for the upper and lower blocks are $x_{u,0} = 0.5\,x_{02}$, $x_{l,0} = 0.5\,x_{01}$ pursuant to VF posturing, whereas $x_{b,0} = 3$ mm (Galindo *et al.*, 2017). Furthermore, force subscripts k, d, e, and kc represent the mechanical forces produced by the springs, dampers, flow pressures, and elastic coupling between the upper and lower masses, respectively. An additional spring force, F_{Col}, is introduced during vocal fold collision to capture the effects of impact between opposing upper/lower cover masses. The definitions of the forces are described in detail in the appendix of Galindo *et al.* (2017) and are thus omitted here for brevity. To compute the aerodynamic driving forces, the intraglottal pressures on the upper/lower cover masses are derived from the subglottal pressure, P_s, and the supraglottal pressure, P_e, according to the formulation introduced in the appendix of Titze (2002).

C. Physiological rules for the triangular body-cover model

The empirical rules introduced by Titze and Story (2002) are applied and extended for the computation of the TBCM parameters. These rules introduce dynamic muscle control based on mapping the normalized activation levels of the intrinsic muscles, introduced at the end of Sec. II A, into the relevant mechanical parameters (e.g., linear stiffness, mass distribution, glottal convergence, and VF length, thickness, and depth). Besides controlling the TBCM vibrations, the rules also have an impact on other aerodynamic and acoustic quantities in scenarios involving the tissue-flow-acoustic interaction in the glottis (Lowell and Story, 2006; Zañartu *et al.*, 2014).

Originally, Titze and Story (2002) begin by prescribing the VF elongation and VP adduction as a function of normalized activation levels a_{TA}, a_{CT}, and a_{LC}, with the latter combining the effects of both the LCA and PCA muscles. The geometrical parameters and fiber stresses are then obtained from the VF elongation, after which the stiffness and mass distributions in the VF model are computed. In this work, a few methodological modifications are considered to extend the rules for the TBCM. The first difference consists of including independent activation levels for the five intrinsic muscles; hence, the activation set $\mathbf{a} = [a_{LCA}, a_{IA}, a_{PCA}, a_{CT}, a_{TA}]$ controls the activation of the laryngeal musculature. Moreover, internal stresses in the VF tissues, specifically for the TA muscle, LIG, and MUC, are dynamically obtained by solving a modified Kelvin model (see the Appendix), whereas the VF elongation and VP adduction are obtained from the laryngeal posturing according to Eqs. (5) and (3a), respectively; hence, the elongation and adduction rules in Titze and Story (2002) are replaced. With this information, the remaining rules in
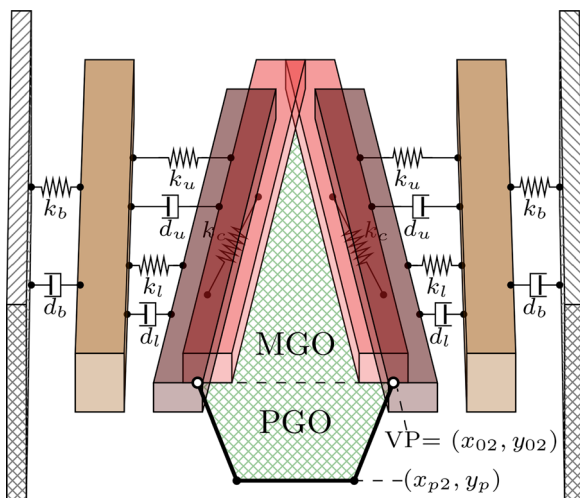
20   J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*

Titze and Story (2002) are thus applied for computing the nodal point, thickness, and depth for each block, and glottal convergence, from which the biomechanical parameters (i.e., mass and spring values) of the TBCM are derived. Notably, the shear moduli of the body and cover are set to $\mu_b = 600$ Pa and $\mu_c = 300$, respectively, which differs from the values selected by Titze and Story (2002).

Based on VF adduction and symmetry with respect to the midsagittal plane, the glottal area for the upper and lower cover masses are

$$A_u = 2(1 - \alpha_u)L_g(\tilde{x}_u + 0.5(1 + \alpha_u)x_{01}), \tag{7a}$$

$$A_l = 2(1 - \alpha_l)L_g(\tilde{x}_l + 0.5(1 + \alpha_l)x_{02}), \tag{7b}$$

where $\tilde{x}_u = x_u - x_{u,0}$ and $\tilde{x}_l = x_l - x_{l,0}$ are block displacements relative to their rest positions. Additionally, α_u and α_l are the proportions of mass length for the upper and lower blocks undergoing collision at the given time, where $0.0 \le \alpha_u, \alpha_l \le 1.0$. As a result, the area for the membranous glottal opening is $A_{MGO} = \min\{A_u, A_l\}$. The effects of laryngeal posture on the posterior cartilaginous portion of the glottis are also simulated. Following Titze (2006), a trapezoidal shape is assumed for the posterior glottal opening, thus the resulting area can be computed as follows:

$$A_{PGO} = \max\{0, \min\{(x_{p1} + x_{01}), (x_{p2} + x_{02})\}(y_{02} - y_p)\}, \tag{8}$$

where x_{p1} is the posterior wall half-width at the bottom, x_{p2} is the posterior wall half-width at the top, and y_p is the posterior wall position along the longitudinal axis. The total glottal area comprises both the membranous and the cartilaginous parts,

$$A_g = A_{MGO} + A_{PGO}. \tag{9}$$

D. Interactions at the glottis and acoustic wave propagation

To capture the physics of human phonation, the three-way interaction between sound, flow, and VF tissue is considered. For computing the air volume velocity, U_g, through the glottal area, A_g [see Eq. (9)], the solver proposed in Zañartu et al. (2014) considering the acoustic driving pressures and the posterior glottal opening is applied, with the inclusion of the corrections made by Lucero and Schoentgen (2015). As shown in Zañartu et al. (2014), in the one-dimensional flow approximation the solution obtained for a domain with two separate orifices for the posterior and membranous areas is equivalent to solving for the volume velocity through the total glottal area, A_g.

The wave reflection analog scheme is selected for describing the propagation of one-dimensional, planar acoustic waves in the time domain throughout the equivalent subglottal and supraglottal systems (Zañartu, 2006). These tracts are discretized as the concatenation of a finite number of short uniform cylinders with variable cross-sectional areas. The area functions obtained from magnetic resonance imaging data for a male participant during sustained phonation are applied to simulate the supraglottal tract for different vowels (Story, 2008) and the subglottal system (Zañartu, 2006). For simulating an equivalent excised-larynx scenario within the same framework, the non-interactive control scenario introduced by Titze and Palaparthi (2016) is also considered, where a very wide cross-sectional area ($30\,\text{cm}^2$) is set for every cylinder in both the subglottal and the supraglottal systems. Boundary conditions based on a resistive lung termination and a parallel resistive-inertive element at the lips are also defined in the simulations.

E. Numerical implementation

The implemented model involving the muscle control of the larynx posture and vocal fold function depends upon several anatomical and biomechanical parameters. Model parameters were selected to simulate the physiology of male phonation. The parameters considered in this work are reported in Table I. As in Galindo et al. (2017), the simulations were performed using a truncated Taylor-series approximation to solve the differential equations. A sampling frequency of 44.1 kHz was employed. A colored random source component was included for modeling the turbulent aspiration noise generated at the glottis (Galindo et al., 2017).

III. RESULTS

This section introduces illustrative simulations with the proposed voice production model. The effects of intrinsic muscle activation on the laryngeal posture are first described, then simulations for sustained vowels and articulatory gestures are subsequently analyzed. An example resembling antagonistic muscle behavior during sustained phonation is also discussed.

TABLE I. Anatomical and biomechanical parameters required for implementing the dynamic simulation of the laryngeal posturing and glottis configuration.

	Definition	Value
CAJ	Mass and moment of inertia	$M_{ac} = 1.4 \times 10^{-3}$ kg, $I_{ac} = 1.6 \times 10^{-6}$ kg m^2.
	Vocal process cadaveric position	$\bar{x}_0 = 4$ mm, $\bar{y}_0 = 0$ mm.
	CAJ center coordinates	$x_{CAJ} = 10.1$ mm, $y_{CAJ} = -10.1$ mm.
	Translational/rotational dampings	$d_x = 0.02$ s, $d_y = 0.02$ s, $\delta = 0.02$ s.
CTJ	Mass and moment of inertia	$M_{cc} = 10^{-2}$ kg, $I_{cc} = 10^{-5}$ kg m^2.
	CT/TA moment arms	$r_{TA} = 16.1$ mm, $r_{CT} = 11.1$ mm.
	CT angle relative to TA	$\phi = 45°$, $\cos\phi = 0.76$.
	Translational/rotational viscous times	$t_t = 0.04$ s, $t_r = 0.04$ s.
Glottis	Rest VF length	$L_0 = 16$ mm.
	Posterior wall half-width coordinates	$x_{p2} = 3.2$ mm, $y_p = -2.5$ mm.

J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi et al.   21

A. Vocal fold posturing

The biomechanical scheme for the prephonatory laryngeal posturing utilizes, principally, the arytenoid cartilage accommodation for controlling glottal adduction and abduction. Figure 3 shows the effect of the individual activation of the five intrinsic muscles on the (right) arytenoid posture, as described by the position of the cricoarytenoid joint and the vocal process. Figure 3 also includes an equivalent LCA/IA *adductory* complex given by $a_{LCA} = a_{IA} = a_{Add}$, i.e., the coupled 1:1 activation for LCA and IA muscles, as in previous studies (Chhetri *et al.*, 2012, 2014; Geng *et al.*, 2020; Palaparthi *et al.*, 2019). The muscle activation is incremented in normalized steps of 0.1 from 0 to 1 for each case. Note that the displacement of the cricoarytenoid joint is smaller and different in nature than those of the vocal process.

Herein, VF adduction is characterized by the medialization and (positive) counterclockwise rotation of the arytenoid cartilage, and is controlled by the LCA muscle and, to a lesser extent, the TA muscle. Glottal abduction is achieved *via* the PCA muscle where concurrent vocal process lateralization and clockwise arytenoid rotation is observed. Furthermore, anterior-posterior displacements of the vocal process are determined by the antagonist effects of the TA and CT muscles. Although activation of the IA muscle minimally alters the position of the vocal process, it does displace the cricoarytenoid joint medially and caudally, thus reducing the posterior glottal opening. The adductory complex combines both the LCA and IA muscles, thus allowing for jointly reducing the membranous and posterior glottal areas.

The effects on VF adduction due to muscle coactivation in the model are illustrated in Fig. 4. The resulting movement of the vocal process (refer to the attached schematic) from the incremental activation of the adductory complex is shown for different coactivation states of the remaining muscles. For each path, activation levels for a_{Add} increased from 0 to 1 in normalized steps of 0.1. The resulting movement of the vocal process with no coactivation is drawn in black in the center of the figure, whereas the two topmost black paths and the two bottommost black paths describe the coactivation with the TA and CT muscles, respectively. Results involving two cases of PCA coactivation are also included for contrast.
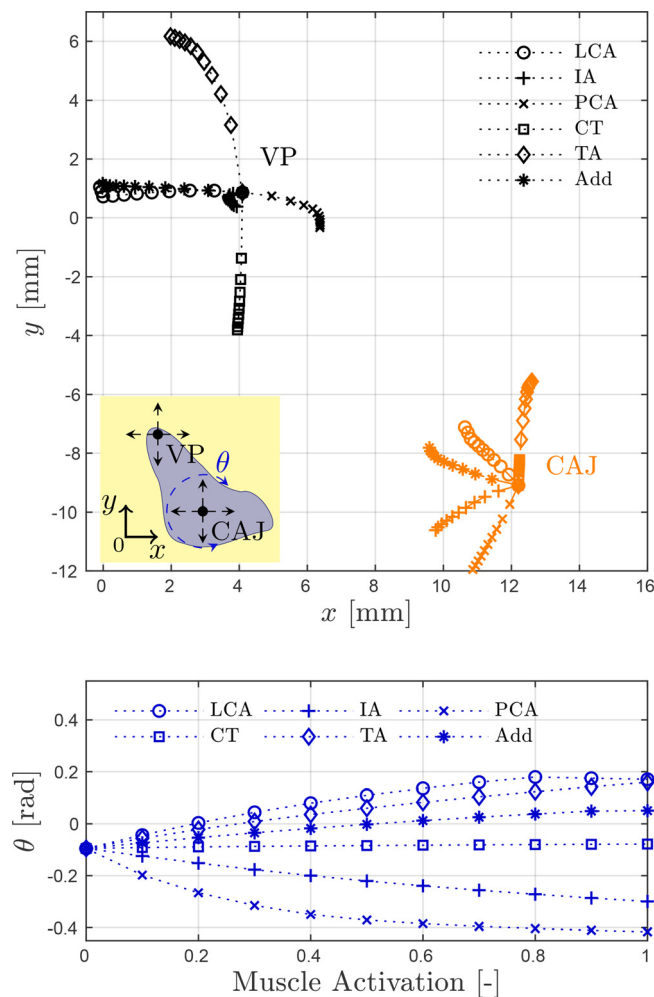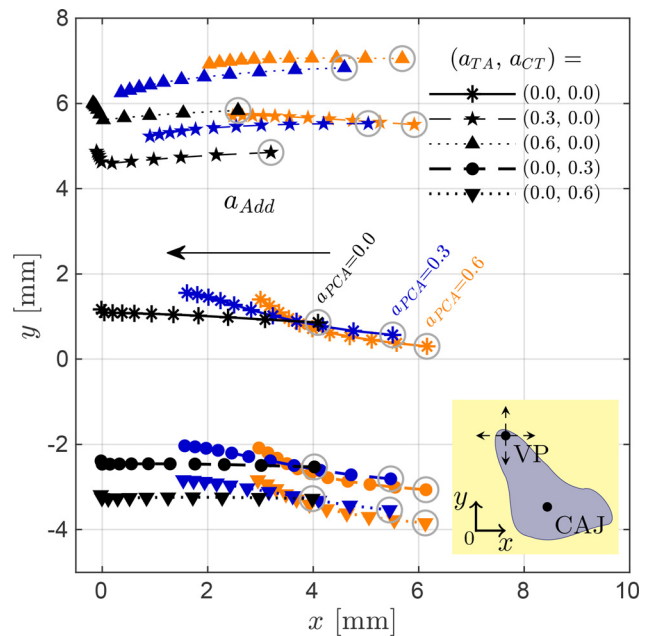

FIG. 3. (Color online) Simulated accommodation of the (right) arytenoid cartilage obtained by the independent activation of the five intrinsic muscles and the adductory complex. *Top:* Cartesian displacement of the cricoarytenoid joint (CAJ) center, $(x_{CAJ} + \xi, y_{CAJ} + \psi)$, and the vocal process (VP), (x_{02}, y_{02}). The inset schematic illustrates the VP movements resulting from the displacement and rotation of the CAJ. *Bottom:* Rotation angle θ for the CAJ. Concurrent beginning of the paths indicates zero muscle activation.

FIG. 4. (Color online) Effects of laryngeal muscle coactivation on glottal adduction. Vocal process (VP) coordinates, (x_{02}, y_{02}), produced through the activation of the adductory complex for the non-coactivation case (solid line with medium markers), and for two activation levels for both TA muscle (dashed and dotted lines with small markers) and CT muscle (dashed and dotted lines with large markers). Three PCA activations are drawn in black ($a_{PCA} = 0.0$), dark color ($a_{PCA} = 0.3$), and light color ($a_{PCA} = 0.6$) lines. Markers enclosed in circles indicate the results for null adductory complex activation ($a_{Add} = 0.0$). The inset schematic illustrates the VP Cartesian movements.

22   J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*

These results further illustrate the antagonist role of the TA and CT muscles, where ventral and dorsal displacements of the vocal process are produced *via* the activation of the TA and CT muscles, respectively, and where the TA muscle increases the total adduction of the vocal process. Moreover, Figure 4 also illustrates that coactivation of the PCA muscle introduces a noticeable lateral displacement in the vocal process path with increasing activation, and slightly changes the direction of its path.

In Fig. 5, muscle activation maps depicting VF elongation and vocal process distance for pairs of intrinsic muscle groups are presented. VF elongation is computed by Eq. (5), whereas the distance between the vocal processes is measured with respect to the rest distance at zero muscle activation. A coupled 1:1 activation was applied for every muscle in a group, for example, an activation level $\bar{a}$ set to group

LCA/IA implies $a_{LCA} = a_{IA} = \bar{a}$. The coactivation of the main phonatory laryngeal muscles is investigated in Fig. 5, whereas the effect of PCA activation is described further bellow. Similar representations have been applied in the past for investigating muscle groups having antagonistic or synergistic functions in VF posturing in experiments with excised canine larynges, (Chhetri *et al.*, 2012; Chhetri *et al.*, 2014), and in simulations involving a three-dimensional finite element model of a canine larynx (Geng *et al.*, 2020). Given that the proposed biomechanical model of the larynx was partially based on canine models (Titze and Hunter, 2007), contrast against these prior studies is feasible and meaningful.

Results for VF elongation (from VF strain in the left column of Fig. 5) show that the proposed model captures the antagonistic effects of the CT and TA muscles, where VF strain can be increased/decreased through corresponding CT/TA muscle activations. It can be observed that the main adductor muscles have a minor effect on VF elongation. The largest VF elongation (around +20%) is obtained from maximum CT activation and zero TA activation. Similar strain values were obtained in simulations using finite element models (Geng *et al.*, 2020), whereas studies with canine models have reported larger elongation values (Chhetri *et al.*, 2012; Chhetri *et al.*, 2014). In contrast, the shortest VF elongation results from maximum TA activation and zero CT activation. The minimum VF strain values (around −35%) are lower than those reported in the benchmark studies. However, the strain contour lines for all activation conditions resemble the diagonal patterns reported by other authors (Chhetri *et al.*, 2012; Geng *et al.*, 2020).

Vocal process distance indicates the adduction of the vocal folds. The plots in Fig. 5 show the resulting values for the distance between the vocal processes, measured in percentage relative to the distance at rest (i.e., null activation for all intrinsic muscles). As expected, a strong synergistic adductory role played by the TA, LCA, and IA muscles is noticeable. Individual TA activation has a less significant effect on glottal adduction than the coactivation of the LCA/IA complex. However, the equal activation of the TA, LCA, and IA muscles as a group compresses the previous gradual posturing of the vocal processes at the region of lower activation levels. Despite the considerable anatomical and functional oversimplifications in the biomechanical model of the larynx, our simulations of vocal process distance are in agreement with previous experimental and high-order modeling studies (Chhetri *et al.*, 2012; Chhetri *et al.*, 2014; Geng *et al.*, 2020). Two important differences are worth noting. First, the previous studies indicate that a non-zero vocal process distance is obtained even for strong cases of glottal adduction (Geng *et al.*, 2020; Moisik and Gick, 2017). The implemented model, however, produces zero distance values for a range of muscle activation conditions. In addition, the current laryngeal model is not capable of reproducing the abductory effect of the CT muscle on VF adduction reported by other authors (Chhetri *et al.*, 2012; Chhetri *et al.*, 2014; Geng *et al.*, 2020).
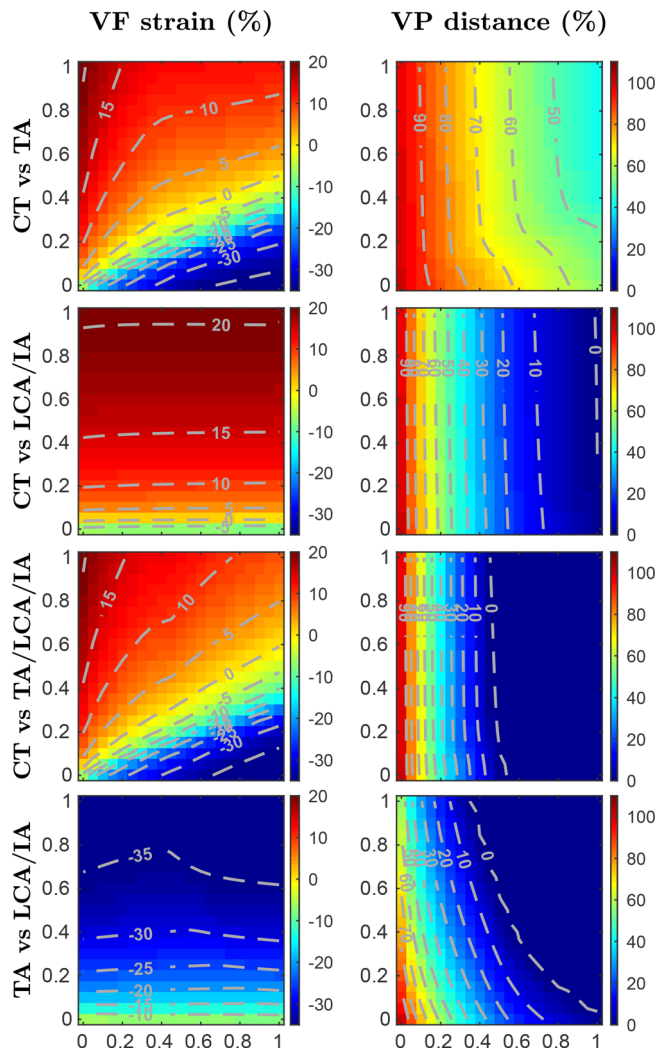

FIG. 5. (Color online) Muscle activation plots for parametric coactivation of main phonatory intrinsic muscles. Vocal fold strain (left column) and vocal process distance (right column) are depicted as functions of the paired coactivation of muscle groups. The rows show four activation scenarios with contour lines included for clarity. For each row, A vs B refers to the abscissa and ordinate axes indicating activation levels for muscle groups A and B, respectively; null activations are set for the remaining intrinsic muscles.

J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*   23

The PCA muscle is the main larynx abductor, with a key role during inspiratory laryngeal movements, whereas during typical phonation, it plays only a minor role (Poletto *et al.*, 2004). Muscle activation plots portraying the VF elongation and vocal process distance for adductory/abductory antagonist muscle groups are depicted in Fig. 6, following the previously mentioned methodology. The first row shows the results from the coactivation of the PCA muscle versus LCA/IA group. It is observed that the distance between the vocal processes ranges from nearly 0% for full activation of the LCA/IA group and no activation of the PCA muscle, to around 150% for the opposite case, whereas the changes in VF strain are minimal. Notice that, in the case of equal activation, the abductor force from the PCA muscle is slightly superior to the antagonist force produced by combining the LCA and IA muscles. The second row illustrates the results obtained by including the TA muscle in the adductor group. The results show that the TA muscle acts to reinforce the adductor forces that are opposite to the PCA muscle action, producing an appreciable region with null vocal process distance; moreover, TA muscle inclusion yields a dramatic VF shortening of up to −45% strain for full activation of the four muscles considered. Further comparisons depict that the PCA and CT muscles actuate differently during larynx accommodation, providing evidence of no antagonist effect; the PCA muscle abducts the vocal process producing no appreciable change in VF strain, whereas the CT muscle mainly elongates the VF with a minimal modification of the distance between the vocal processes. These results are not included here for conciseness.
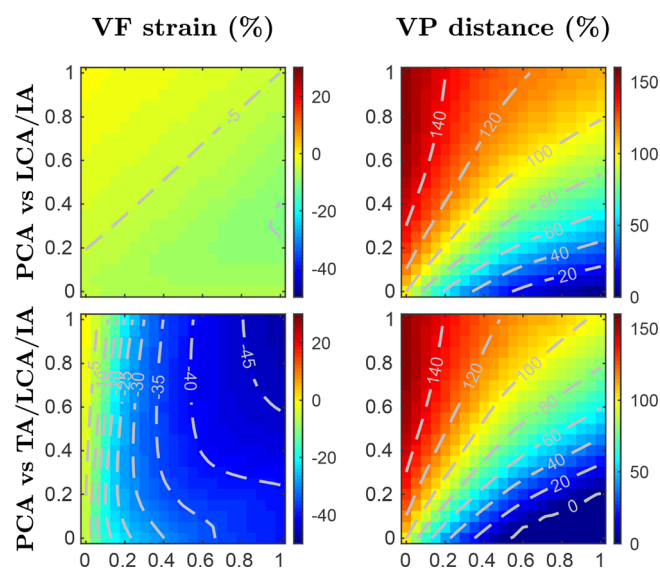

FIG. 6. (Color online) Muscle activation plots for parametric coactivation of adductory/abductory intrinsic muscles. Vocal fold strain (left column) and vocal process distance (right column) are depicted as functions of the paired coactivation of muscle groups. The rows show two activation scenarios with contour lines included for clarity. For each row, A vs B refers to the abscissa and ordinate axes indicating activation levels for muscle groups A and B, respectively; null activations are set for the remaining intrinsic muscles.

B. Sustained phonatory gestures

Sustained vowel simulations are performed to investigate basic glottal function (self-sustained oscillation and pitch) as a function of the laryngeal muscle activation and subglottal pressure, P_s. For simplicity, the prephonatory laryngeal posture is fixed through the constant activations $a_{LCA} = a_{IA} = 0.5$ and $a_{PCA} = 0.0$ for the main adductor/abductor muscles, whereas VF adjustment is controlled *via* the parametric activation of a_{CT} and a_{TA} in the range from 0 to 1.

For each condition, 600 ms of simulated phonation data were obtained. In general, the simulations evidenced transients and fluctuations at the beginning, which were not relevant for the present analysis; thus, the first portion of data was discarded and only the last 200 ms steady segment was considered. An autocorrelation method was applied for assessing periodicity and fundamental frequency for the simulated glottal area waveform, A_g (Mehta *et al.*, 2015). In Fig. 7, muscle activation plots are shown for fundamental frequency as a function of TA and CT muscle coactivation, for three physiological subglottal pressure levels $P_s = [0.8, 1.4, 2.0]$ kPa. Acoustic coupling for vowels {/æ/, /i/, /a/, /u/, /o/, /e/} as well as for a non-interactive acoustic system was also considered to further assess the model behavior. Results show that in the low subglottal pressure condition, increasing CT activation leads to less stable, or even unachievable, VF oscillations (shown as blank portions in Fig. 7). Self-sustained oscillations are strengthened by increasing the subglottal pressure level. The relation between high CT activation and phonation onset pressure has been described in excised larynx experiments (Chhetri *et al.*, 2014), as well as in high-order simulations (Palaparthi *et al.*, 2019). Results also show that the acoustic coupling facilitates self-sustained VF oscillations and raises pitch in comparison with the no vocal tract (non-interactive) scenario, where an increased region without periodic vibrations was observed.

Fundamental frequency contours resulting from coactivation of the TA and CT muscles globally resemble those reported in Titze and Story (2002) and Lowell and Story (2006). CT activation is the main contributor to changes in fundamental frequency and the highest values are obtained for low values of TA activation. This result was expected and it has been reported before (Lowell and Story, 2006; Titze and Story, 2002). However, the original rules for muscle control of the body-cover VF model yield fundamental frequencies greater than 500 Hz, whereas the frequency range in Fig. 7 is more restricted, showing maximum values around 350 Hz. Fundamental frequencies up to 400 Hz are obtained for scenarios with increased gottal adduction. In addition, prior reported muscle activation plots featured a noticeable downward bending of the contour lines in the lower left region (low CT and TA). The implemented model also shows this non-linear response for the fundamental frequency in the region of weak TA activation and moderate-to-strong CT activation; however, bending in the contour
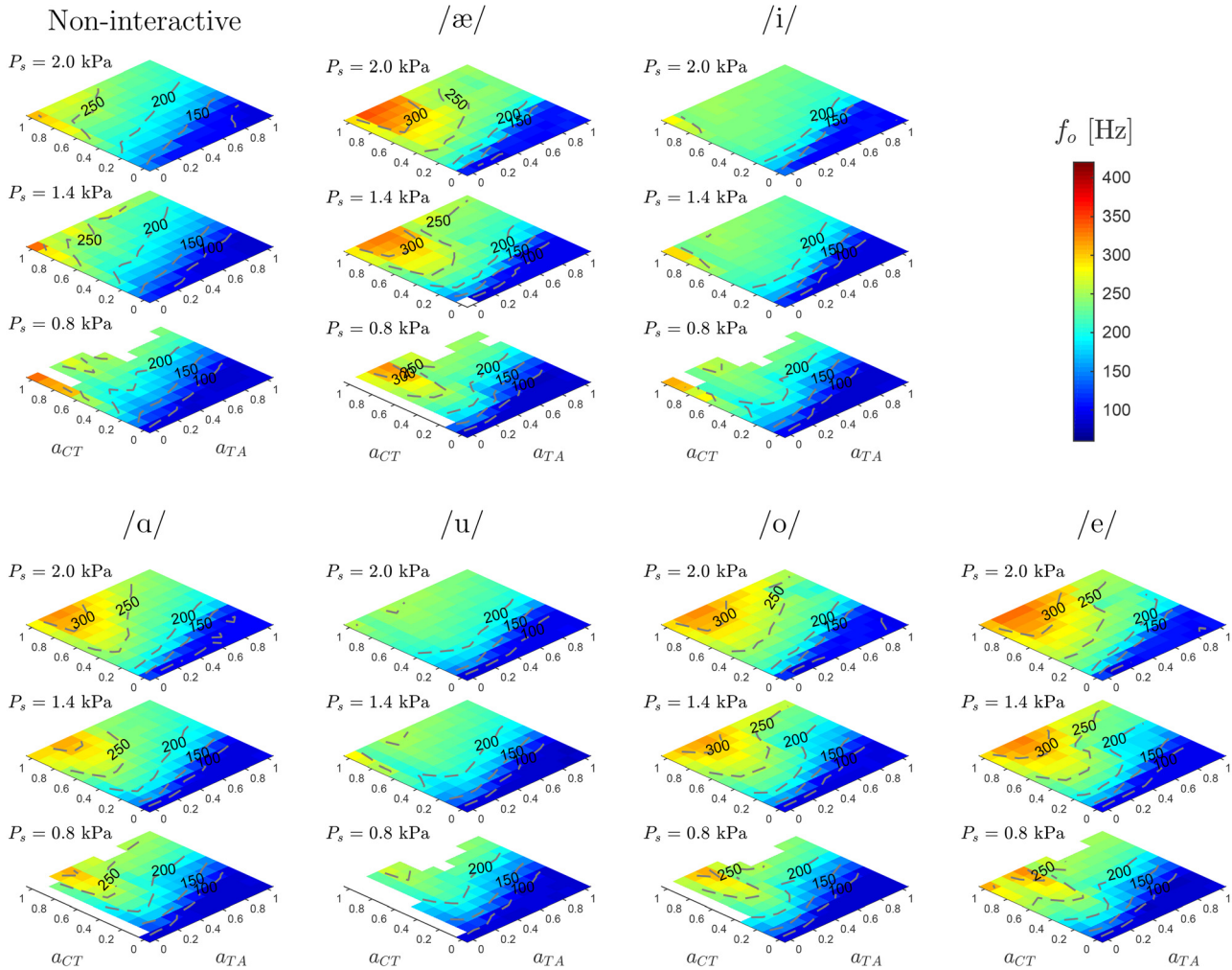
24   J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*

FIG. 7. (Color online) Muscle activation plots with fundamental frequency, f_0, for CT versus TA activation, considering different vocal tract shapes and subglottal pressures $P_s = [0.8, 1.4, 2.0]$ kPa. In all the cases $a_{LCA} = a_{IA} = 0.5$, and $a_{PCA} = 0.0$. The non-interactive case has no vocal tract and represents an excised larynx scenario. Isofrequency contours are drawn for clarity.

lines is less pronounced in comparison with the results reported by Titze and Story (2002). This is more noticeable for the non-interactive vocal tract case.

C. Dynamic phonatory gestures

Simulations of dynamic (articulatory) gestures combine posturing and phonation and are key to assessing the overall performance of the model. Simulations of repetitive voicing-devoicing gestures, i.e., /hi-hi-hi-hi/, are shown in Fig. 8. Model parameters are set to produce a modal male voice, where $P_s = 800$ Pa, $a_{TA} = 0.2$, $a_{CT} = 0.1$, and $a_{PCA} = 0.0$, for a (male) vowel /i/ for a direct comparison with Titze and Hunter (2007).

As before, the equivalent LCA/IA adductory complex is controlled in time with $a_{LCA} = a_{IA} = a_{Add}$. The three columns in Fig. 8 show simulations for weak (left), moderate (middle), and strong (right) VF adduction, respectively. Top panels depict the muscle control signal a_{Add} for every case, with peak amplitude indicating the adduction level. Waveform outputs of glottal area A_g, glottal volume velocity

U_g, and radiated pressure P_o for the three simulated adduction strengths are shown in the following three rows, respectively. Spectrograms describing time-frequency content for the radiated pressure signals are included in the bottom row.

Despite the overly simple and fairly square activation signals, the simulated variables capture the gradual VF adduction/abduction characteristic of the voicing/devoicing gesture. The noticeable delay in area and airflow waveforms with respect to variations in the control signal, a_{Add}, is a product of the combined effects of inertia in the biomechanical model of the larynx and the time constants in the laryngeal tissue model. Radiated sound pressure is further delayed as a consequence of acoustic wave propagation throughout the vocal tract.

Changes in VF adduction are primarily seen in the voice onset time and, to a lesser extent, in the voice offset time. A higher magnitude in a_{Add} speeds up the onset of VF oscillations and delays the offset of voicing. VF adduction condition also affects phonation time and the spectral richness of the radiated sound. The latter is in accordance with previous studies illustrating that, for a constant driving pressure,
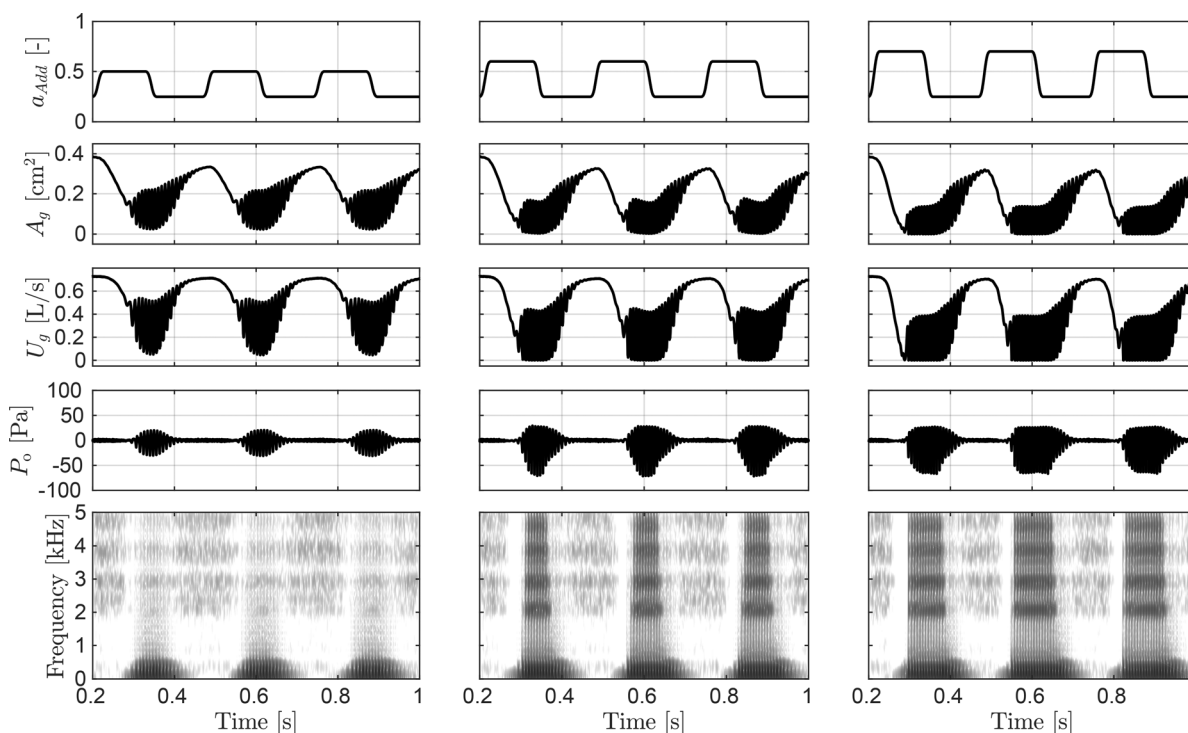
FIG. 8. Simulated muscle control of voicing-devoicing for a /hi-hi-hi-hi/ gesture. From left to right, columns correspond to simulations for weak, moderate, and strong VF adduction with maximum activation levels 0.5, 0.6, and 0.7, respectively. Adductory signals a_{Add} are shown in the top row. The following three rows show the output signals for the glottal area A_g, glottal volume velocity U_g, and the radiated acoustic pressure P_o. The bottom row depicts the wideband spectrogram of P_o.

weak to moderate adduction conditions modify PGO area in the range of small to large, which in turn produces a considerable reduction of the radiated sound pressure level (Galindo *et al.*, 2017; Zañartu *et al.*, 2014).

Note that simulations are built around muscle control signals for the LCA and IA muscles only. However, it is well known that during speech production all the laryngeal muscles coactivate in complex ways (Moisik and Gick, 2017; Movahhedi *et al.*, 2021; Zhang, 2016). As illustrated before, other muscles, such as PCA, TA, and CT, have an effect on VF adduction and can also alter voicing onset and offset (Poletto *et al.*, 2004; Titze and Hunter, 2007). The proposed model allows for incorporating all five muscle activations in a time varying fashion, which can improve the physiological relevance when modeling articulatory gestures. Nevertheless, the resulting output with the simplified muscle control input is still in agreement with Titze and Hunter (2007) and Poletto *et al.* (2004).

D. Antagonistic muscle behavior

As the proposed model allows for analyzing the effects of selected levels of antagonistic muscle activation, we herein simulate conditions that have the same vocal fold posturing, but with different underlying antagonistic muscle activations. We argue that this exercise can lead to future studies exploring disproportionate coactivation of antagonist intrinsic muscles that can be related to NPVH.

The ability of the model to attain the same prephonatory VF posturing for various muscle activation conditions is illustrated in Fig. 9, where the displacement paths for the right vocal process for three simulated laryngeal conditions are drawn. The dotted rectangle in Fig. 9 shows that three distinctly different activation conditions (reported in Table II) can produce almost the same vocal process position, and result in a visually similar VF configuration. This idea can also be extended into paths of glottal adduction. Figure 9 shows three paths obtained when varying the adductory complex a_{Add} that correspond to different activations of all intrinsic muscles. These paths compare a given prephonatory state with two other conditions involving lower and higher activation levels, respectively. A slight increase in abduction is seen in the latter case for each marker (displacement to the right) do to PCA activation; although the paths overlap, this indicator shows that the increased tension scenario yields a more abducted condition for every step.

Even though VF posturing can be visually preserved, changes in muscle activation are expected to alter the TBCM parameters differently and, thus, the resulting vocal function. To investigate this idea, sustained phonation is simulated for the three activation sets (denoted as a_I, a_{II}, and a_{III} in Table II) that result in the same VF posture although associated with different degrees of coactivation of the antagonist intrinsic muscles. For these simulations, a driving pressure $P_s = 800$ Pa is applied, and tract area functions for ten different vowels.

To describe multiple aspects of vocal function for each set, various parameters were computed, including (1) VF
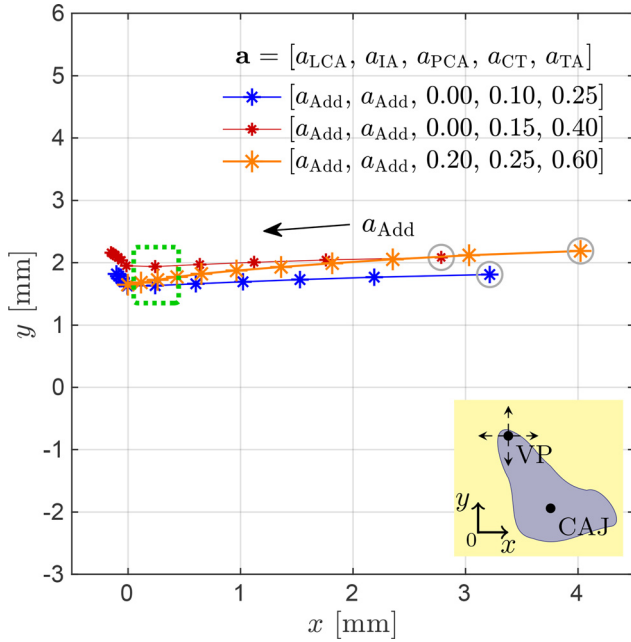
FIG. 9. (Color online) Displacement paths for the right vocal process (VP), (x_{02}, y_{02}), for three simulated tension states in the larynx. Trajectories correspond to laryngeal postures produced by parametric activation of the intrinsic muscles following the activation set $\mathbf{a} = [a_{LCA}, a_{IA}, a_{PCA}, a_{CT}, a_{TA}]$. The paths describe variations for the normalized activations a_{Add} in the range of 0 to 1 in steps of 0.1. The inset schematic illustrates the VP Cartesian movements. Configurations enclosed in the dotted rectangle are considered for voiced sound simulations.

posturing: VF strain, vocal process distance, and membranous and posterior portions of the glottal area; (2) VF kinematics: fundamental frequency (f_0), open quotient (OQ), and closing quotient (ClQ); (3) Glottal aerodynamics: unsteady glottal airflow (AC Flow), stable glottal flow (DC Flow), maximum flow declination rate (MFDR), and amplitude of the first harmonic relative to the second harmonic ($H_1 - H_2$); (4) Voice acoustics: sound pressure level (SPL),

TABLE II. Three sets of appreciable different intrinsic muscle activation levels producing nearly the same simulated vocal process configurations.

Set	a_{LCA}	a_{IA}	a_{PCA}	a_{CT}	a_{TA}
$\mathbf{a}_I$	0.50	0.50	0.00	0.10	0.25
$\mathbf{a}_{II}$	0.40	0.40	0.00	0.15	0.40
$\mathbf{a}_{III}$	0.80	0.80	0.20	0.25	0.60

and low-to-high (L/H) spectral ratio. These measures have been shown to be valuable for the assessment of vocal hyperfunction (Cortés *et al.*, 2018; Galindo *et al.*, 2017; Mehta *et al.*, 2019; Mehta *et al.*, 2015; Zañartu *et al.*, 2014).

The resulting measures of vocal function are reported in Table III. As expected, given that the vocal process position is almost the same, the parameters describing the laryngeal posture are similar for the three conditions; the only evident difference is the increased PGO area for set $\mathbf{a}_{II}$ that results from the lower activation of the LCA and IA muscles. Moreover, a significant increase in the minimum glottal area is observed for the $\mathbf{a}_{II}$ condition, which leads to a higher DC Flow. Kinematic measures in Table III show that fundamental frequency and closing quotient show minor changes with the activation conditions, whereas OQ is noticeably increased for the $\mathbf{a}_{II}$ condition.

Aerodynamic measures in Table III show that the highest MFDR values are obtained, as expected, for the more adducted cases, i.e., with less PGO leakage. Conversely, increasing the PGO area decreases MFDR, suggesting less efficient acoustic excitation. Interestingly, AC Flow and $H_1 - H_2$ show minor variations, suggesting negligible differences in the vocal tract coupling and the source spectral tilt. In addition, acoustic measures provide complementary evidence of the reduced phonatory efficiency for set $\mathbf{a}_{II}$, yielding a noticeable decrease in SPL levels together with increased L/H spectral ratios, which are signs of less spectral energy in the high frequency region.

TABLE III. Effects of three different activation conditions on sustained phonation. Reported prephonatory posture parameters were obtained from the biomechanical larynx model. Kinematic, aerodynamic, and acoustic mean (SD) parameters were computed from simulated phonatory data considering tract area functions for vowels [æ ʌ ɑ e ɛ ɪ i o u ʊ] and driving pressure $P_s = 800$ Pa.

		Activation sets		
		$\mathbf{a}_I$	$\mathbf{a}_{II}$	$\mathbf{a}_{III}$
Posture	Strain ϵ [%]	−9.5	−11.3	−10.3
	VP Distance [%]	5.3	6.0	6.3
	A_{MGO} [mm^2]	0.8	0.9	1.0
	A_{PGO} [mm^2]	1.4	3.0	1.1
Kinematic	f_0 [Hz]	116.4 (4.7)	105.1 (4.4)	115.1 (3.7)
	OQ [%]	81.7 (7.6)	98.6 (3.8)	79.7 (11.7)
	ClQ [%]	24.0 (8.5)	33.9 (5.9)	25.5 (8.4)
Aerodynamic	AC Flow [mL/s]	395.2 (34.1)	367.6 (35.0)	372.2 (28.1)
	DC Flow [mL/s]	11.1 (1.1)	65.6 (1.7)	6.1 (1.0)
	MFDR [L/s^2]	710.6 (146.9)	375.7 (92.7)	648.4 (133.2)
	H1-H2 [dB]	12.9 (1.5)	12.7 (1.1)	12.2 (1.4)
Acoustic	SPL [dB]	84.9 (3.5)	77.2 (3.5)	84.5 (3.4)
	L/H Ratio [dB]	31.0 (14.1)	41.4 (4.7)	31.7 (14.1)

J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*   27

IV. DISCUSSION

The proposed scheme allows for controlling prephonatory posturing and phonatory gestures in a triangular body-cover model of vocal folds by means of the independent activation of all of the intrinsic laryngeal muscles. It combines prior efforts to model laryngeal posturing (Titze and Hunter, 2007), empirical rules relating muscle activation to biomechanical parameters in a lumped-element model (Titze and Story, 2002), and a triangular body-cover VF model (Galindo et al., 2017). The scheme thus provides a flexible and physiologically relevant way to control the self-sustained fully interactive voice production model for both sustained vowels and time-varying glottal gestures. As a result, the approach allows for exploring the role of antagonistic muscle pairs in phonation, which has direct implications for studying normal and disordered muscle behavior in phonation.

In this study, it was first shown that the implemented framework can reproduce prior findings with excised larynx experiments and high-order numerical simulations for sustained vowels and simple articulatory gestures. In addition, we applied the approach to exemplify the role of antagonistic muscle pairs and the effects of antagonistic muscle activation in the context of a common manifestation of NPVH, with simulations in agreement with prior clinical data. Thus, despite its low-order complexity, the scheme provides a physiologically-inspired tool that could be used to gain insights into the physical mechanisms underlying intrinsic muscle function in phonation. This becomes particularly useful for extending previous efforts to model vocal hyperfunction (Galindo et al., 2017; Zañartu et al., 2014) to account for elevated, unbalanced, and poorly regulated activity of the intrinsic laryngeal muscles, as seen in NPVH (Hillman et al., 2020). We hypothesize that NPVH is associated with a disproportionate coactivation of antagonist intrinsic muscles that result in a more abducted prephonatory VF posturing, which in turn results in a reduced loudness that is compensated with elevated subglottal pressure and further muscle tension. Subsequent efforts will explore the use of the proposed model to investigate these ideas and aid in delineating the etiology and pathophysiology of vocal hyperfunction. The model has also potential for developing improved estimation frameworks based on biomechanical models of human phonation applicable to the prevention and treatment of VH (Alzamendi et al., 2020; Ibarra et al., 2021).

An evident limitation regarding the implemented muscle control scheme is worth highlighting. The effects of the coupled activation of the five intrinsic muscles on the glottal posture and vocal fold configuration are simulated by applying the biomechanical model by Titze and Hunter (2007) at a low-to-moderate computational cost. However, this model is built around the effective displacement and rotation of the laryngeal cartilages on a transverse glottal plane, and is not thus able to describe in full detail the three-dimensional (3D) glottal adjustment, in contrast

with, for example, the anatomically accurate finite element models. Current evidence supports that the arytenoid movement relative to an axis located at the CAJ and perpendicular to the transverse glottal plane [see Fig. 1(a)] is highly constrained and that a rocking-sliding motion about the long axis of the cricoid cartilage is more accurate (Geng et al., 2020; Yin and Zhang, 2014). The latter can explain in part the differences in the simulated vocal fold posturing discussed in Sec. III A, and it is expected to also affect the three way coupling between tissue, airflow, and sound at the glottis. Furthermore, the two-dimensional (2D) scheme neglects the changes in the medial surface shape and the vertical height of the superior edge of the vocal folds (Movahhedi et al., 2021), and underestimates the anatomical constraints with extrinsic structures which can play a role during phonation (Moisik and Gick, 2017).

Two important aspects regarding the biomechanical modeling of the laryngeal tissues for the modified Kelvin model are highlighted. First, model parameters selection is key for producing reliable responses. Although a number of references report biomechanical parameters for the main laryngeal tissues in humans (Hunter and Titze, 2007; Hunter et al., 2004; Palaparthi et al., 2019; Titze, 2006), the significant variability among the different reported values is striking. The lack of consistent human data required a compromise that considered laryngeal function in both animals and humans for the model. Thus, the ad hoc values reported in Table IV correspond to those tissue parameters that yield the best performance simultaneously for both the laryngeal posture model and the TBCM; notwithstanding, significant differences with prior works are obtained, as evidenced in the simulated muscle activation plots with vocal process distance and fundamental frequency. In addition, each laryngeal tissue was modeled as a single one-dimensional fiber. This is a limitation to accurately capture the distributed portions of the intrinsic muscles, especially for the thyrovocalis and thyromuscularis portions of the TA muscle. This could be extended by adding independent submodels representing the different muscle portions, at the

TABLE IV. Parameters for simulating laryngeal tissues according to the modified Kelvin model. Passive stress-strain response and active stress for the five intrinsic laryngeal muscles: CT, cricothyroid; TA, thyroarytenoid; LCA, lateral cricoarytenoid; IA, interarytenoid; and PCA, posterior cricoarytenoid-, the vocal ligament (LIG) and mucosa (MUC) are considered.

Parameter	Laryngeal muscles and tissues						
	CT	LCA	TA	IA	PCA	LIG	MUC
σ_0 [kPa]	2.2	3.0	2.0	2.0	5.0	2.0	1.0
σ_2 [kPa]	5.0	59.0	1.5	30.0	55.0	1.4	20.0
B [−]	7.0	4.0	6.5	3.5	5.3	13.0	4.4
ϵ_1 [−]	−0.5	−0.5	−0.5	−0.5	−0.5	−0.5	−0.5
ϵ_2 [−]	−0.06	−0.06	−0.05	−0.06	−0.05	−0.3	−0.3
σ_m [kPa]	300	100	150	100	100	—	—
ϵ_m [−]	0.0	0.4	0.2	0.4	0.4	—	—

28   J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*

expense of increasing the model complexity. We opted to keep the model with the single one-dimensional fiber assumption for simplicity.

Note that the proposed scheme for the TBCM model continues to be an approximation and has limitations that could be addressed in future studies, such as constraints for cartilage displacements (Geng *et al.*, 2020; Hunter *et al.*, 2004), superior-inferior accommodation of the larynx during phonation (Moisik and Gick, 2017), medial bulging, and anterior-posterior gradient, among others.

Further assessment using heuristic or random control rules could provide additional insights into the muscle control of the laryngeal function to disentangle the relations between laryngeal function and the resulting acoustic output. Other future avenues for exploration include the connection with a neurophysiological muscle activation that incorporates natural neurological fluctuations in the activation of intrinsic laryngeal muscles (Manriquez *et al.*, 2019), and simulations of /VCV/ gestures to compute relative fundamental frequency.

V. CONCLUSIONS

A physiological scheme for controlling the mechanical properties of a triangular body cover model of the vocal folds through the independent activation of the five intrinsic laryngeal muscles is proposed. The approach builds upon prior efforts that describe rules for controlling low-order models, vocal fold posturing, and a triangular vocal fold model. The scheme provides a flexible and physiologically relevant way to control the self-sustained fully interactive voice production model for both sustained vowels and time-varying glottal gestures. At the same time, the resulting model allows for exploring the role of antagonistic muscle pairs in phonation. The model simulations are in agreement with prior studies using excised larynx experiments and high-order simulations. Using the proposed approach, we illustrate that different states of activation can lead to the same vocal fold posturing, albeit with highly different stress states. These similar posturing scenarios, however, have clear differences in the resulting kinematic, aerodynamic, and acoustic measures of vocal function. The resulting model is a relevant tool that can provide key insights into the physical mechanisms underlying normal and disordered phonation.

ACKNOWLEDGMENTS

APPENDIX: MODIFIED KELVIN MODEL

The modified Kelvin model applied for simulating the laryngeal tissue is briefly described. It is a one-dimensional, biomechanical model for the internal stress-strain response in fibrous tissue given by Eq. (A1). The main simulated variables are the axial stress, σ_i, and the axial strain, ϵ_i, for each tissue $i \in \mathcal{I}$. Equation (A1a) represents the total (passive plus active) stress, whereas Equation (A1b) describes active stress due to the internal contractile properties in the intrinsic muscles:

$$t_s \dot{\sigma}_i + \sigma_i = [\sigma_{ai} + \sigma_{pi} + Et_p \dot{\epsilon}_i], \tag{A1a}$$

$$t_a \dot{\sigma}_{ai} + \sigma_{ai} = a_i \sigma_m \max\left\{0,\, 1 - b(\epsilon_i - \epsilon_m)^2\right\}, \tag{A1b}$$

where a dot over a variable indicates time derivative, σ_{ai} is the active stress, and σ_{pi} is the passive viscoelastic stress corresponding to the fiber deformation ϵ_i modeled as follows (Hunter *et al.*, 2004; Titze, 2006):

$$
\begin{aligned}
\sigma_{pi} &= -\frac{\sigma_0}{\epsilon_1}(\epsilon_i - \epsilon_1), \quad \epsilon_i \leq \epsilon_2, \\
&= -\frac{\sigma_0}{\epsilon_1}(\epsilon_i - \epsilon_1) + \sigma_2\left[e^{B(\epsilon_i - \epsilon_2)} - 1 - B(\epsilon_i - \epsilon_2)\right], \\
&\quad \epsilon_i > \epsilon_2.
\end{aligned}
\tag{A2}
$$

For non-contractile tissue (i.e., vocal ligament and mucosa) the active component is set to zero.

Tissue-specific dynamical properties and stress-strain response characteristics in the Kelvin model of Eq. (A1) can be specified independently by its parameters (Hunter *et al.*, 2004; Titze, 2006): t_s is a time-series constant, t_p is a parallel time constant, t_a is the internal activation time constant, A_i is the cross-sectional area, σ_m is the maximum isometric active stress, ϵ_m is the strain at maximum contractile stress, b is a coefficient for active stress, σ_0 is the stress at zero strain, σ_2 is a scale factor for the exponential function, B is an exponential strain constant, ϵ_1 is the strain at zero stress, and ϵ_2 is the strain at which the nonlinear exponential function begins. Moreover, $E = d\sigma_{pi}/d\epsilon_i$ is the nonlinear tangent Young's modulus. The resulting fiber force magnitude is $F_i = A_i \sigma_i$, where the force direction coincides with the longitudinal axis in the fiber. The parameters set in this work for modeling laryngeal tissue are reported in Table IV, whereas the remaining parameters are set as originally introduced in Titze (2006).

Alzamendi, G. A., Manríquez, R., Hadwin, P. J., Deng, J. J., Peterson, S. D., Erath, B. D., Mehta, D. D., Hillman, R. E., and Zañartu, M. (**2020**). "Bayesian estimation of vocal function measures using laryngeal high-speed videoendoscopy and glottal airflow estimates: An *in vivo* case study," J. Acoust. Soc. Am. **147**(5), EL434–EL439.

Birkholz, P., Kröger, B. J., and Neuschaefer-Rube, C. (**2011a**). "Articulatory synthesis of words in six voice qualities using a modified two-mass model of the vocal folds," in *First International Workshop on Performative Speech and Singing Synthesis*, p3s-2011, March 14–15, Vancouver, Canada.

Birkholz, P., Kröger, B. J., and Neuschaefer-Rube, C. (**2011b**). "Synthesis of breathy, normal, and pressed phonation using a two-mass model with a

J. Acoust. Soc. Am. **151** (1), January 2022

Alzamendi *et al.*   29

triangular glottis," in *Proceedings of Interspeech 2011: 12th Annual Conference of the International Speech Communication Association*, August 27–31, Florence, Italy, pp. 2681–2684.

Chhetri, D. K., Neubauer, J., and Berry, D. A. (**2012**). "Neuromuscular control of fundamental frequency and glottal posture at phonation onset," J. Acoust. Soc. Am. **131**(2), 1401–1412.

Chhetri, D. K., Neubauer, J., Sofer, E., and Berry, D. A. (**2014**). "Influence and interactions of laryngeal adductors and cricothyroid muscles on fundamental frequency and glottal posture control," J. Acoust. Soc. Am. **135**(4), 2052–2064.

Cortés, J. P., Espinoza, V. M., Ghassemi, M., Mehta, D. D., Van Stan, J. H., Hillman, R. E., Guttag, J. V., and Zañartu, M. (**2018**). "Ambulatory assessment of phonotraumatic vocal hyperfunction using glottal airflow measures estimated from neck-surface acceleration," PLoS One **13**(12), e0209017.

Erath, B. D., Zañartu, M., Stewart, K. C., Plesniak, M. W., Sommer, D. E., and Peterson, S. D. (**2013**). "A review of lumped-element models of voiced speech," Speech Commun. **55**(5), 667–690.

Espinoza, V. M., Mehta, D. D., Stan, J. H. V., Hillman, R. E., and Zañartu, M. (**2020**). "Glottal aerodynamics estimated from neck-surface vibration in women with phonotraumatic and nonphonotraumatic vocal hyperfunction," J. Speech Lang. Hear. Res. **63**(9), 2861–2869.

Espinoza, V. M., Zañartu, M., Stan, J. H. V., Mehta, D. D., and Hillman, R. E. (**2017**). "Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction," J. Speech Lang. Hear. Res. **60**(8), 2159–2169.

Farley, G. R. (**1996**). "A biomechanical laryngeal model of voice F_0 and glottal width control," J. Acoust. Soc. Am. **100**(6), 3794–3812.

Galindo, G. E., Peterson, S. D., Erath, B. D., Castro, C., Hillman, R. E., and Zañartu, M. (**2017**). "Modeling the pathophysiology of phonotraumatic vocal hyperfunction with a triangular glottal model of the vocal folds," J. Speech Lang. Hear. Res. **60**(9), 2452–2471.

Geng, B., Pham, N., Xue, Q., and Zheng, X. (**2020**). "A three-dimensional vocal fold posturing model based on muscle mechanics and magnetic resonance imaging of a canine larynx," J. Acoust. Soc. Am. **147**(4), 2597–2608.

Gömmel, A., Butenweg, C., Bolender, K., and Grunendahl, A. (**2007**). "A muscle controlled finite-element model of laryngeal abduction and adduction," Comput. Methods Biomech. Biomed. Eng. **10**(5), 377–388.

Hillman, R. E., Stepp, C. E., Van Stan, J. H., Zañartu, M., and Mehta, D. D. (**2020**). "An updated theoretical framework for vocal hyperfunction," Am. J. Speech Lang. Pathol. **29**(4), 2254–2260.

Huber, J. E., Stathopoulos, E. T., and Sussman, J. E. (**2004**). "The control of aerodynamics, acoustics, and perceptual characteristics during speech production," J. Acoust. Soc. Am. **116**(4), 2345–2353.

Hunter, E. J., and Titze, I. R. (**2007**). "Refinements in modeling the passive properties of laryngeal soft tissue," J. Appl. Physiol. **103**(1), 206–219.

Hunter, E. J., Titze, I. R., and Alipour, F. (**2004**). "A three-dimensional model of vocal fold abduction/adduction," J. Acoust. Soc. Am. **115**(4), 1747–1759.

Ibarra, E. J., Parra, J. A., Alzamendi, G. A., Cortés, J. P., Espinoza, V. M., Mehta, D. D., Hillman, R. E., and Zañartu, M. (**2021**). "Estimation of subglottal pressure, vocal fold collision pressure, and intrinsic laryngeal muscle activation from neck-surface vibration using a neural network framework and a voice production model," Front. Physiol. **12**, 1419.

Lester, R. A., Daliri, A., Enos, N., Abur, D., Lupiani, A. A., Letcher, S., and Stepp, C. E. (**2020**). "The relation of articulatory and vocal auditory & motor control in typical speakers," J. Speech Lang. Hear. Res. **63**(11), 3628–3642.

Lowell, S. Y., and Story, B. H. (**2006**). "Simulated effects of cricothyroid and thyroarytenoid muscle activation on adult-male vocal fold vibration," J. Acoust. Soc. Am. **120**(1), 386–397.

Lucero, J. C., and Schoentgen, J. (**2015**). "Smoothness of an equation for the glottal flow rate versus the glottal area," J. Acoust. Soc. Am. **137**(5), 2970–2973.

Manriquez, R., Peterson, S. D., Prado, P., Orio, P., Galindo, G. E., and Zanartu, M. (**2019**). "Neurophysiological Muscle Activation Scheme for Controlling Vocal Fold Models," IEEE Trans. Neural Syst. Rehabil. Eng. **27**(5), 1043–1052.

Mehta, D. D., Espinoza, V. M., Van Stan, J. H., Zañartu, M., and Hillman, R. E. (**2019**). "The difference between first and second harmonic amplitudes correlates between glottal airflow and neck-surface accelerometer signals during phonation," J. Acoust. Soc. Am. **145**(5), EL386–EL392.

Mehta, D. D., Van Stan, J. H., Zañartu, M., Ghassemi, M., Guttag, J. V., Espinoza, V. M., Cortés, J. P., Cheyne, H. A., II, and Hillman, R. E. (**2015**). "Using ambulatory voice monitoring to investigate common voice disorders: Research update," Front. Bioeng. Biotechnol. **3**, 155.

Moisik, S. R., and Gick, B. (**2017**). "The quantal larynx: The stable regions of laryngeal biomechanics and implications for speech production," J. Speech Lang. Hear. Res. **60**(3), 540–560.

Movahhedi, M., Geng, B., Xue, Q., and Zheng, X. (**2021**). "Effects of cricothyroid and thyroarytenoid interaction on voice control: Muscle activity, vocal fold biomechanics, flow, and acoustics," J. Acoust. Soc. Am. **150**(1), 29–42.

Palaparthi, A., Smith, S., and Titze, I. R. (**2019**). "Mapping thyroarytenoid and cricothyroid activations to postural and acoustic features in a fiber-gel model of the vocal folds," Appl. Sci. **9**(21), 4671.

Poletto, C. J., Verdun, L. P., Strominger, R., and Ludlow, C. L. (**2004**). "Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures," J. Appl. Physiol. **97**(3), 858–866.

Story, B. H. (**2008**). "Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002," J. Acoust. Soc. Am. **123**(1), 327–335.

Story, B. H., and Titze, I. R. (**1995**). "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**(2), 1249–1260.

Titze, I. R. (**2002**). "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," J. Acoust. Soc. Am. **111**(1), 367–376.

Titze, I. R. (**2006**). *The Myoelastic Aerodynamic Theory of Phonation*, 1st ed. (National Center for Voice and Speech, Salt Lake City, UT).

Titze, I. R., and Hunter, E. J. (**2007**). "A two-dimensional biomechanical model of vocal fold posturing," J. Acoust. Soc. Am. **121**(4), 2254–2260.

Titze, I. R., and Palaparthi, A. (**2016**). "Sensitivity of Source-Filter Interaction to Specific Vocal Tract Shapes," IEEE/ACM Trans. Audio Speech Lang. Process. **24**(12), 2507–2515.

Titze, I. R., Story, B., Smith, M., and Long, R. (**2002**). "A reflex resonance model of vocal vibrato," J. Acoust. Soc. Am. **111**(5), 2272–2282.

Titze, I. R., and Story, B. H. (**2002**). "Rules for controlling low-dimensional vocal fold models with muscle activation," J. Acoust Soc. Am. **112**, 1064–1076.

Van Stan, J. H., Ortiz, A. J., Cortés, J. P., Marks, K. L., Toles, L. E., Mehta, D. D., Burns, J. A., Hron, T., Stadelman-Cohen, T., Krusemark, C., Muise, J., Fox-Galalis, A. B., Nudelman, C., Zeitels, S., and Hillman, R. E. (**2021**). "Differences in daily voice use measures between female patients with nonphonotraumatic vocal hyperfunction and matched controls," J. Speech Lang. Hear. Res. **64**(5), 1457–1470.

Yin, J., and Zhang, Z. (**2014**). "Interaction between the thyroarytenoid and lateral cricoarytenoid muscles in the control of vocal fold adduction and eigenfrequencies," J. Biomech. Eng. **136**, 1110061–11100610.

Zañartu, M. (**2006**). "Influence of acoustic loading on the flow-induced oscillations of single mass models of the human larynx," Master's thesis, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.

Zañartu, M., Galindo, G. E., Erath, B. D., Peterson, S. D., Wodicka, G. R., and Hillman, R. E. (**2014**). "Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction," J. Acoust. Soc. Am. **136**(6), 3262–3271.

Zhang, Z. (**2016**). "Mechanics of human voice production and control," J. Acoust. Soc. Am. **140**(4), 2614–2635.

Alzamendi *et al.*