Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction[a)]

Matías Zañartu[b)] and Gabriel E. Galindo
Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile

Byron D. Erath
Department of Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, New York 13699

Sean D. Peterson
Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

George R. Wodicka[c)]
Weldon School of Biomedical Engineering, Purdue University, West Lafayette, Indiana 47907

Robert E. Hillman[d)]
Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, Massachusetts 02114

(Received 4 June 2014; revised 16 October 2014; accepted 3 November 2014)

Despite the frequent observation of a persistent opening in the posterior cartilaginous glottis in normal and pathological phonation, its influence on the self-sustained oscillations of the vocal folds is not well understood. The effects of a posterior gap on the vocal fold tissue dynamics and resulting acoustics were numerically investigated using a specially designed flow solver and a reduced-order model of human phonation. The inclusion of posterior gap areas of 0.03–0.1 cm^2 reduced the energy transfer from the fluid to the vocal folds by more than 42%–80% and the radiated sound pressure level by 6–14 dB, respectively. The model was used to simulate vocal hyperfucntion, i.e., patterns of vocal misuse/abuse associated with many of the most common voice disorders. In this first approximation, vocal hyperfunction was modeled by introducing a compensatory increase in lung air pressure to regain the vocal loudness level that was produced prior to introducing a large glottal gap. This resulted in a significant increase in maximum flow declination rate and amplitude of unsteady flow, thereby mimicking clinical studies. The amplitude of unsteady flow was found to be linearly correlated with collision forces, thus being an indicative measure of vocal hyperfunction. [http://dx.doi.org/10.1121/1.4901714]

PACS number(s): 43.70.Bk, 43.70.Aj, 43.70.Dn [BHS] Pages: 3262–3271

I. INTRODUCTION

A. Posterior glottal opening

Incomplete glottal closure is known to be ubiquitous in both normal and disordered voices. The lack of complete closure during the closed phase of the vocal fold cycle results in airflow leaking through the glottis into the vocal tract, thus changing the aerodynamic, acoustic, and subsequent dynamic behavior of the larynx. Incomplete glottal closure can occur between the vibrating segments of the vocal folds (membranous glottis) and/or between the nonvibrating arytenoid cartilages (cartilaginous glottis), which is referred to as a posterior glottal opening (PGO). In normal voices, a membranous leak can be the result of an abducted vocal fold process configuration, which results in a triangular glottal shape. This configuration is common in female speakers,[1–3] children,[4] and voice pathologies that alter prephonatory adduction, such as vocal fold paralysis, abductor spasmodic dysphonia, and muscle tension dysphonia. A membranous leak can also occur due to organic vocal fold lesions (nodules, polyps) and superior laryngeal nerve injuries. On the other hand, incomplete glottal closure due to a PGO results from vocal posturing by creating a gap confined to the posterior cartilaginous glottis. The PGO cannot always be observed via laryngeal endoscopy because it is sometimes hidden by the arytenoid cartilages, but it is consistently revealed by a steady flow component (also referred to as DC or minimum flow).[5–7] The PGO is present in normal modal phonation and becomes more apparent in higher frequency registers and pathological conditions.

The aerodynamic differences between leakage through the membranous portion of the vocal folds and leakage through a PGO were studied using a driven parametric model by Cranen *et al.*[8,9] As noted before, it was found that both types of leaks increased DC flow and source-filter interaction,[10,11] and introduced fluctuations during the closed phase of the cycle.[11] However, a leak in the membranous

[a)]Part of this work was presented at the 165th Meeting of the Acoustical Society of America in Montreal, Quebec, Canada.
[b)]Author to whom correspondence should be addressed. Electronic mail: matias.zanartu@usm.cl
[c)]Also at School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907.
[d)]Also at Harvard Medical School and Harvard-MIT Division of Health Sciences and Technology, Boston, MA 02114.

portion gave rise to a steeper decay of the glottal flow spectrum, whereas a PGO decreased the low-frequency content and maintained the higher frequency content with respect to the fully-closed scenario.[8,9] In spite of other efforts devoted to understand the acoustic and aerodynamic effects of incomplete glottal closure,[2,3,12] the overall role of glottal leakage on the net energy transfer during phonation remains unclear. Progress in that direction was made by Park and Mongeau[13] in their experimental study of the influence of a PGO on the net energy transfer using a flow-driven, physiological-scale synthetic vocal fold model with no vocal tract. They found that a PGO of 0.08 cm^2 reduced the net energy transfer to roughly 10%. However, the lack of a vocal tract eliminated the presence of source-filter interactions that have been noted to be of importance in this phenomenon.[10,11]

B. Numerical modeling

Even though self-sustained models of the vocal folds are designed to provide insights into the mechanisms that control phonation in normal and pathological cases,[14] limited efforts have been made to include the effects of incomplete glottal closure in these representations. It is noted that the aforementioned work of Cranen *et al.*[8,9,11] imposed a glottal area waveform that was obtained from a parametric model[8,9] or a two-mass model,[11] but did not use three way interactions between airflow, tissue, and sound, as is commonly employed in self-sustained models. Thus, these efforts were unable to explore the influence of the PGO on the vocal fold dynamics. A few other studies that have incorporated some type of incomplete glottal closure in low-dimensional, self-sustained models, have included pre-contact changes in the stiffness,[15] an anterior-posterior feature that restricted the vibration of the vocal folds,[16] a nonlinear damping coefficient,[17] and a gradual anterior-posterior closure due to a triangular glottis.[18] These distinct approaches have been used to mimic polyps and nodules,[15,19] to minimize the unrealistically large amplitudes produced when increasing the degree of abduction to reproduce consonant-vowel-consonant gestures,[16,20] and to synthesize different voice qualities.[21] More recently, a self-sustained model with PGO was proposed,[22] to predict AC and DC flows, as compared with inverse filtered glottal airflow from human recordings. However, this work did not include an acoustic propagation model and hence did not provide insights into the influence of the PGO on the self-sustained oscillations of the vocal folds in a fully fluid-structure-acoustic scenario.

Despite the aforementioned efforts, incomplete glottal closure, and particularly PGO, remains largely neglected in studies using self-sustained numerical models. In addition, inclusion of the changes in the net energy transfer from flow and sound to the vibrating tissue introduced due to incomplete closure has not been studied with a fully coupled model that accounts for flow-sound-tissue interactions. These acoustic interactions have been noted to be of importance in this phenomenon[10,11] and must be considered. Investigating the effects of a PGO on the vocal fold dynamics in this fashion will help to understand the phenomena affecting the driving forces of vocal fold oscillation. As noted by Park and Mongeau,[13] a PGO can reduce the net energy transfer significantly. This reduction in energy may come with a reduction in loudness, which might spur compensations from the subject. These attempts to compensate may contribute to vocal hyperfunction, which can lead to the development of a voice disorder. An abnormally large PGO is probably caused by the hyperfunction in the first place.

C. Vocal hyperfunction

Vocal hyperfunction is associated with a majority of the most common voice disorders. At a basic level it refers to patterns of vocal misuse/abuse caused by excessive or imbalanced forces in the muscles involved in phonation.[6] In a common scenario, attempts to compensate for a deficit in voice production (e.g., reduced loudness) in this fashion could contribute to a vicious cycle in which increased muscular and aerodynamic forces cause further deterioration in vocal function (e.g., vocal fatigue, vocal fold tissue damage, dysphonia, etc.) and the need to drive the system at ever increasing levels.[6] Two types of vocal hyperfunction that can be quantitatively described and differentiated from each other and normal voice production have been proposed; adducted hyperfunction and non-adducted hyperfunction.[6]

Adducted hyperfunction is associated with the formation of benign vocal fold lesions, such as nodules and polyps, and is accompanied by abnormalities in various vocal function measures. This behavior is produced by stiff and tightly approximated vocal folds, and accompanied by increased lung pressure. All these conditions are expected to produce abnormally high vocal fold collision forces that lead to tissue trauma and the formation of benign vocal fold lesions. Since measuring collision forces during *in vivo* recordings is not trivial, these forces are indirectly estimated via aerodynamic measures, such as maximum flow declination rate (MFDR) and the amplitude of unsteady flow (AC flow), both being related to vocal fold closing velocity and amplitude. Despite this intuitive rationale, there is limited evidence that relates these measures to the actual collision forces.

Non-adducted hyperfunction is associated with vocal fatigue and dysphonia, but with an absence of vocal fold tissue trauma. Other clinical terms for non-adducted hyperfunction encountered in the literature include functional dysphonia and muscle tension dysphonia. Non-adducted hyperfunction is produced when high levels of stiffness and tension in the vocal folds are observed, but the folds are not completely approximated. Muscle fatigue may occur with voice use, but there is no vocal fold trauma, and therefore, no development of secondary organic pathology. The most consistent finding for this type of hyperfunction is a significant increase in the minimum flow (or DC flow), which reflects the reduction in glottal closure.

Using a Rothenberg mask,[23] salient measures of adducted and non-adducted hyperfunction have been detected in patients by relating aerodynamic and acoustic parameters from recordings of a single sustained vowel.[6] Detection of abnormal behavior was performed by contrasting data from pathological subjects to that of a normative population, using normal and regressed Z-scores.[6] Z-scores are measures of the number of standard deviations a given data point deviates from the mean of a normative set.

Methods to detect hyperfunctional voice disorders based on long-term recordings are also possible.[24–26]

D. Aims of the study

In this study, the effects of a PGO on tissue dynamics, energy transfer, acoustic interactions, and glottal aerodynamics were numerically investigated. We first introduce a simple tool that facilitates the inclusion of a PGO in self-sustained models of phonation. With this tool, we explore the changes in the fluid-structure-acoustic energy exchange due to the presence of a PGO and evaluate the effect on the resulting loudness. In an effort to relate the numerical model to actual clinical data, this study aims to contrast the acoustic and aerodynamic effects of a PGO with actual human recordings to gain further insights into its potential role in vocal hyperfunction.

The manuscript outline is as follows: Details of the reduced-order vocal fold model, selected measures, and analysis tools are presented in Sec. II, including a derivation of the PGO flow solver; numerical simulations and their contrast with clinical data are presented and discussed in Sec. III; and conclusions are put forth in Sec. IV.

II. METHODS

A. Numerical model selection

Self-sustained models of the vocal folds are designed to provide insights into the mechanisms that control phonation in normal and pathological cases. Low-dimensional models are more commonly used, as they efficiently capture the most dominant modes of vibration and are expected to reproduce many fundamental aspects of phonation with acceptable accuracy at lower computational cost. The three-mass body-cover model,[27] one of the most accepted low-dimensional models, was used to evaluate the proposed PGO scheme. A schematic of the model representation, including the addition of the PGO, is shown in Fig. 1. This model is an extension of the classical two-mass model[28] that better represents physiological aspects of the vocal folds, and has been used to study source-filter interaction,[29,30] voice pathologies,[19,31] inverse filtering,[32] and muscle activation,[33] among others. The equations of motion for the vocal fold model are taken exactly as in the original paper[27] and are thus omitted for brevity.

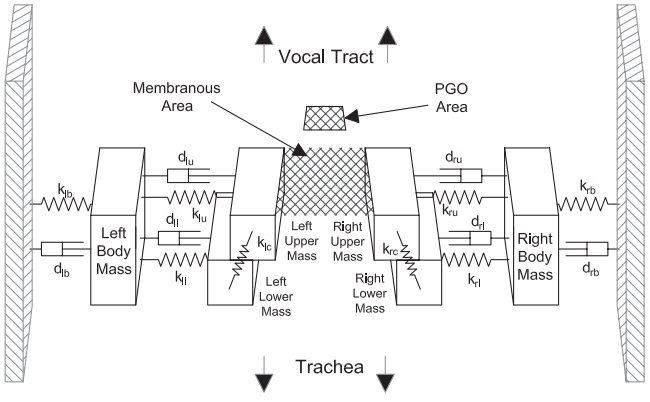

FIG. 1. Three-dimensional representation of the body cover model showing the posterior glottal opening, the vocal fold masses (body and cover), and the membranous area.

Model parameters were selected to yield a male modal voice using muscle activation principles,[34] thus selecting a 10% cricothyroid and 20% thyroarytenoid muscle activation. A wave-reflection analog scheme[35] was used to account for sound propagation and interaction, based on a sustained vowel /e/.[36] The subglottal area function was adapted from respiratory system measurements of human cadavers[37] and includes the trachea, bronchi, and a resistive termination impedance (zeroth and first airway generations). Unless otherwise stated, the simulation parameter of the model were set as the following: Speed of sound = 350 m/s; simulation time = 200 ms; sampling frequency = 70 kHz; and subglottal pressure = 800 Pa.

Although it is well-established that the presence of incomplete glottal closure is associated with an increase in turbulent noise,[3] and that the cepstral peak prominence (CPP) is highly correlated to the presence of turbulence in voice,[38] turbulent noise was not included in these initial investigations due to the apparent small influence of turbulence on vocal fold dynamics.[39] Thus, the present study focused only on the main dipole component of voice production,[40] leaving for future investigations the addition of turbulent sound sources due to the PGO.

The effects of the posterior gap were investigated through parametrical variations of the PGO area. To isolate the effect of the posterior gap, no other mechanisms of incomplete closure were included. Slightly larger than normal posterior gaps were used as a first approximation of pathological conditions in which gaps could extend into the membranous glottis.

B. Airflow through the posterior glottal opening

The volumetric flow rate through the combined membranous portion of the vocal folds and the posterior glottal gap, subjected to acoustic driving pressures, is derived using the control volume depicted in Fig. 2. The geometry shown in Fig. 2 is a two-dimensional schematic representation of the more physiological geometry shown in Fig. 1, and is meant for illustration of the control volume only. The inferior surface of the control volume spans the exit of the glottis A_d, and the gap A_g, while the superior surface is placed sufficiently far downstream that the velocity of the fluid and the

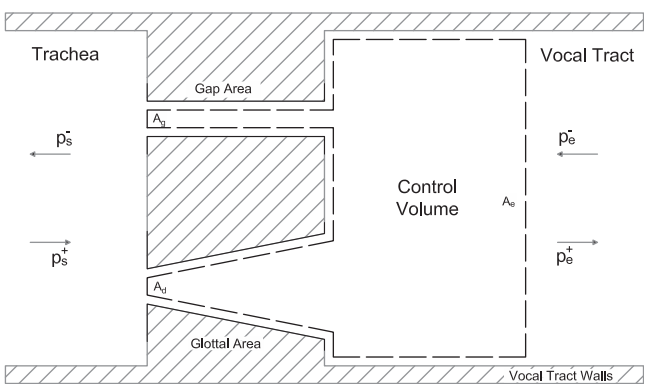

FIG. 2. Diagram representing the vocal fold area, the trachea area, and the posterior glottal opening used for calculation of the flow.

pressure field can once again be considered uniform. The area at this downstream location is A_e. Thus, the complex viscous effects associated with the interaction of the glottal jet with flow emanating from the gap is fully contained within the control volume.

The jet of air that passes through both orifices exhausts into the supraglottal tract, and it is assumed that the pressure at the exit plane is uniform. It is further assumed that the fluid is incompressible and inviscid, and thus the frictional losses leading into the glottis and PGO are negligible. Consequently, from Bernoulli's equation, the velocity at the exit of the gap V_g and the velocity at the exit of the glottis V_d are also equal. Using the control volume in Fig. 2, conservation of mass then yields the velocity at the exit of the vocal tract as $V_e = V_d((A_d + A_g)/A_e)$. Conservation of linear momentum for the control volume reduces to the sum of the forces acting on the volume being balanced by the momentum flux through the control volume boundaries. This assumes that a steady state condition applies. Conservation of linear momentum reduces to

$$(P_e - P_d) = \frac{1}{2}\rho\left(\frac{Q}{A_d + A_g}\right)^2 \times \left[2\left(\frac{A_d + A_g}{A_e}\right)\left(1 - \left(\frac{A_d + A_g}{A_e}\right)\right)\right], \quad (1)$$

where P_d is the pressure at the glottal exit plane (which is equal to the pressure at the gap exit), P_e is the pressure in the vocal tract, $Q = V_d(A_d + A_g)$ is the total volumetric flow rate of air through the system, and ρ is the fluid density. The pressure P_d can be expressed in terms of the sub glottal pressure P_s as

$$P_d = P_s - \frac{1}{2}\rho\left(\frac{Q}{A_d + A_g}\right)^2. \quad (2)$$

To enable comparison with similar approaches,[41] a kinetic loss coefficient k_e is defined as

$$k_e = 2\left(\frac{A_d + A_g}{A_e}\right)\left(1 - \left(\frac{A_d + A_g}{A_e}\right)\right), \quad (3)$$

allowing conservation of linear momentum to be expressed as

$$(P_s - P_e) = \frac{1}{2}\rho\left(\frac{Q}{A_d + A_g}\right)^2 (1 - k_e). \quad (4)$$

The subglottal pressure and the vocal tract pressure are equal to the sum of the forward and backward traveling acoustic waves in those regions that, following Liljencrants wave reflection analog,[42] can be expressed as

$$2P_s = 2p_s^+ - \rho\frac{cQ}{A_s}, \quad (5)$$

$$P_e = 2p_e^- + \rho\frac{cQ}{A_e}, \quad (6)$$

where c is the speed of sound of the air, and p_s^+ and p_e^- are the forward traveling and backward traveling acoustic waves in the subglottal and vocal tract regions, respectively. Substituting Eqs. (5) and (6) into Eq. (4) and rearranging yields a quadratic equation for the volumetric flow rate

$$Q^2 + Q\frac{2c(A_d + A_g)^2}{A^*(1 - k_e)} - \frac{4(p_s^+ - p_e^-)(A_d + A_g)^2}{\rho(1 - k_e)} = 0, \quad (7)$$

where $A^* = [A_s A_e/(A_s + A_e)]$ is the equivalent vocal tract area. The solutions to this equation are

$$Q = \frac{c(A_d + A_g)}{1 - k_e}\left[-\left(\frac{A_d + A_g}{A^*}\right) \pm \sqrt{\left(\frac{A_d + A_g}{A^*}\right)^2 + \frac{4(1 - k_e)}{\rho c^2}(p_s^+ - p_e^-)}\right], \quad (8)$$

which is identical to Titze's flow solver[41] with $k_t = 1 - k_e$ and the glottal area increased by the gap area.

The aerodynamic domain was prescribed as flow through two separate orifices (posterior gap and membranous vocal folds) that merge in the supraglottal tract, with the governing flow equations determined from a control volume analysis based on conservation of mass and linear momentum. The proposed method is based on a wave-reflection analog (WRA) scheme to account for the acoustic pressures in the system. Note that the proposed method does not alter the general structure of the governing dynamical equations for flow through only a membranous glottis due to the inviscid flow assumption through the glottis and the gap; the current approach is equivalent to solving for the glottal airflow using the total glottal area, i.e., summing both posterior gap and membranous glottal areas. This simple alteration for modeling flow through a PGO is readily applicable to any self-sustained vocal fold model that similarly employs one-dimensional, incompressible, and inviscid flow assumptions. Thus, source-filter interaction[43] is responsible for any changes in the net energy transfer and consequent vocal fold dynamics in this approach.

The proposed airflow solver can be used in conjunction with any voice production model that resolves the acoustic field using a WRA propagation scheme. It is important to include a fully interactive representation, where there is a three-way interaction between sound, flow, and vocal fold tissue, i.e., level 2 of interaction.[43] As discussed by Titze,[41,43] the source-filter interaction is primarily controlled by the coupling parameter $\zeta = a_g/A^*$, where a_g is the total glottal area and A^* is the equivalent vocal tract area as defined before, leading to Titze's asymptotic flow solutions for an uncoupled scenario given by

$$Q_{\text{uncoupled}} = a_g\sqrt{\frac{4}{k_t\rho}(p_s^+ - p_e^-)}, \quad (9)$$

and for a highly coupled scenario,

$$Q_{\text{coupled}} = \frac{2A^*}{\rho c}(p_s^+ - p_e^-). \quad (10)$$

C. Selected measures of vocal function

Numerous acoustic parameters have been proposed to detect the associated breathiness produced by excessive incomplete glottal closure, including cepstral peak prominence (CPP),[38] amplitude of the first harmonic relative to the second harmonic (H_1-H_2)[3], and harmonic richness factor (HRF).[44] Aerodynamic parameters such as, maximum flow declination rate, subglottal pressure (Ps), steady flow rate (DC flow), and unsteady flow rate (AC flow) allow for identifying hyperfunctional voices that are in many cases a product of incorrect compensations due to excessive incomplete glottal closure.[6]

To evaluate the effects of the posterior gap, selected parameters were computed, including fundamental frequency (F0), maximum flow declination rate, radiated sound pressure level, steady and unsteady glottal airflow components, spectral tilt, and net energy transfer (denoted as Π_m). Sound pressure level (SPL) is obtained at the lips and was projected to a 15 cm distance by subtracting 30 dB, based upon our empirical observations. The energy transfer was computed as in previous studies[13,45] and given by Eq. (11), where $p_i(t)$ is the driving pressure acting on mass i (upper or lower), T is the period, and $v_i(t)$ is the velocity of the ith mass,

$$\Pi_{m_i}(t) = p_i(t)v_i(t), \tag{11}$$

$$\Pi_{m_{i\,\mathrm{RMS}}} = \frac{1}{T}\sqrt{\int_0^T \Pi_{m_i}(t)^2 dt}. \tag{12}$$

D. Modeling hyperfunction

Vocal hyperfunction is associated with compensations that are believed to contribute to increased muscular and contact forces that can cause further deterioration in vocal function. Although there are different hyperfunctional mechanisms, in this study, vocal hyperfunction was associated with increased lung pressure to regain vocal loudness, assuming that this could only be accomplished by increasing the activity of the respiratory and laryngeal muscles involved in producing phonation. Thus, the compensation was represented as an increase in the subglottal pressure to match a given SPL target with a 0.1 dB of error. The results of the simulation were then compared with actual measurements performed on a normative population to identify abnormal behavior based on regressed Z-scores.[6] A Z-score measures the numbers of standard deviations (σ_N) between a given data point (x) and the mean of a given normative ($\bar{x}_N$) data set,

$$Z = \frac{x-\bar{x}_N}{\sigma_N}, \tag{13}$$

where a Z-score with a magnitude higher than 2 (>2 standard deviations) is typically considered indicative of abnormality.[6]

A normative set was taken from the data reported by Perkell *et al.*,[7] considering parameters such as MFDR, AC flow, and DC flow for two loudness condition (normal and increased). Because of the lack of intermediate values of loudness, the mean ($\bar{x}_N$) and standard deviation (σ_N) for each parameter were linearly projected between the two loudness conditions, producing a regressed Z-score given by the distance to the regression line, normalized by the standard deviation. While variance was projected linearly between the two loudness conditions, outside this range the variances were maintained on the values indicated by Perkell *et al.*,[7] avoiding irregularities due to a zero variance projection.

III. RESULTS

A. Resulting waveforms

The predicted effects of a PGO on the resulting glottal airflow are depicted in Fig. 3, where a posterior glottal opening of 0.05 cm^2 was considered. It is noted that the addition of the gap yields smooth contours of the glottal airflow during the closed portion of the cycle, which is in agreement with experimental observations from inverse filtering in human subjects.[8,32] However, it is noted that this smoothing of the glottal airflow may occur from multiple other conditions and is not uniquely tied to the PGO effects.

The resulting airflow is contrasted with two other flow solvers that account for a PGO in an idealized fashion: An uncoupled Bernoulli flow solver that is proportional to the glottal area; and a simplified flow solver for highly coupled scenarios that is proportional to the incident transglottal pressure.[41] These idealized cases are asymptotic conditions that are representative of minimum and maximum flow-sound coupling[41] and are described by Eqs. (9) and (10), respectively. It is noted that the resulting glottal airflow (solid line in Fig. 3) better matches the highly coupled solver (dotted line in Fig. 3), suggesting that the posterior glottal gap significantly increases source-filter interactions, which becomes more evident for large gap areas. It is also noted that for large gap areas, the transglottal pressure term becomes smoother and relatively proportional to the glottal area, further reducing the harmonic content of the source spectra. These findings are explained by the coupling parameter[43] given by $\zeta = a_g/A^*$. Note that the PGO area adds to the total glottal area, such that $a_g = A_d + A_g$, which leads to a higher coupling parameter ζ and a highly coupled scenario.

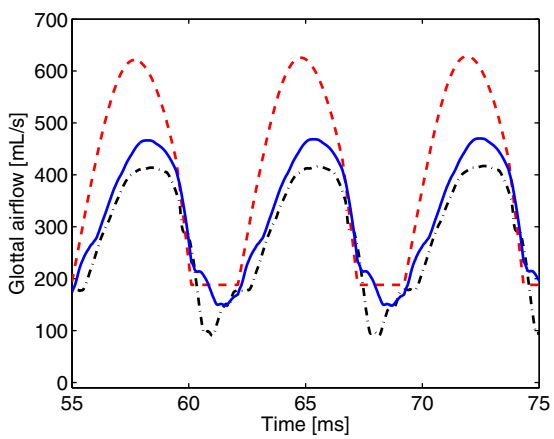

FIG. 3. (Color online) Resulting glottal airflow with proposed PGO method (–), uncoupled Bernoulli flow solution (- -), and highly coupled airflow solution (-.-).

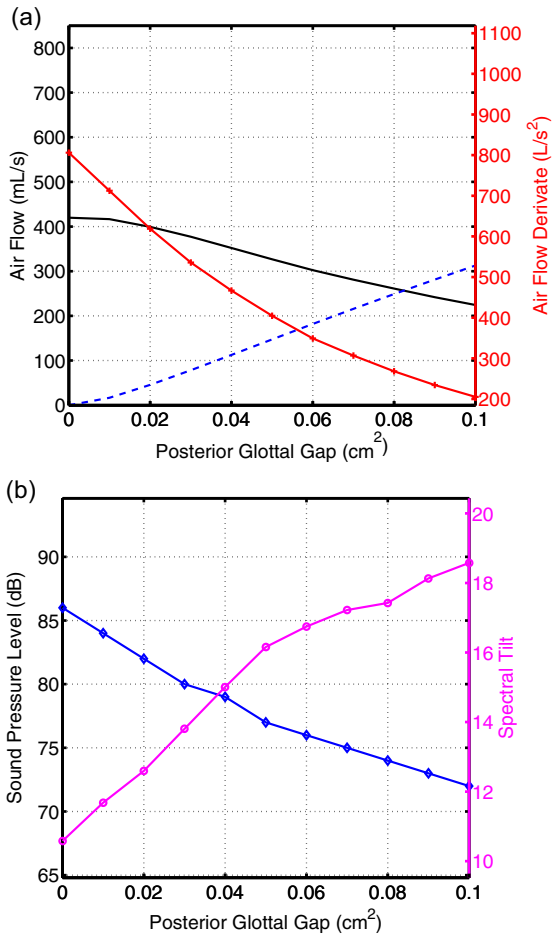

FIG. 4. (Color online) Effect of the posterior glottal gap on selected glottal measures: (a) AC flow (–, left axis), DC flow (- -, left axis), MFDR (-+-, right axis) (b) SPL (-◇-, left axis) and Spectral tilt (-○-, right axis).

This discussion follows the same principles that are used to claim that a narrow vocal tract can increase the source filter-interaction.[41,43] Note that this scenario has a smaller DC and AC flow components than the uncoupled Bernoulli flow.

B. Acoustic, aerodynamic, and energy transfer effects

The acoustic and aerodynamic effects of a parametric variation of the posterior gap area are presented in Fig. 4. As expected, the increment in gap area increases the DC component of the glottal airflow linearly. In contrast, a decay in AC flow and MFDR is observed with larger gap areas. The gap area has similar acoustic effects in the source spectra and radiated SPL.

The energy transfer from the flow to the vocal fold tissue is also affected by increments in gap area, as observed in Fig. 5. Larger gap areas reduce the net energy that drives the vocal fold vibration, where the superior cover mass (mass 2) suffers a more rapid decay due to the more predominant influence of the downstream pressure on the driving force. For this mass, an increase of 0.1 cm^2 resulted in a reduction in energy transfer from 1.22 J/s to 0.22 J/s, i.e., a decrease of 82%. A similar decrease in energy transfer (90%) was observed in previous studies for synthetic vocal fold models.[13] Note that due to the incorporation of a normal posterior gap opening (e.g., 0.03–0.05 cm^2), the body-cover model

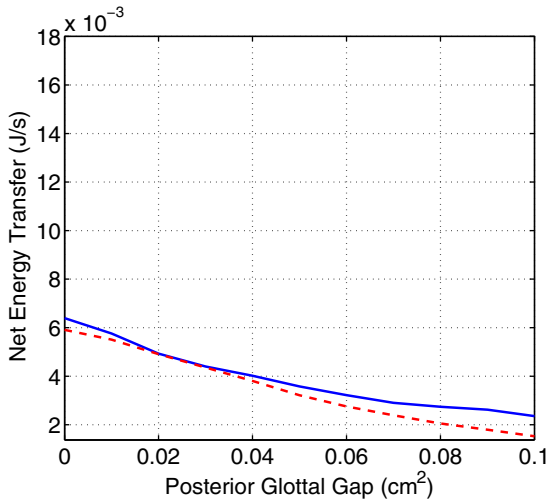

FIG. 5. (Color online) Net energy transfer vs PGO. Upper cover mass (–), lower cover mass (- -).

produced acoustic and aerodynamic outputs that were squarely in the range of normal human speech production,[7] which was not the case for the non-gap scenario. This is not a singularity of the particular condition or model that was used, as it was seen for other conditions and vocal fold models, and it is associated to the non-natural fully closed glottal condition most lumped models produce in chest register.

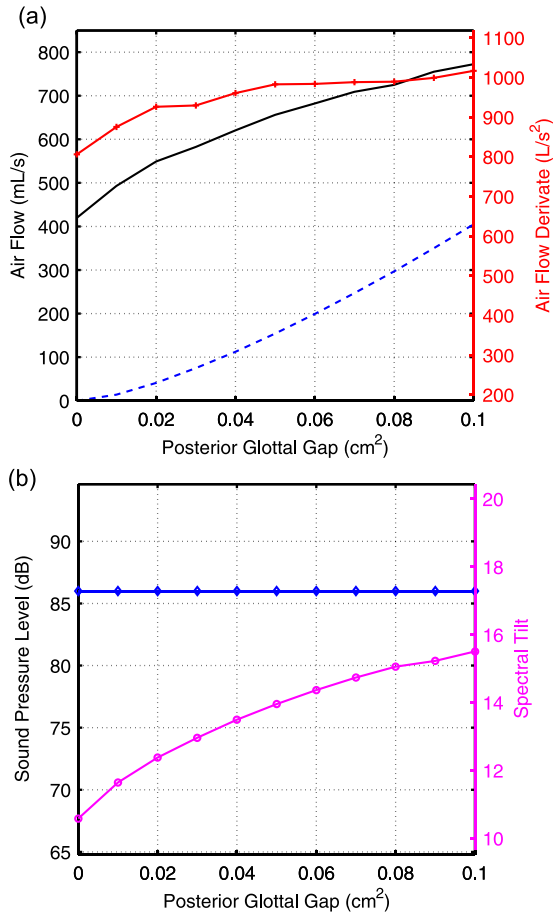

FIG. 6. (Color online) Effect of the posterior glottal opening on selected glottal measures with sub glottal pressure compensation: (a) AC flow (–, left axis), DC flow (- -, left axis), MFDR (-+-, right axis), (b) SPL (-◇-, left axis) and Spectral tilt (-○-, right axis).

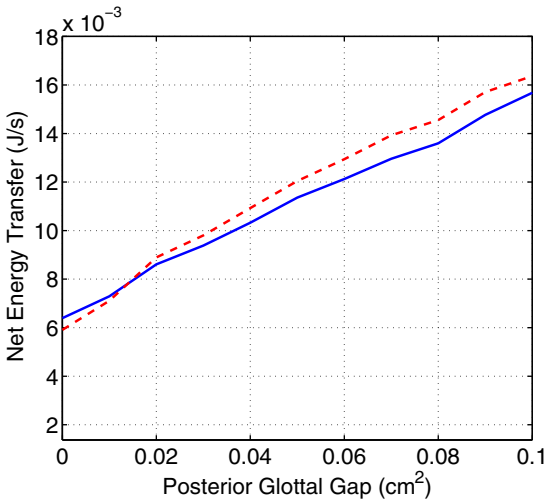

FIG. 7. (Color online) Net energy transfer vs PGO. Upper cover mass (–), lower cover mass (- -).

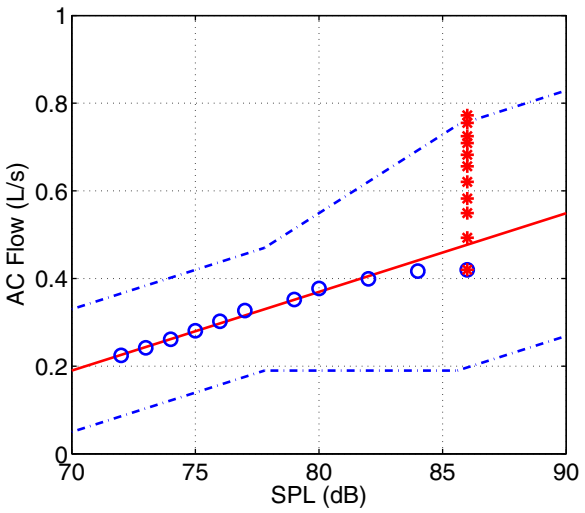

FIG. 8. (Color online) Illustration of the computation of regressed Z-scores for AC Flow. Mean values of AC Flow are linearly related to SPL (–), and ±2 standard deviations across SPL (- -). Simulations for various gap areas for the non-compensated scenario are shown in circles, and those for the compensated scenario in asterisks.

C. Vocal hyperfunction: Effects on selected measures

It is noted from the previous sections that PGO yields a decrease in the energy transfer from the airflow to the vocal fold tissue that in turn results in a reduced SPL. A compensation for the reduced SPL is introduced by increasing subglottal pressure up to the point where a target SPL is achieved. A target given by the no-gap scenario was selected, i.e., 86 dB SPL. This scenario is referred to as the "compensated" case, which results in higher aerodynamic measures and increased energy transfer, as shown in Fig. 6 and Fig. 7, respectively. These figures are the paired compensated scenarios previously presented in Fig. 4 and Fig. 5. Larger gap areas require a more significant compensation effort and higher values for all selected aerodynamic measures, e.g., the AC flow increases over 76%, DC flow increases almost linearly from 0 to 400 mL/s, and MFDR clearly indicates a more dynamic behavior in comparison with the non gap case and non-compensated case. Figure 6(b) illustrates that the target SPL is maintained for all gap areas and a slight increase in spectral tilt is observed, indicating that the proposed compensation does not make a considerable alteration in the resulting voice quality of the main dipole component.

The observed changes in aerodynamic measures due to the lung pressure compensation are produced by increased net energy transfer from airflow to tissue, as shown in Fig. 7. The magnitude of the net energy transfer (RMS) is a monotonically increasing function of the PGO and almost linearly related to the increasing lung pressure. This finding is explained by the construction of Eq. (11), which implies a direct increase in the net energy transfer due to pressure acting on the vocal folds. Since the sub glottal pressure is increased in order to obtain the targeted SPL, the net energy transfer is also driven to a higher value proportional to the increased pressure.

D. Vocal hyperfunction: Contrast with normative data

The effects of sub glottal pressure compensation can be assessed in terms of actual human data via Z-scores. The rationale to obtain the regressed Z-scores for one of the selected measures (AC Flow) is shown in Fig. 8. In this figure, the thresholds are fixed by 2 standard deviations (dotted lines) and are depicted along with the linear relation obtained between AC flow deviations and SPL. The mean values and the standard deviations are obtained from human data in Perkell *et al.*[7] The data points from the non-compensated simulations for various gap areas (circles) fall within the normal range, and the compensated ones for a given target (asterisks) fall, in some cases, outside of the normal range. This means that hyperfunctional subjects will be outside of a normal range for this measure while having a normal SPL. In other words, all subjects achieve the same normal SPL but those with hyperfunction exceed the normal ranges on other measures of vocal function due to the extra effort required to compensate for incomplete glottal closure.

Table I illustrates the idea of hyperfunctional compensations, where a scenario without a posterior gap is compared with cases with gap areas of 0.03 cm^2 (normal gap opening) and 0.1 cm^2 (excessive gap opening). As noted before, when compensating for the reduction in SPL in the latter case to match the no-gap scenario with an increased subglottal pressure, an increase in MFDR, AC flow, and DC flow is observed. This observation is in agreement with measurements on adducted hyperfunctional subjects.[6]

The same principle is used for the selected gap areas and conditions from Table I, yielding Z-scores for MFDR, AC flow, and DC flow. Those scores with magnitudes higher than 2 are considered abnormal; it is noteworthy that DC flow and AC flow are the only scores in that range. DC flow becomes abnormally high for a large gap area, with and without lung pressure compensations. On the other hand, AC flow exhibits abnormally high values only when compensated with increased lung pressure. It is observed that MFDR becomes much higher with compensations, but not sufficiently high to be considered pathological. All of these observations are in agreement with the findings of

TABLE I. Simulation of hyperfunction due to the compensation of the reduction in SPL by an increased subglottal pressure. The values in parentheses are percentage differences with respect to the no gap scenario, which was the target for SPL.

Parameter	Units	No gap	Normal	Large gap	Hyperfunction
A_{min}	cm^2	0	0.03	0.1	0.1
P_s	Pa	800	800 (0%)	800 (0%)	1581 (98%)
SPL	dB	86	80 (−50%)	72 (−80%)	86 (0%)
F_0	Hz	147.7	144.9 (−2%)	141.1 (−4%)	150.9 (2%)
MFDR	L/s^2	806.1	535.8 (−34%)	205.3 (−75%)	1016.7 (26%)
DC flow	mL/s	0	78	313	405
AC flow	mL/s	420	377 (−10%)	224 (−47%)	772 (84%)
$H_1 - H_2$	dB	12.4	15.5 (25%)	20.6 (66%)	17.3 (39%)
$\Pi_{RMS_{m1}}$	J/s	1.19	0.84 (−29%)	0.52 (−57%)	2.74 (131%)
$\Pi_{RMS_{m2}}$	J/s	1.22	0.71 (−42%)	0.22 (−82%)	2.57 (110%)
Z - MFDR	–	0.56	0.68	0.79	1.40
Z - AC flow	–	−0.41	0.09	−0.02	2.11
Z - DC flow	–	−1.6	−0.03	4.66	6.49

previous studies,[6] where DC flow is always independent of the type of hyperfunction, AC flow becomes much higher when compensated with increased lung pressure, and MFDR follows a similar trend with a less salient behavior.

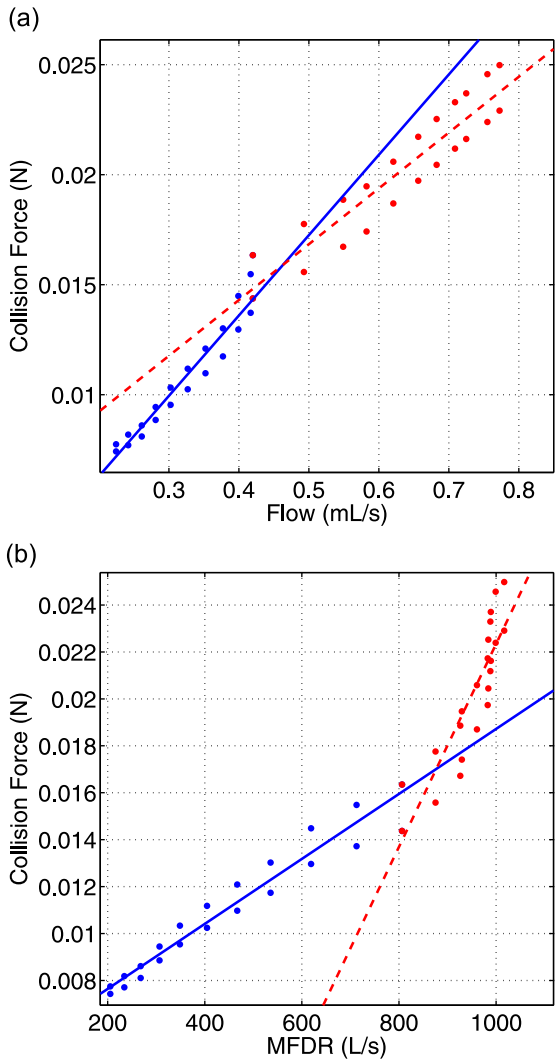

FIG. 9. (Color online) Relation between measures and collision forces: (a) Collision force vs ac-flow, not compensated (–, $R^2 = 0.93570$), SPL compensated (- -, $R^2 = 0.83806$). (b) Collision force vs MFDR, not compensated (–, $R^2 = 0.94006$), SPL compensated (- -, $R^2 = 0.73927$).

Note that DC flow is almost pathologically low when no gap is present, implying that some degree of incomplete glottal closure is needed to mimic normal behavior in numerical models.

Given that the compensated scenario mimics adducted hyperfunction, it is of interest to evaluate if the common measures such as AC flow and MFDR are correlated with the actual collision forces, as typically suspected. Several simulations with and without compensations are shown in Fig. 9, where it is observed that increased AC flow is linearly related to contact forces. Alternatively, MFDR exhibits a similar trend but with lower correlation, which is in agreement with the lower Z-score and smaller prevalence of this measure in human recordings.[6] These results suggest that high AC flow is a good indicator of high vocal fold collision forces that may lead to the formation of benign vocal fold lesions, such as nodules and polyps.

IV. CONCLUSIONS

The effects of a posterior glottal gap on the glottal aerodynamics, energy transfer, tissue dynamics, and resulting acoustics were numerically investigated using a specially designed flow solver and a reduced-order model of human phonation. The proposed representation of the posterior glottal gap yielded results that are in agreement with observations in both human subjects, for normal and hyperfunctional voices,[6,7] and rubber model experiments.[13] The proposed method has the additional benefit that it can be readily applied to any self-sustained model of the vocal folds with acoustic interaction. In this modeling approach, the presence of a posterior glottal opening produced changes in the resulting upstream and downstream acoustic pressures that, in turn, affected the glottal airflow and net energy transfer to the vibrating vocal fold tissue. The ratio between total glottal area and the equivalent vocal tract area, normally referred to as the coupling parameter,[43] became larger due to the offset in the total glottal area, for which the posterior gap opening constituted a scenario with high source-filter interaction. The inclusion of a posterior glottal opening produced rather expected changes in selected glottal measures, such as the presence of non-zero minimum flow, a steeper spectral

tilt, and a reduction in maximum flow declination rate and unsteady flow amplitude. However, a significant reduction in the energy transfer from the fluid to the vocal folds was also observed and resulted in a considerable reduction of the radiated sound pressure level from 6 to 14 dB for a range of small to large gaps, respectively. Thanks to the incorporation of a normal posterior gap opening, the body-cover model produced acoustic and aerodynamic outputs that were in the range of normal human speech production,[7] which was not the case without the gap. Compensating for a large reduction in sound pressure level (due to a large posterior gap) with higher subglottal pressure resulted in a behavior that mimicked adducted hyperfunction. In particular, a hyperfunctional behavior was observed as an abnormally high maximum flow declination rate, amplitude of unsteady flow, and minimum flow, when contrasted with the no-gap scenario and a normative data set from recordings of human subjects.[7] Insights into how these measures of vocal function are related to actual collision forces were also put forth, where the amplitude of unsteady flow was shown to be linearly correlated to the impact forces and thus was a good indicator of the potential formation of organic pathologies in the vocal folds. Future studies will explore other compensation mechanisms to increase radiated sound pressure such as vocal tract constrictions and vocal fold posturing, as well as variations arising from asymmetric tissue conditions.

ACKNOWLEDGMENTS

The authors acknowledge the support from CONICYT grants FONDECYT 11110147, STIC-AMSUD 13STIC-08, and BASAL FB00008, as well as the NIH grant R21/R33-DC011588.

[1]S. Linville, "Glottal gap configurations in two age groups of women," J. Speech Hear. Res. **35**, 1209–1215 (1992).
[2]E. B. Holmberg, R. E. Hillman, and J. S. Perkell, "Comparisons among aerodynamic, electroglotto graphic, and acoustic spectral measures of female voice," J. Speech Hear. Res. **38**, 1212–1223 (1995).
[3]H. M. Hanson, "Glottal characteristics of female speakers: Acoustic correlates," J. Acoust. Soc. Am. **101**, 466–481 (1997).
[4]B. H. Story and K. Bunton, "Production of child-like vowels with nonlinear interaction of glottal flow and vocal tract resonances," Proc. Meet. Acoust. **19**, 060303 (2013).
[5]E. B. Holmberg, R. E. Hillman, and J. S. Perkell, "Glottal air-flow and transglottal air-pressure measurements for male and female speakers in soft, normal, and loud voice," J. Acoust. Soc. Am. **84**, 511–529 (1988).
[6]R. E. Hillman, E. B. Holmberg, J. S. Perkell, M. Walsh, and C. Vaughan, "Objective assessment of vocal hyperfunction: An experimental framework and initial results," J. Speech Hear. Res. **32**, 373–392 (1989).
[7]J. S. Perkell, R. E. Hillman, and E. B. Holmberg, "Group differences in measures of voice production and revised values of maximum airflow declination rate," J. Acoust. Soc. Am. **96**, 695–698 (1994).
[8]B. Cranen and J. Schroeter, "Modeling a leaky glottis," J. Phon. **23**, 165–177 (1995).
[9]B. Cranen and J. Schroeter, "Physiologically motivated modelling of the voice source in articulatory analysis/synthesis," Speech Comm. **19**, 1–19 (1996).
[10]M. Rothenberg, "Source-tract acoustic interaction in breathy voice," in *Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control*, edited by I. R. Titze and R. C. Scherer (The Denver Center for the Performing Arts, Denver, CO, 1984), pp. 465–481.
[11]B. Cranen and L. Boves, "On subglottal formant analysis," J. Acoust. Soc. Am. **81**, 734–746 (1987).
[12]D. H. Klatt and L. C. Klatt, "Analysis, synthesis and perception of voice quality variations among male and female talkers," J. Acoust. Soc. Am. **87**, 820–856 (1990).
[13]J. B. Park and L. Mongeau, "Experimental investigation of the influence of a posterior gap on glottal flow and sound," J. Acoust. Soc. Am. **124**, 1171–1179 (2008).
[14]B. D. Erath, M. Zañartu, K. C. Stewart, M. W. Plesniak, D. E. Sommer, and S. D. Peterson, "A review of lumped-element numerical models of voiced speech," Speech Comm. **55**, 667–690 (2013).
[15]X. Pelorson, A. Hirschberg, A. P. J. Wijnands, and H. M. A. Bailliet, "Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation," J. Acoust. Soc. Am. **96**, 3416–3431 (1994).
[16]R. S. McGowan, L. L. Koenig, and A. Lofqvist, "Vocal tract aerodynamics in /aCa/ utterances: Simulations," Speech Commun. **16**, 67–88 (1995).
[17]J. C. Lucero and L. L. Koenig, "Simulations of temporal patterns of oral airflow in men and woman using a two-mass model of the vocal folds under dynamic control," J. Acoust. Soc. Am. **117**, 1362–1372 (2005).
[18]P. Birkholz, B. J. Kröger, and C. Neuscheafer-Rube, "Synthesis of breathy, normal, and pressed phonation using a two-mass model with a triangular glottis," in *Proceedings of the Interspeech 2011* (Florence, Italy, 2011), pp. 2681–2684.
[19]J. Kuo, "Voice source modeling and analysis of speakers with vocal-fold nodules," Ph.D. thesis, Division of Health Sciences and Technology, Harvard-MIT, Cambridge, MA, 1998.
[20]J. C. Lucero, "Oscillation hysteresis in a two-mass model of the vocal folds," J. Sound Vib. **282**, 1247–1254 (2005).
[21]P. Birkholz, B. J. Kröger, and C. Neuschaefer-Rube, "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* (2011), pp. 1–8.
[22]R. Scherer, B. Frazer, and G. Zhai, "Modeling flow through the posterior glottal gap," Proc. Meet. Acoust. **19**, 060240 (2013).
[23]M. Rothenberg, "A new inverse-filtering technique for deriving the glottal air flow wave-form during voicing," J. Acoust. Soc. Am. **53**, 1632–1645 (1973).
[24]D. D. Mehta, M. Zañartu, S. W. Feng, H. A. Cheyne, and R. E. Hillman, "Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform," IEEE Trans. Biomed. Eng. **59**, 3090–3096 (2012).
[25]M. Zañartu, J. C. Ho, D. D. Mehta, R. E. Hillman, and G. R. Wodicka, "Subglottal impedance-based inverse filtering of speech sounds using neck surface acceleration," IEEE Trans. Audio Speech Lang. Proc. **21**, 1929–1939 (2013).
[26]M. Ghassemi, J. H. Van Stan, D. D. Mehta, M. Zañartu, H. A. Cheyne, R. E. Hillman, and J. V. Guttag, "Learning to detect vocal hyperfunction from ambulatory neck-surface acceleration features: Initial results for vocal fold nodules," IEEE Tran. Biomed. Eng. **61**, 1668–1675 (2014).
[27]B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**, 1249–1260 (1995).
[28]K. Ishizaka and M. Matsudaira, "Fluid mechanical considerations of vocal fold vibration," in *Speech Communication Research Laboratory*, Monograph No. 8 (Santa Barbara, CA, 1972), pp. 1–152.
[29]I. R. Titze and B. H. Story, "Acoustic interactions of the voice source with the lower vocal tract," J. Acoust. Soc. Am. **101**, 2234–2243 (1997).
[30]I. R. Titze and A. S. Worley, "Modeling source-filter interaction in belting and high-pitched operatic male singing," J. Acoust. Soc. Am. **126**, 1530–1540 (2009).
[31]D. D. Mehta, M. Zañartu, T. F. Quatieri, D. D. Deliyski, and R. E. Hillman, "Investigating acoustic correlates of human vocal fold phase asymmetry through mathematical modeling and laryngeal high-speed videoendoscopy," J. Acoust. Soc. Am. **130**, 3999–4009 (2011).
[32]P. Alku, C. Magi, S. Yrttiaho, T. Bäckström, and B. Story, "Closed phase covariance analysis based on constrained linear prediction for glottal inverse filtering," J. Acoust. Soc. Am. **125**, 3289–3305 (2009).
[33]I. R. Titze, "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," J. Acoust. Soc. Am. **111**, 367–376 (2002).
[34]I. R. Titze and B. H. Story, "Rules for controlling low-dimensional vocal fold models with muscle activation," J. Acoust. Soc. Am. **112**, 1064–1076 (2002).
[35]B. H. Story, "Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract," Ph.D. thesis, University of Iowa, Iowa City, IA, 1995.
[36]H. Takemoto, K. Honda, S. Masaki, Y. Shimada, and I. Fujimoto, "Measurement of temporal changes in vocal tract area function from 3D

cine-MRI data," J. Acoust. Soc. Am. **119**, 1037–1049 (2006).
[37]E. R. Weibel, *Morphometry of the Human Lung* (Springer, New York, 1963), pp. 136–142.
[38]J. Hillenbrand, R. Cleveland, and R. Erickson, "Acoustic correlates of breathy vocal quality," J. Speech Hear. Res. **37**, 769–778 (1994).
[39]Z. Zhang, L. Mongeau, and S. H. Frankel, "Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes," J. Acoust. Soc. Am. **112**, 1652–1663 (2002).
[40]L. Mongeau, N. Franchek, C. H. Coker, and R. A. Kubli, "Characteristics of a pulsating jet through a small modulated orifice, with application to voice production," J. Acoust. Soc. Am. **102**, 1121–1133 (1997).
[41]I. R. Titze, "Parameterization of the glottal area, glottal flow, and vocal fold contact area," J. Acoust. Soc. Am. **75**, 570–580 (1984).
[42]J. Liljencrants, "Speech synthesis with a reflection-type line analog," Ph.D. thesis, Royal Institute of Technology, Stockholm, Sweden, 1985.
[43]I. R. Titze, "Nonlinear source-filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**, 2733–2749 (2008).
[44]D. G. Childers and C. K. Lee, "Vocal quality factors: Analysis, synthesis, and perception," J. Acoust. Soc. Am. **90**, 2394–2410 (1991).
[45]S. L. Thomson, L. Mongeau, and S. Frankel, "Aerodynamic transfer of energy to the vocal folds," J. Acoust. Soc. Am. **118**, 1689–1700 (2005).