

Vocal instabilities in a three-dimensional body-cover phonation model

Zhaoyan Zhang^{a)}

Department of Head and Neck Surgery, University of California, Los Angeles, 31-24 Rehabilitation Center, 1000 Veteran Avenue, Los Angeles, California 90095-1794, USA

(Received 4 June 2018; revised 17 August 2018; accepted 20 August 2018; published online 7 September 2018)

The goal of this study is to identify vocal fold conditions that produce irregular vocal fold vibration and the underlying physical mechanisms. Using a three-dimensional computational model of phonation, parametric simulations are performed with co-variations in vocal fold geometry, stiffness, and vocal tract shape. For each simulation, the cycle-to-cycle variations in the amplitude and period of the glottal area function are calculated, based on which the voice is classified into three types corresponding to regular, quasi-steady or subharmonic, and chaotic phonation. The results show that vocal folds with a large medial surface vertical thickness and low transverse stiffness are more likely to exhibit irregular vocal fold vibration when tightly approximated and subject to high subglottal pressure. Transition from regular vocal fold vibration to vocal instabilities is often accompanied by energy redistribution among the first few vocal fold eigenmodes, presumably due to nonlinear interaction between eigenmodes during vocal fold contact. The presence of a vocal tract may suppress such contact-related vocal instabilities, but also induce new instabilities, particularly for less constricted vocal fold conditions, almost doubling the number of vocal fold conditions producing irregular vibration. © 2018 Acoustical Society of America.

<https://doi.org/10.1121/1.5053116>

[JFL]

Pages: 1216–1230

I. INTRODUCTION

While phonation often involves periodic vocal fold vibration, irregular vocal fold vibration occurs in both normal and pathologic voices (Scherer, 1989; Berry, 2001). Linguistically, irregular vocal fold vibration may signal changes in meaning or offer cues to the grammatical structure of utterances (Gordon and Ladefoged, 2001; Gerratt and Kreiman, 2001; Keating *et al.*, 2015; Kuang, 2017). Clinically, while often observed in pathological conditions, vocal instabilities may also occur in functional dysphonia without obvious anatomical, neurological, or other organic deviations. Changes in vocal fold physiology, e.g., swollen vocal folds due to inflammation, often lead to irregular vocal fold vibration. The goal of this study is to identify vocal fold conditions that are more likely to lead to irregular vocal fold vibration. Identification of such conditions would allow us to better understand how we control and use our voice and to improve diagnosis and treatment of voice disorders in the clinic.

Our current understanding of vocal instabilities largely comes from excised larynx experiments and stroboscopic observation of vocal fold vibration in humans. In an early excised larynx experiment, van den Berg and Tan (1959) demonstrated the importance of the coordination between the subglottal pressure and laryngeal adjustments in maintaining regular vocal fold vibration. They observed that without simultaneous increase in thyroid cartilage compression, increasing the glottal flow often caused the vocal folds to be

blown wide apart and the shape of the glottis became irregularly curved during vibration. A more systematic experimental study was described in Isshiki (1989, 1998). Using excised larynges, he was able to map out regions of normal, breathy, and rough voice qualities in the three-dimensional parameter space of the subglottal pressure, vocal fold stiffness, and initial glottal opening area. He showed that for a given vocal fold stiffness and initial glottal opening area, increasing subglottal pressure led to voice production of a rough quality, whereas increasing vocal fold stiffness expanded the region of normal voice in the parameter space. However, considering the anisotropic nature of vocal fold mechanical properties, it is unclear which vocal fold stiffness component (transverse or longitudinal) was referred to and how it was manipulated in these studies. The perceptual evaluation of different voice qualities was not fully described either.

In summarizing previous stroboscopic investigations of vocal fold vibration, Laver (1980) argued that irregular vocal fold vibration, creak and harsh voices in his terminology, often occurs when the vocal folds are under strong medial compression, due to the actions of the lateral cricoarytenoid and thyroarytenoid muscles, and possibly weak longitudinal tension, due to relaxation of the cricothyroid muscle and strong action of the thyroarytenoid muscle. He also noted that creak and harsh voices are often produced with increased vertical thickness of the vocal folds, due to strong adduction of the vocal folds and possibly also the false folds.

One limitation of these studies is that the control parameters of voice production (i.e., vocal fold stiffness and geometry) were often not measured, probably due to difficulties in

^{a)}Electronic mail: zyzhang@ucla.edu

directly measuring such properties in excised larynges or live humans. Thus, despite these previous studies, it still remains unclear what vocal fold conditions are more likely to produce irregular vocal fold vibration.

Herzel and colleagues (Herzel, 1993; Herzel *et al.*, 1994) showed that irregular vocal fold vibration may be understood in the framework of nonlinear dynamics as coupled oscillators. They showed that regular vocal fold vibration occurs when vocal fold eigenmodes are synchronized to vibrate at the same frequency, whereas irregular vibration occurs when vocal fold eigenmodes are no longer synchronized or synchronized to two or multiple frequencies (Berry *et al.*, 1994). The coupled oscillators are not necessarily limited to vocal fold eigenmodes, and can include resonances of the sub- or supra-glottal tracts, particularly when the fundamental frequency of vocal fold vibration is close to one of the resonances of the vocal tract (Ishizaka and Flanagan, 1972; Zhang *et al.*, 2006; Titze, 2008; Zanartu *et al.*, 2011; Wade *et al.*, 2017). While this nonlinear dynamics approach has the appeal of providing a theoretical framework for better understanding irregular vocal fold vibration, it has only been applied to simplified lumped-element models of the vocal folds, and has yet to provide much insight into the individual roles of realistic vocal fold parameters in determining when vocal fold eigenmodes become synchronized or desynchronized.

In the present study, voice production is simulated in a large range of stiffness and geometry conditions, with the goal to identify vocal fold conditions that are likely to produce irregular vocal fold vibration. The three-dimensional body-cover model of the vocal folds from a previous study (Zhang, 2017a) is used. While vocal instabilities often occur in left-right asymmetric conditions, this study focuses on vocal instabilities that appear naturally in normal healthy voices and thus only considers left-right symmetric vocal fold conditions. To investigate the effect of source-tract interaction, voice simulations are performed for three vocal tract conditions (without a vocal tract, and with vocal tract shapes corresponding to the /a/ and /i/ sounds). In the following, the computational model, simulation conditions, and data analysis are described in Sec. II. The results for conditions without a vocal tract are presented in Sec. III A, whereas results with a vocal tract are described in Sec. III B, followed by discussions in Sec. IV.

II. METHOD

A. Computational model and simulation conditions

The same three-dimensional vocal fold model as in Zhang (2017a) is used in this study, as shown in Fig. 1. As this study focuses on vocal instabilities that occur in normal healthy voices, left-right symmetry in vocal fold properties (geometry, material properties, and position) about the glottal midline is imposed so that only one vocal fold is modeled in this study. The vocal fold length along the anterior-posterior (AP) direction is 17 mm. The posterior cross-sectional geometry of the vocal fold model is similarly defined as in Zhang (2017a), with medial-lateral depths of 6 and 1.5 mm for the body and cover layers, respectively, and

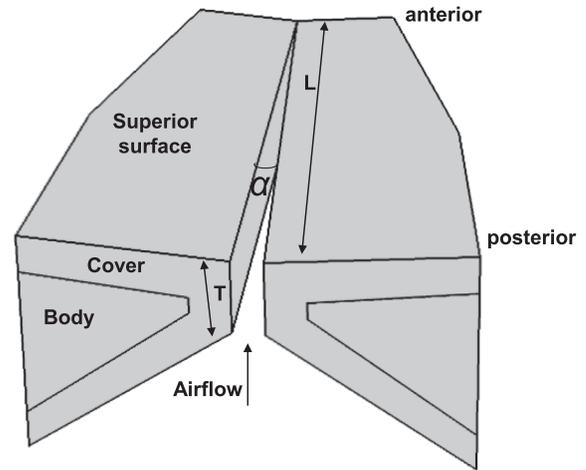


FIG. 1. The three-dimensional vocal fold model and key geometric control parameters.

a total depth of 7.5 mm. The vertical thickness of the medial surface T in the inferior-superior direction varies between 1 and 4.5 mm (Table I). In Zhang (2017a), the cover layer thickness at the lateral boundary varied with the medial surface vertical thickness T . In the present study, this dependence is removed and the cover layer thickness at the lateral boundary is set to be constant at 0.5 mm, based on measurements in Wu and Zhang (2016). Similar to that in Zhang (2017a), the vocal fold cross-section tapers quadratically toward the anterior direction, with the total depth reduced to 6.56 mm in the middle coronal plane and 3.75 mm at the anterior surface of the vocal folds, which leads to a continuously reduced body-layer depth along the AP direction while the cover layer depth remains constant at 1.5 mm. The medial surfaces of the two vocal folds form an initial glottal angle α , changes in which control the resting glottal opening or degree of vocal fold approximation. The vocal fold model is fixed at the lateral surface and the two side surfaces at the anterior and posterior ends.

Three vocal tract conditions are considered in this study. In the first condition, no sub- or supra-glottal tracts are included. Any vocal instabilities observed in this condition thus originate from laryngeal mechanisms alone. This condition also serves as a baseline condition in which source-tract interaction is absent. For the other two vocal tract conditions, a vocal tract with the shape corresponding to either /a/ or /i/ sound is included. The vocal tract is modeled as a one-dimensional waveguide (Story, 1995), and the cross-sectional area functions reported in Story *et al.* (1996) are used. The two vocal tract conditions also include an 11-cm

TABLE I. Ranges of model control parameters. For all conditions, the vocal fold density is 1030 kg/m^3 , the AP Poisson's ratio is 0.495, and $E_{ap} = 4 G_{ap}$ is assumed.

Transverse Young's modulus	$E_t = [1, 2, 4] \text{ kPa}$
Cover AP shear modulus	$G_{apc} = [1, 10, 20, 30, 40] \text{ kPa}$
Body AP shear modulus	$G_{apb} = [1, 10, 20, 30, 40] \text{ kPa}$
Vertical thickness	$T = [1, 2, 3, 4.5] \text{ mm}$
Initial glottal angle	$\alpha = [0^\circ, 1.6^\circ, 4^\circ]$
Subglottal pressure	$P_s = 50\text{--}2400 \text{ Pa}$ (18 conditions)

long uniform subglottal tract. The subglottal tract length is intentionally set to be slightly shorter than that reported in humans in order to avoid possible subglottal tract interaction with the vocal folds (Zhang *et al.*, 2006).

The model formulation regarding glottal fluid-structure-acoustics interaction has been described in our earlier studies (Zhang, 2015, 2016, 2017a). The reader is referred to these papers for details of the model formulation. Briefly, each vocal fold layer is modeled as a transversely isotropic, nearly incompressible, linear material with a plane of isotropy perpendicular to the AP direction. The material control parameters for each vocal fold layer include the transverse Young's modulus E_t , the AP Young's modulus E_{ap} , the AP shear modulus G_{ap} , and density. The density of the vocal fold is assumed to be 1030 kg/m^3 . The AP Poisson's ratio is assumed to be 0.495. As in Zhang (2017a), to reduce the number of conditions to be investigated, $E_{ap} = 4 G_{ap}$ is also assumed and the transverse Young's moduli of the two layers are assumed to be identical in the present study. Thus, the mechanical properties of the two-layer vocal fold are determined by three remaining moduli: the transverse Young's modulus E_t , the cover-layer AP shear modulus G_{apc} , and the body-layer AP shear modulus G_{apb} . For both layers, a constant loss factor of 0.4 is used, similar to Zhang (2015, 2016). The glottal flow is modeled as a one-dimensional quasi-steady glottal flow model taking into consideration of viscous loss, as described in detail in Zhang (2015, 2017a). This voice production model has been shown to agree well with experimental observations (Zhang, *et al.*, 2002; Zhang and Luu, 2012; Farahani and Zhang, 2016).

One key feature of our model is the use of an eigenmode-base formulation to significantly reduce the degree of freedom of the governing equations. Specifically, the vocal fold displacement U is approximated as linear superposition of the first N *in vacuo* eigenmodes φ_i of the vocal folds

$$U(t) = \sum_{i=1}^N q_i(t)\varphi_i, \quad (1)$$

where q is the generalized coordinates or temporal coefficients. Our previous study (Zhang, 2017b) showed that normally $N > 100$ is sufficient. In this study, $N = 300$ is used in order to better resolve vocal fold contact. In addition to reducing the degree of freedom, this approach also allows us to investigate participation of individual vocal fold eigenmodes in vocal fold vibration, changes in which often accompany vibratory regime changes (e.g., Berry *et al.*, 2006; Tokuda *et al.*, 2007; Zhang, 2009; Zhang and Luu, 2012).

For each of the three vocal tract conditions, simulations are performed with parametric variations in the medial surface thickness, initial glottal angle, subglottal pressure, and the three material moduli (E_t , G_{apb} , G_{apc}), as in Zhang (2017a) and shown in Table I. The range of material moduli is based on previous numerical and experimental studies (Titze and Talkin, 1979; Alipour-Haghighi and Titze, 1991; Alipour *et al.*, 2000; Zhang *et al.*, 2017). The range of medial surface vertical thickness is based on values used in

previous numerical studies (Titze and Talkin, 1979; Alipour *et al.*, 2000) and estimations from experimental measurements (Hollien and Curtis, 1960; Sidlof *et al.*, 2008). Note that five values in the AP shear modulus are considered for each of the body and cover layers, which leads to 25 AP stiffness conditions. In total, 16 200 conditions are investigated for each vocal tract condition, with a total of 48 600 conditions. For each condition, a half-second voice production is simulated at a sampling rate of 44 100 Hz, with the subglottal pressure linearly increased from zero to a target value in 30 time steps and then kept constant.

B. Perturbation analysis and voice type classification

For each phonating condition, perturbation analysis is performed using the last 0.3 s of each simulation by which time vocal fold vibration has reached steady-state or nearly steady-state. To focus on instabilities of vocal fold vibration, perturbation analysis is performed on the glottal area function. The peaks in the glottal area function are extracted using the "findpeaks" function in MATLAB (version 2017a), with a user-specified prominence factor. The prominence factor determines the minimal prominence a local maximum needs to have in relative to neighboring data in order to be identified as a peak. A prominence factor of 0.4 is used in the data reported below, based on a trial and error process conducted by the author using about 100 voice samples to make sure the identified voice types (see below) are largely consistent with perceptual evaluations. The amplitudes of identified peaks are then assembled into an amplitude vector, whereas the time intervals between consecutive peaks are calculated and assembled into a period vector. For each of the amplitude and period vectors, a perturbation measure P1 is calculated as the ratio between the standard deviation and mean. While P1 measures the overall level of perturbation, it will have a large value for voices with a subharmonic but otherwise regular vibration pattern (i.e., multiple peaks within one period of vocal fold vibration; an example is shown in the middle row of Fig. 2). To better identify subharmonic voices, a second perturbation measure P2 is calculated for each amplitude and period vector x of length N as follows:

$$P2 = \min(z_1, z_2, \dots, z_{10}),$$

$$z_i = y_i / \text{mean}(x); \quad y_i = \sum_{j=1}^{N-i} |x(j) - x(j+i)| / (N-i). \quad (2)$$

For a periodic vocal fold vibration with multiple peaks within one cycle, P2 would be zero whereas P1 would have a very large value.

Based on the values of P1 and P2, the corresponding vector can be classified to be either regular (with both P1 and P2 below a threshold value), subharmonic (large P1 value but small P2 value), or aperiodic (both P1 and P2 are above a threshold value). In this study, a threshold value of 0.05 is used, which is close to the mean values of jitter and shimmer reported in healthy adults (Brockmann *et al.*,

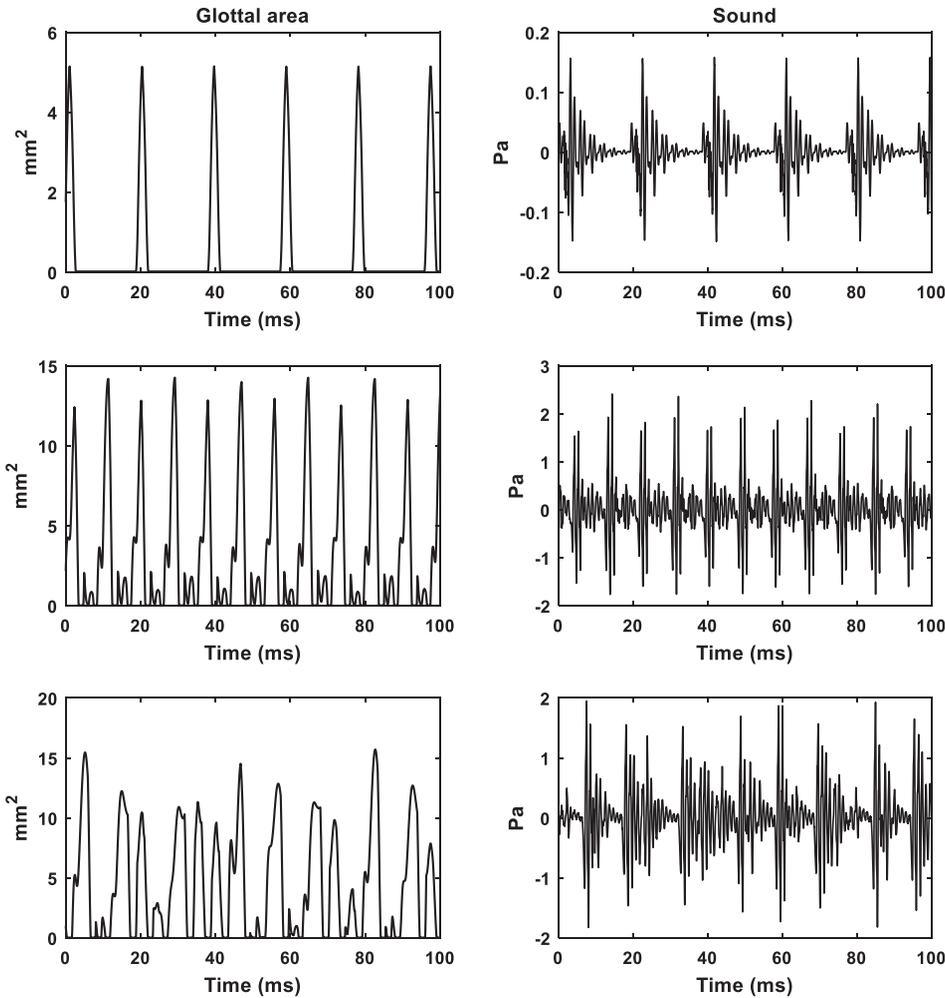


FIG. 2. Example glottal area functions and sound waveforms of the three voice types. Type 1 voices (top row) exhibit periodic vocal fold vibration. For type 2 voices (middle row), the glottal area function has multiple peaks over one oscillation cycle but otherwise is almost periodic. Type 3 voices (bottom row) show no obvious periodicity.

2008). For each simulation, the resulting voice is then categorized into three voice types (Titze, 1995): type 1 voice if both the amplitude and period vectors are regular based on their P1 and P2 values, type 3 voice if either the amplitude or period vector or both are aperiodic, and type 2 voice otherwise (i.e., when neither the amplitude nor period is aperiodic and at least one of them is subharmonic). Examples of the three voice types are shown in Fig. 2. In the following, vocal instability is considered to occur when either a type 2 or type 3 voice is identified.

III. RESULTS

A. Vocal instabilities in the absence of a vocal tract

Figure 3 shows the P2 values for the amplitude vector as a function of the control parameters and the mean glottal opening area A_{g0} . Although not shown, the P2 values for the period vector have similar trends. As expected, the cycle-to-cycle perturbation increases with increasing subglottal pressure, particularly for subglottal pressures above about 1 kPa. Figure 3 shows that the perturbation decreases with decreasing vertical thickness of the vocal fold medial surface, increasing transverse stiffness of the vocal fold in the coronal plane, and increasing glottal angle (i.e., reduced degree of vocal fold approximation). There is no obvious effect of the AP stiffness of the vocal fold in either the cover or body

layer, except for a slight effect of increasing perturbation with increasing the cover-layer AP stiffness. In general, Fig. 3 shows that vocal fold vibration is more likely to exhibit large perturbation in constricted vocal fold conditions when the mean glottal opening during phonation is small. These observations are confirmed by stepwise linear regression (Table II), which reveals significant effects on amplitude perturbation of the medial surface vertical thickness, subglottal pressure, transverse stiffness, and glottal angle, and a moderate effect (with a relatively small coefficient and a high p-value) of the cover-layer AP stiffness. The regression also shows a moderate effect of the body-layer AP stiffness on period perturbation.

The P1 measure of perturbation generally has a similar dependence on the model control parameters and the mean glottal opening area. Linear regression shows that a similar strong dependence of $P1_{amp}$ on the subglottal pressure, vertical thickness, initial glottal angle, and transverse stiffness, and a moderate dependence on the cover-layer AP stiffness. The effect of the body-layer AP stiffness on the P1 measures, for both amplitude and period, is consistently small and insignificant (p-values larger than 0.5).

Figure 4 shows the voice types in all vocal fold conditions with a transverse stiffness $E_t = 1$ kPa in the absence of a vocal tract. The voice types 1, 2, and 3 are denoted by black triangles, green squares, and red circles, respectively

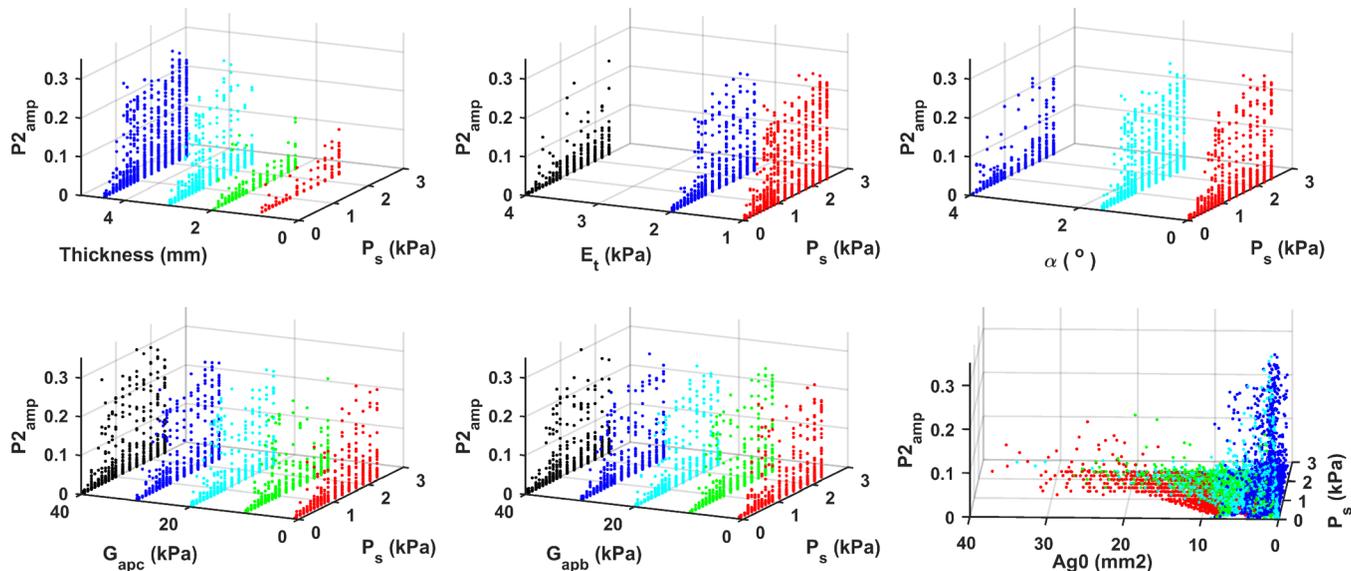


FIG. 3. (Color online) Dependence of the amplitude perturbation $P2_{amp}$ on model control parameters (subglottal pressure P_s , vertical thickness, transverse stiffness E_t , initial glottal angle α , AP shear moduli in the body and cover layers G_{apb} and G_{apc}) and the mean glottal opening area Ag_0 .

(for interpretation of the references to color, the reader is referred to the online version of this article). For convenience of data presentation, the 25 body-cover AP stiffness conditions are referred to by a G_{ap} index (refer to Table II in Zhang, 2017a) shown along the vertical axis in each panel of Fig. 4. The G_{ap} index is related to the 25 AP stiffness conditions as follows: with increasing G_{ap} index, the value of G_{apb} cyclically changes from 1, 10, 20, 30, to 40 kPa, whereas G_{apc} remains constant at 1 kPa for G_{ap} index = 1–5, changes to 10 kPa for G_{ap} index = 6–10, 20 kPa for G_{ap} index = 11–15, 30 kPa for G_{ap} index = 16–20, and 40 kPa for G_{ap} index = 21–25. In other words, the G_{ap} index is assembled in such a way so that the AP stiffnesses vary with increasing G_{ap} index in groups of five, with each group consisting of conditions with identical values of the cover-layer AP stiffness but increasing body-layer AP stiffness. Thus, the effects of increasing body-layer AP stiffness are demonstrated by the within-group variations along the G_{ap} index axis, whereas the effects of increasing cover-layer AP stiffness are demonstrated by cross-group variations. Figure 4 shows that type 2 or 3 voices are more likely to occur for higher

subglottal pressure and thicker vocal folds, and smaller glottal angles, consistent with the trends in Fig. 3. Again, there is no clear or consistent effect of the AP stiffness.

The effect of the transverse stiffness on voice types can be illustrated by comparing Fig. 4 to Fig. 5, which shows a similar voice type map but for conditions with $E_t = 4$ kPa. Increasing E_t from 1 to 4 kPa significantly reduces the number of conditions producing type 2 or 3 voices, and further limits them to conditions of thicker folds and tighter vocal fold approximation. For conditions with $E_t = 1$ kPa, out of a total of 5400 conditions examined, there are 3478 conditions producing sustained phonation, in which the numbers of type 1, 2, and 3 voices are 2,729 (78.5%), 445 (12.8%), and 304 (8.7%), respectively. For $E_t = 4$ kPa, these numbers change to 2036 (92.1%), 123 (5.6%), and 52 (2.4%), respectively, with significant decreases in both type 2 and type 3 voices.

Figures 6 and 7 (left columns) show the percentages of vocal fold conditions that produce type 2 and type 3 voices, respectively, as a function of various combinations of control parameters. Larger symbol sizes indicate higher percentages of vocal fold conditions that produce the specific voice type (type 2 in Fig. 6 and type 3 in Fig. 7). The percentage is calculated as the ratio between the number of vocal fold conditions producing the voice type of interest (type 2 or 3) and the total number of phonating vocal fold conditions for a given control parameter combination. In general, regular vocal fold vibration is more likely to occur for thin vocal folds under conditions of larger transverse stiffness, weaker vocal fold approximation, and lower subglottal pressures, consistent with the observation in Figs. 3–5.

In contrast, vocal instabilities (type 2 or type 3) are more likely to occur at constricted vocal fold conditions (thicker folds under tight approximation) with small transverse stiffness. Compared with type 2 voices, type 3 voices (Fig. 7, left column) are more restricted to conditions of large vertical medial surface thickness. In general, the percentage of type 2 voices is higher than that of type 3 voices, except at high

TABLE II. Linear regression coefficients and p -values (in parenthesis) between the P2 measures of the amplitude and period perturbations and physiological controls for conditions without a vocal tract. Similar results are observed for conditions with a vocal tract.

	Amplitude perturbation	Period perturbation
P_s (kPa)	0.0096 (2.79e-77)	0.0091 (1.56e-61)
T (mm)	0.0086 (3.03e-199)	0.0065 (5.15e-101)
α (deg.)	-0.0072 (1.74e-63)	-0.0069 (2.33e-51)
E_t (kPa)	-0.0054 (1.64e-77)	-0.0044 (1.95e-46)
G_{apc} (kPa)	2.01e-04 (4.92e-14)	1.45e-04 (3.86e-07)
G_{apb} (kPa)	-1.11e-05 (0.67) ^a	-6.01e-05 (0.028)

^aParameters with a large p value and thus exclude from the regression model.

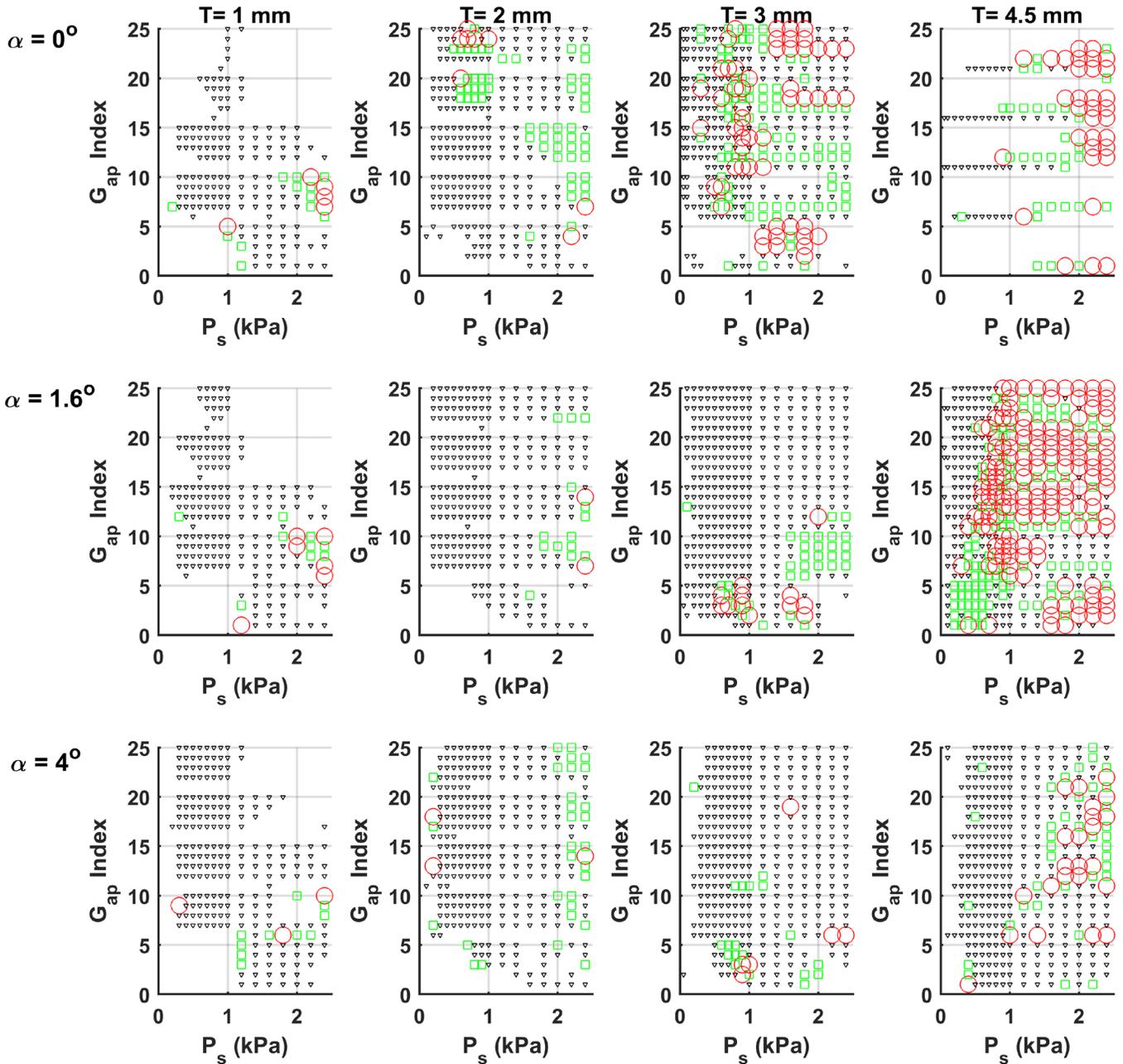


FIG. 4. (Color online) Voice type maps for conditions $E_t = 1$ kPa and no vocal tract. Type 1, type 2, and type 3 voices are denoted by triangles, squares, and circles, respectively. Each of the twelve panels shows the voice type map as a function of the AP stiffness index and subglottal pressure, for a given vertical thickness and initial glottal angle. Regions without symbols indicate conditions in which no phonation is observed.

subglottal pressures at which percentages of type 3 voices can become very high for the most constricted vocal fold conditions. While in general the percentage of voice instabilities increases with increasing subglottal pressure, this trend is more clearly observed for type 3 voices and less so for type 2 voices. In contrast to type 3 voices, the percentage of type 2 voices in constricted vocal folds at low subglottal pressures and low transverse stiffness (e.g., $E_t = 1$ kPa) is comparable to that at moderate subglottal pressures. Cross-examination with Fig. 4 reveals that this high percentage at low subglottal pressures is primarily due to the occurrence of type 2 voices at conditions of low values of G_{apc} and E_t , a large vertical thickness $T = 4.5$ mm, and an intermediate initial glottal angle. Finally, Figs. 6 and 7 also show that vocal fold conditions

near phonation onset are prone to irregular vocal fold vibration.

Figures 8 and 9 (left columns) show the percentages of type 2 and type 3 voices, respectively, at different body-cover AP stiffness conditions. Consistent with observations in Fig. 3, the effect of AP stiffness is general small except when the body or cover AP stiffness is extremely small (at 1 kPa), at which the percentage of vocal instabilities is generally higher than other stiffness conditions. In particular, Fig. 8 shows a significantly higher percentage of type 2 voices at conditions with $G_{apc} = 1$ kPa, especially at low subglottal pressures as discussed earlier.

Since no vocal tract is present, the observed vocal instability is solely due to laryngeal mechanisms. Often,

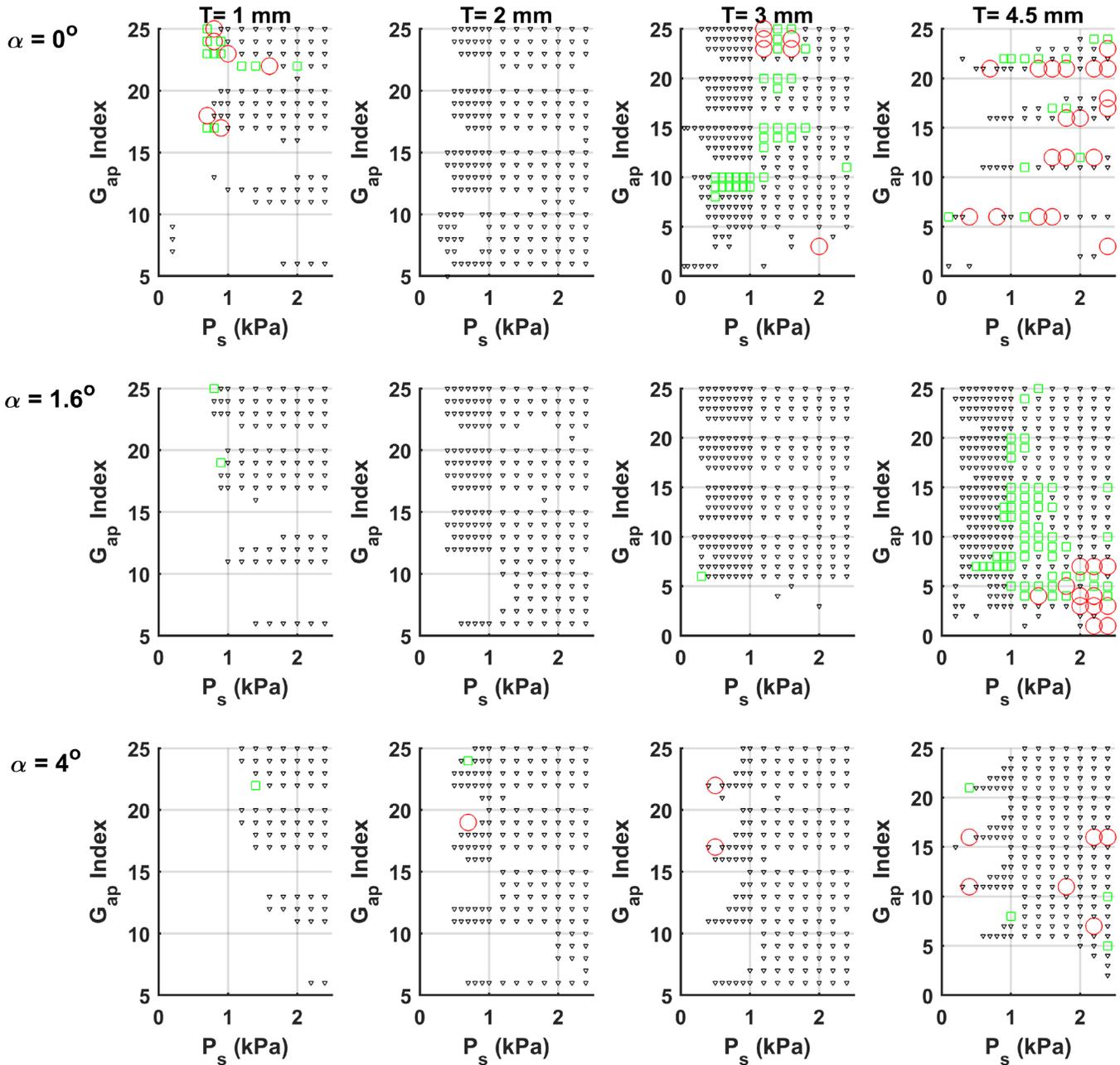


FIG. 5. (Color online) Voice type maps for conditions $E_r = 4$ kPa and no vocal tract. Type 1, type 2, and type 3 voices are denoted by triangles, squares, and circles, respectively. Each of the twelve panels shows the voice type map as a function of the AP stiffness index and subglottal pressure, for a given vertical thickness and initial glottal angle. Regions without symbols indicate conditions in which no phonation is observed.

the occurrence of vocal instability is accompanied by a substantial energy redistribution among the first few vocal fold eigenmodes. An example is given in Fig. 10. As the subglottal pressure gradually increases, vocal fold vibration changes from a regular pattern (conditions 1 and 2 in the figure) to a type-2 subharmonic vibration (condition 3) with a period-3 glottal area waveform (i.e., three notable peaks within one cycle). This change in the vibration pattern is accompanied by a substantial reduction in energy of the first three vocal fold eigenmodes and a moderate boost in energy in the eight to tenth eigenmodes. A further increase of the subglottal pressure to 700 Pa leads to even larger cycle to cycle perturbation and a type-3 vibration in condition 4.

Another example is given in Fig. 11, in which case the F0 of vocal fold vibration is more than doubled (from 112 to 291 Hz) as the subglottal pressure is increased from condition 1 to condition 3, with the transitioning condition 2 exhibiting a period-doubling vibration pattern. The duration of vocal fold closure is also significantly reduced, with the closed quotient reducing from 0.67 in condition 1 to 0.3 in condition 3. Both changes are reminiscent of a chest-falsetto register change, although the change is triggered by a change in the subglottal pressure alone without any simultaneous laryngeal adjustments. Again, such transition is accompanied by significant redistribution in eigenmode energy, in this case a significant reduction in energy of the second eigenmode and a strong energy increase in the first eigenmode.

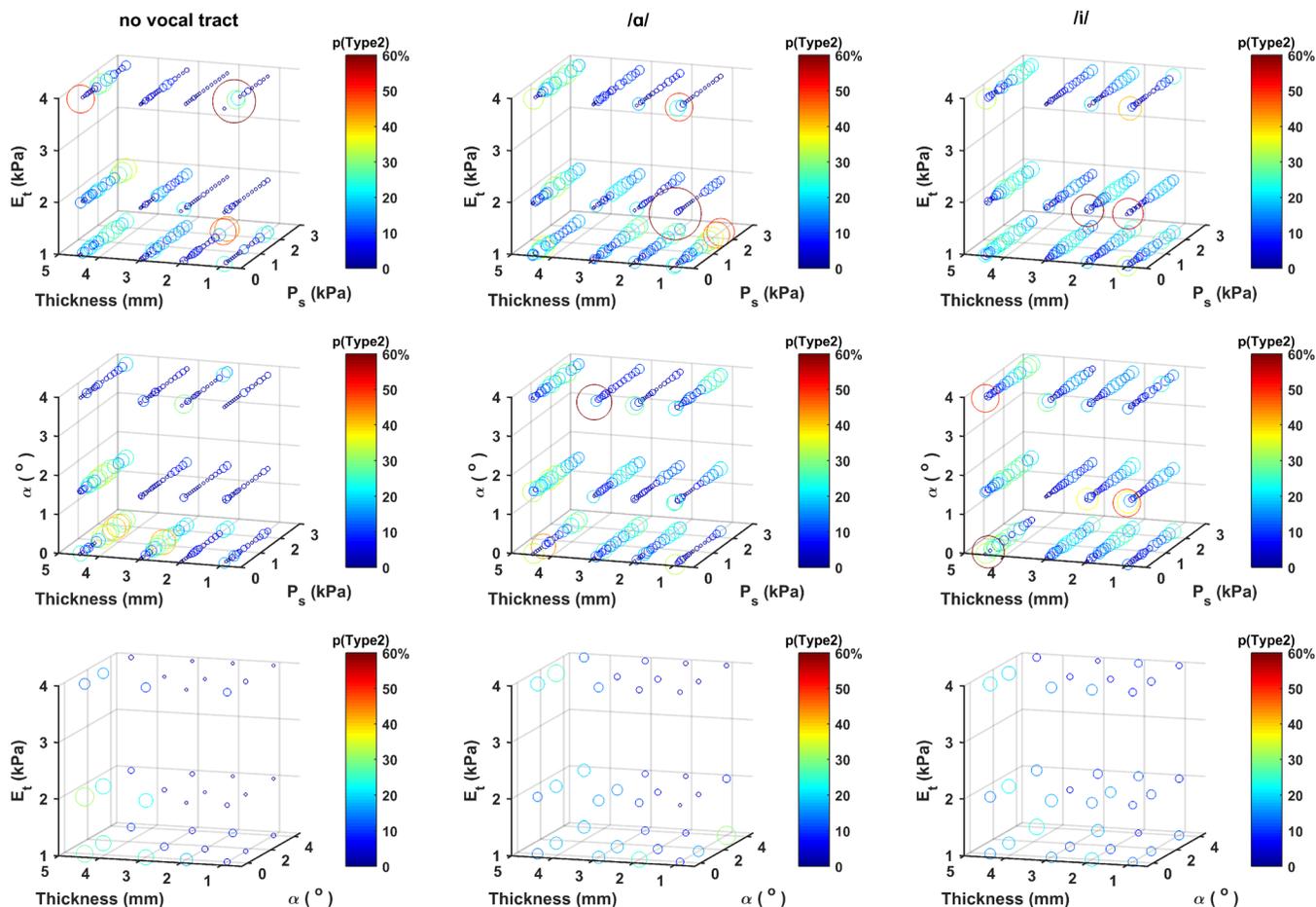


FIG. 6. (Color online) Percentages of vocal fold conditions that produce type 2 voices as a function of the subglottal pressure (P_s), medial surface vertical thickness, transverse stiffness (E_t), and initial glottal angle (α). A larger symbol size indicates a higher percentage of vocal instabilities. The percentages are also color coded in the online version of the paper. The left, middle, and right columns are for conditions without a vocal tract, with a vocal tract corresponding to /a/ and /i/, respectively.

While vocal instabilities in Figs. 10 and 11 are induced by changes in the subglottal pressure, transitions between vocal instabilities and energy redistribution among eigenmodes can also occur due to changes in vocal fold stiffness and geometry, as one traverses the voice type maps in Figs. 4 and 5 along different directions.

It is worth noting that voices with a vocal fry-like vocal fold vibration, in the absence of significant cycle to cycle variations (e.g., the top row in Fig. 2), would be categorized as a type 1 voice, despite that it is often perceived as a creaky voice. This kind of voices is often produced by thick vocal folds at very low subglottal pressures, as shown in the upper rightmost panel in Fig. 4. Increasing subglottal pressure would lead to subharmonics and eventually aperiodic type-3 vocal fold vibration.

B. Vocal instabilities in the presence of a vocal tract

Figure 12 shows the voice type map for $E_t = 1$ kPa with a vocal tract corresponding to the /a/ sound. Comparing to Fig. 4, one overall trend is that the number of vocal fold conditions exhibiting type 2 or type 3 voices is much higher in Fig. 12, particular for the less constricted vocal fold conditions (lower and left panels in the figure). The presence of a vocal tract can either suppress or maintain vocal instabilities that are existing without a vocal tract, or induce new vocal instabilities. As

shown in Table III, out of the 1371 type 2 or type 3 vocal instabilities observed in the absence of a vocal tract, about half (659 conditions for /a/ and 596 for /i/) are suppressed when a vocal tract is included. One such example is given in Fig. 13, which shows voice production for the same vocal fold conditions shown in Fig. 10 but with the /a/ vocal tract. Vocal instabilities are no longer observed in conditions 3 and 4, for which subharmonic and irregular vocal fold vibration are observed without a vocal tract in Fig. 10. Note that the energy distribution among eigenmodes in Fig. 13 is quite different from that in Fig. 10.

Of all the vocal instabilities observed in the presence of a vocal tract, only about one third are existing vocal instabilities (i.e., vocal instabilities are observed for the same vocal fold conditions even without a vocal tract), whereas the remaining are new vocal instabilities that are absent without a vocal tract. In total, the number of vocal fold conditions exhibiting type 2 or type 3 voices in the presence of a vocal tract is almost doubled compared with that in the absence of a vocal tract. Although adding a vocal tract allows more vocal fold conditions to phonate, the percentage of phonating vocal fold conditions that produce either a type 2 or type 3 voice still increases from 15.2% to 21.0% for /a/ and 25.0% for /i/.

The percentages of type 2 and type 3 voices observed under specific combinations of control parameters are shown

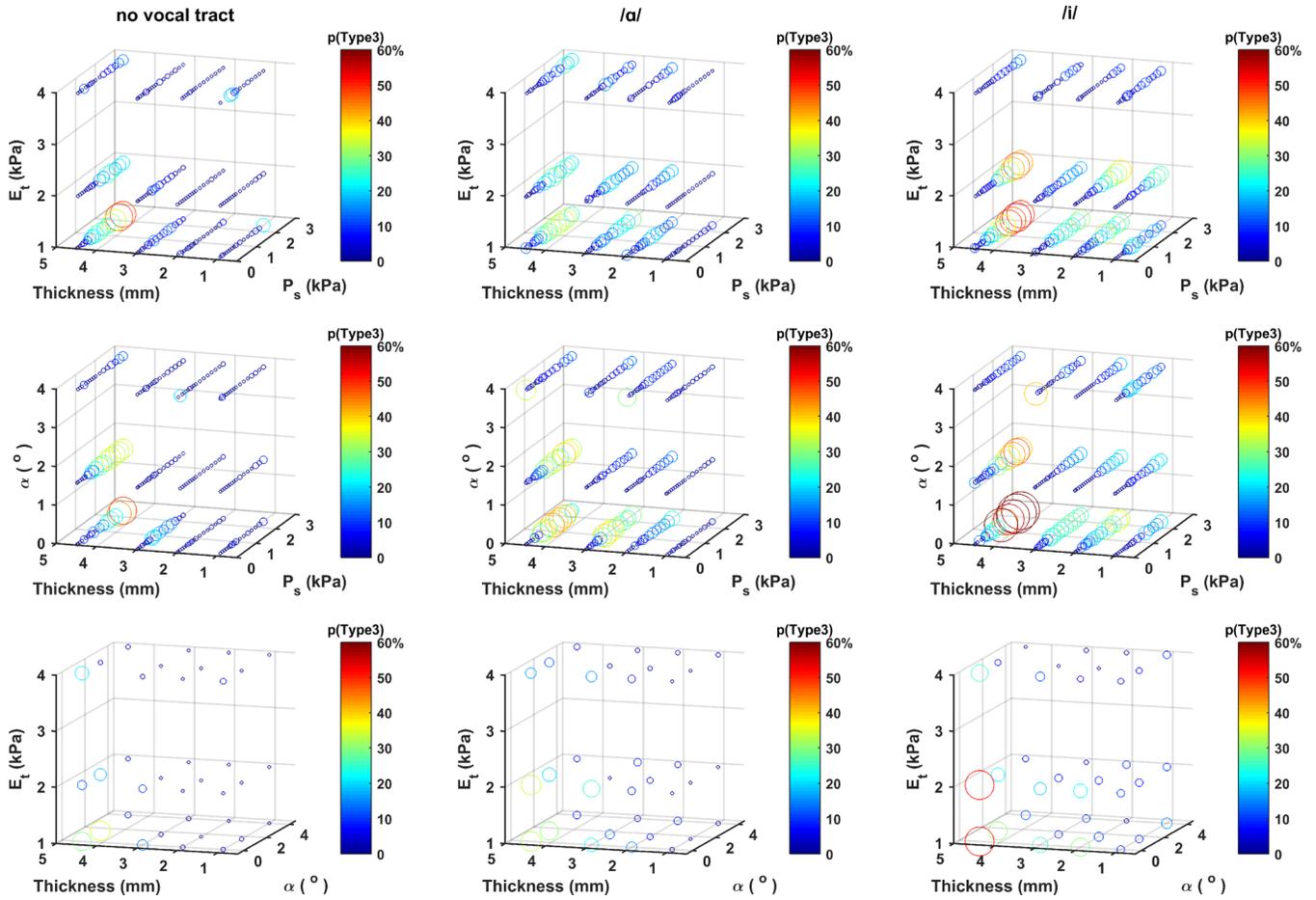


FIG. 7. (Color online) Percentages of vocal fold conditions that produce type 3 voices as a function of the subglottal pressure (P_s), medial surface vertical thickness, transverse stiffness (E_t), and initial glottal angle (α). A larger symbol size indicates a higher percentage of vocal instabilities. The percentages are also color coded in the online version of the paper. The left, middle, and right columns are for conditions without a vocal tract, with a vocal tract corresponding to /a/ and /i/, respectively.

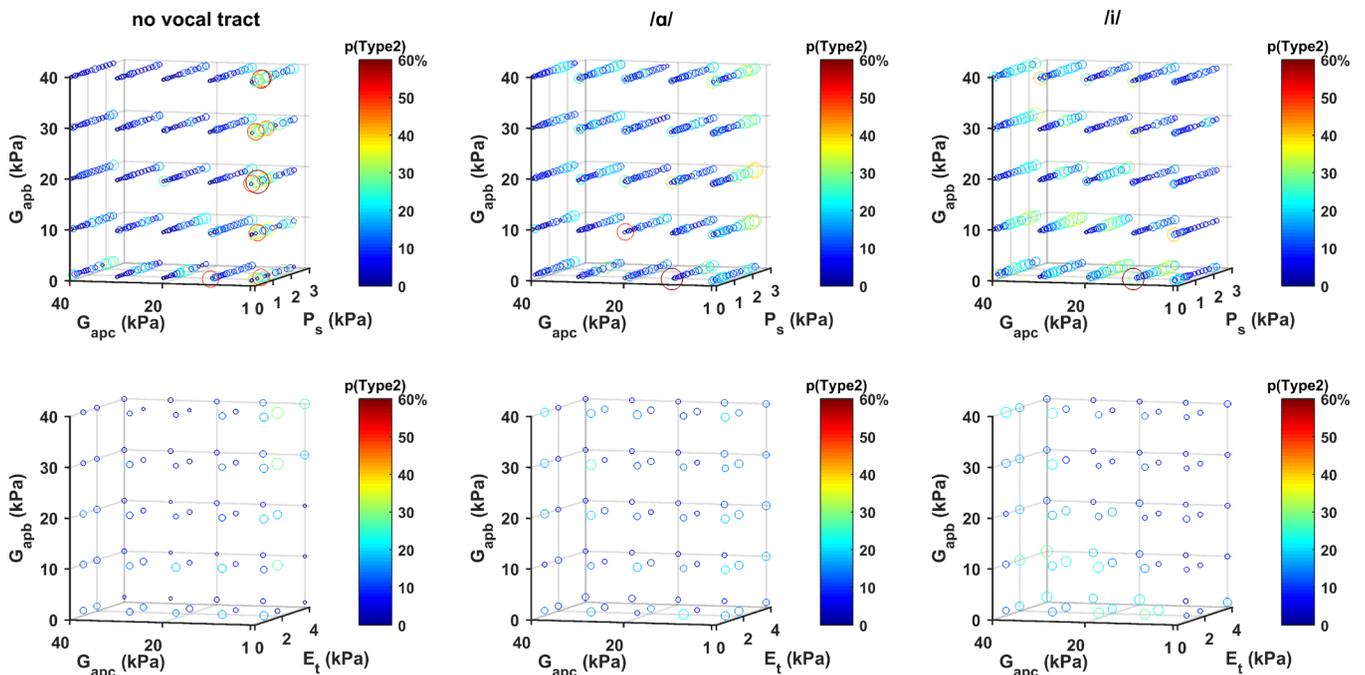


FIG. 8. (Color online) Percentages of vocal fold conditions that produce type 2 voices as a function of the AP stiffnesses of the body and cover layers (G_{apb} and G_{apc}). A larger symbol size indicates a higher percentage of vocal instabilities. The percentages are also color coded in the online version of the paper. The left, middle, and right columns are for conditions without a vocal tract, with a vocal tract corresponding to /a/ and /i/, respectively.

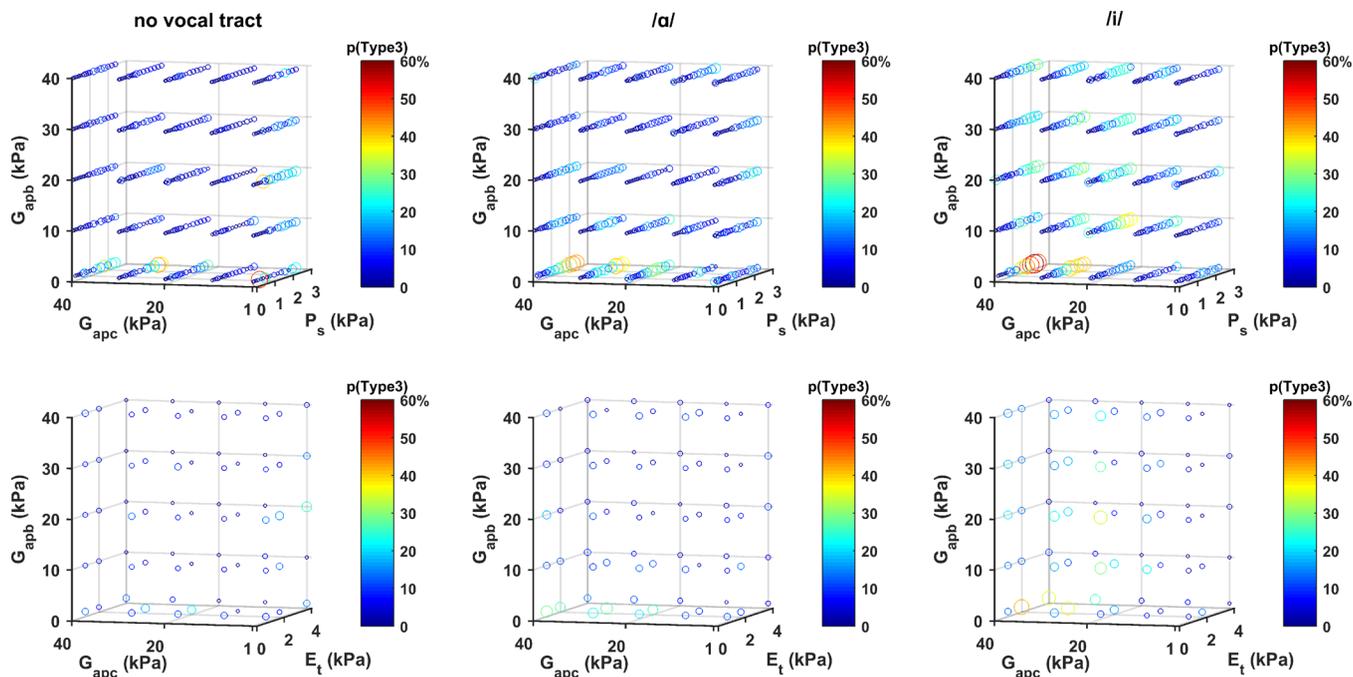


FIG. 9. (Color online) Percentages of vocal fold conditions that produce type 3 voices as a function of the AP stiffnesses of the body and cover layers (G_{apb} and G_{apc}). A larger symbol size indicates a higher percentage of vocal instabilities. The percentages are also color coded in the online version of the paper. The left, middle, and right columns are for conditions without a vocal tract, with a vocal tract corresponding to /a/ and /i/, respectively.

in Figs. 6 and 7, respectively, with the middle columns for /a/ and right columns for /i/. Compared with the case without a vocal tract, the presence of the /a/ vocal tract slightly decreases the percentage of vocal instabilities (particularly type 2 voices) for the most constricted vocal fold conditions (i.e., thickest fold with the smallest transverse stiffness and smallest initial glottal angle), but significantly increases the percentage of vocal instabilities for other less constricted vocal fold conditions. In contrast, the /i/ vocal tract increases the percentage of vocal instabilities for almost all vocal fold conditions. In particular, the percentage of type 3 voices under constricted vocal fold conditions is significantly higher in the presence of the /i/ vocal tract than the other two vocal tract conditions. The presence of a vocal tract also leads to increased percentages of vocal instabilities at certain AP stiffness conditions (Figs. 8 and 9). However, the effect of different AP stiffness conditions remains generally small even in the presence of a vocal tract.

To further understand the mechanisms of vocal instabilities in the presence of a vocal tract, the percentages of new and existing vocal instabilities are separately calculated as a function of different combinations of the model control parameters. For both vocal tracts, the percentages of existing vocal instabilities show a similar dependence on the control parameters as that in the case without a vocal tract (left column, Fig. 6), with high percentage values occurring at the most constricted vocal fold conditions (i.e., thick vocal folds with a low transverse stiffness and a small initial glottal angle). The F_0 of vocal fold vibration also clusters mostly around 0.5, 1, and 2 times of the corresponding F_0 observed for conditions without a vocal tract. Thus, it is reasonable to assume most of the existing vocal instabilities observed in the presence of a vocal tract are due to primarily laryngeal mechanisms.

For the new vocal instabilities, high percentage values are no longer restricted to the most constricted vocal fold conditions, particularly for the /i/ vocal tract. Except for the least constricted vocal fold conditions, all other conditions exhibit moderate percentages of new vocal instabilities. Presumably, the new instabilities could be due to either laryngeal mechanisms, which by themselves are not strong enough and thus require assistance from the presence of a vocal tract to induce vocal instabilities, or entrainment of vocal fold vibration to a vocal tract resonance as demonstrated in previous studies (Ishizaka and Flanagan, 1972; Zhang *et al.*, 2006; Titze, 2008; Zanartu *et al.*, 2011; Wade *et al.*, 2017). However, due to the clustering of vocal fold eigenfrequencies and vocal tract resonances in the frequency range surrounding F_0 , it is often difficult to determine which mechanism, laryngeal or source-tract interaction, plays a dominant role in inducing vocal instabilities. The condition 1 in Figs. 10 and 13 is such an example, in which vocal fold vibration is periodic in Fig. 10 without a vocal tract but becomes subharmonic in the presence of the /a/ vocal tract in Fig. 13. Note again that the energy distribution among vocal fold eigenmodes in condition 1 is different with and without the /a/ vocal tract. In fact, the energy distribution among eigenmodes is quite different between Figs. 10 and 13 even for condition 2, which exhibits regular vibration with and without a vocal tract. In other words, the presence of a vocal tract may still have a significant impact on vocal fold vibration even if the voice type remains unchanged.

Figure 14 shows an example in which vocal instability is clearly induced by strong interaction between vocal fold vibration and a vocal tract resonance. Increasing subglottal pressure causes the F_0 of vocal fold vibration to gradually approach the first formant (around 300 Hz) of the /i/ vocal

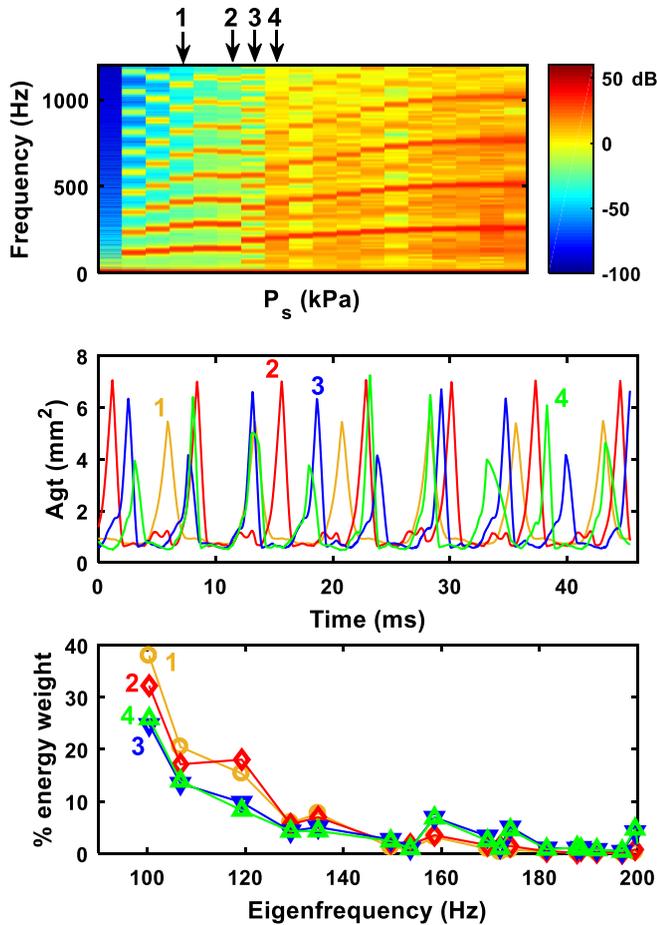


FIG. 10. (Color online) Transition from regular vocal fold vibration (conditions 1 and 2) to period-3 subharmonics (condition 3) and irregular vibration (conditions 4) is accompanied by energy redistribution among vocal fold eigenmodes. Top panel shows the spectrogram of the glottal area function as the subglottal pressure is increased. The middle and bottom panels show the glottal area functions and the percentage energy weights of the first few vocal fold eigenmodes for the four pressure conditions indicated in the top panel.

tract. When the two frequencies are sufficiently close to each other, the vocal fold vibration changes to a period-doubling vibration pattern, with each cycle consisting two peaks. As the subglottal pressure continues to increase, the amplitude difference between the two peaks becomes even larger. It is worth noting that in this case, there is no significant change in the energy distribution among vocal fold eigenmodes, although the energy distribution is quite different from that in conditions without a vocal tract.

IV. DISCUSSION AND CONCLUSIONS

This study has identified the subglottal pressure, medial surface vertical thickness, vocal fold transverse stiffness in the coronal plane, and initial glottal angle (degree of vocal fold approximation) as the primary determinants of vocal instabilities (subharmonic or aperiodic vocal fold vibration). Vocal instabilities are more likely to occur in thick vocal folds with low transverse stiffness under tight vocal fold approximation, with or without a vocal tract. In humans, a thick vocal fold with low transverse stiffness can be achieved by strong activation of the thyroarytenoid muscle coupled with relaxation of the cricothyroid muscle (Hirano

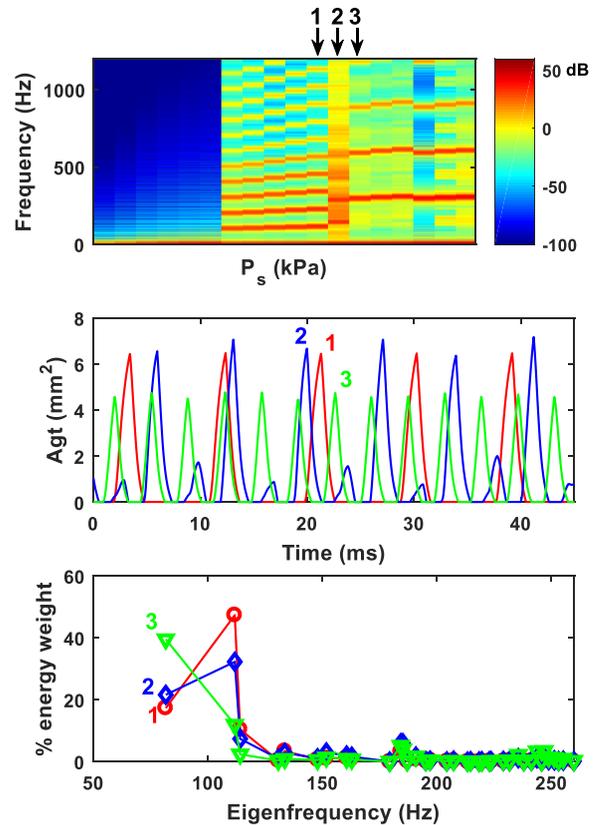


FIG. 11. (Color online) Transition from regular vocal fold vibration in condition 1 to another regular vibration in condition 3 through an intermediate subharmonic condition 2 is accompanied by significant energy redistribution among the first few vocal fold eigenmodes and significant changes in the F0 and closed quotient. Top panel shows the spectrogram as the subglottal pressure is increased. The middle and bottom panels show the glottal area functions and the percentage energy weights of the first few vocal fold eigenmodes for the three pressure conditions indicated in the top panel.

and Kakita, 1985; Hirano, 1988; Vahabzadeh-Hagh *et al.*, 2017; Zhang *et al.*, 2017). Tighter vocal fold approximation can be achieved through increased actions of the lateral cricoarytenoid and thyroarytenoid muscles. Thus, the results of this study show that vocal instabilities are more likely to occur under strong vocal fold adduction. Vocal fold thickening may also occur at conditions of vocal fold swelling (e.g., due to vocal fold inflammation resulting from a cold or in response to phonotrauma). To avoid vocal instabilities, one may relax the lateral cricoarytenoid and thyroarytenoid muscles (which reduces vocal fold approximation, thins the vocal folds, and increases the transverse stiffness), or increase cricothyroid activation (which increases the transverse stiffness and thins the vocal folds). In general, vocal instabilities can also be suppressed by reducing the subglottal pressure, particularly for type 3 voices. However, in this study type 2 voices have also been observed at very low subglottal pressures, for conditions of low overall vocal fold stiffness (both transverse and AP stiffnesses) and large thickness. At such low subglottal pressures, because the laryngeal adjustments mentioned above (especially increasing transverse stiffness) often lead to increased phonation threshold pressure, a simultaneous increase in the subglottal pressure may be required to suppress instability and maintain phonation, as often observed in creak or vocal fry.

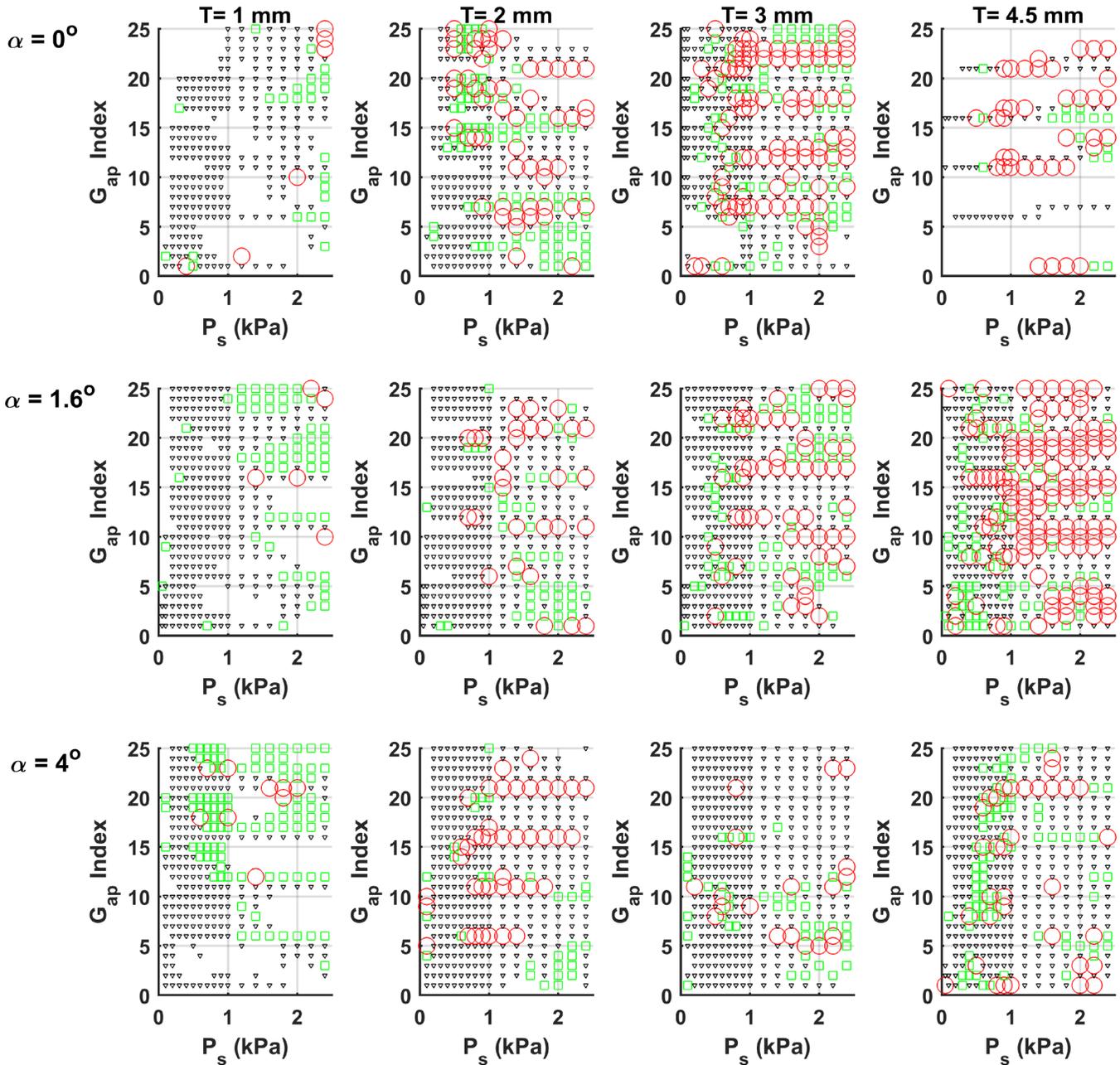


FIG. 12. (Color online) Voice type maps for conditions $E_f = 1$ kPa and an /a/ vocal tract. Type 1, type 2, and type 3 voices are denoted by triangles, squares, and circles, respectively. Each of the twelve panels shows the voice type map as a function of the AP stiffness index and subglottal pressure (P_s), for a given vertical thickness (T) and initial glottal angle (α). Regions without symbols indicate conditions in which no phonation is observed.

TABLE III. Number of vocal fold conditions that exhibit type 2 or type 3 voices and the total number of conditions that phonate under different vocal tract conditions. For conditions with a vocal tract, “existing” refers to vocal fold conditions in which type 2 or type 3 instability has been observed with and without a vocal tract, and “new” refers to vocal fold conditions in which type 2 or type 3 instability is observed only when the corresponding vocal tract is included.

	No vocal tract	/a/		/i/	
		existing	new	existing	new
type 2 or 3	1,317	658	1,839	721	2,222
with phonation	8,693	11,891		11,783	
percentage	15.2%	21.0%		25.0%	

The results of this study are consistent with previous studies. For example, [Isshiki \(1989, 1998\)](#) showed that in excised larynx experiments, for a given vocal fold stiffness and initial glottal opening area, increasing subglottal pressure led to voice production of a rough quality, and increasing vocal fold stiffness or the initial glottal opening allowed normal phonation to be maintained at higher subglottal pressures. This is consistent with the results of our study regarding type 3 voices, if the vocal fold stiffness in the Isshiki study is interpreted as the transverse stiffness in the coronal plane. [Laver \(1980\)](#) has argued that aperiodic vocal fold vibration or harsh voice is produced with strong medial compression, which often leads to thickened vocal folds. This is consistent with

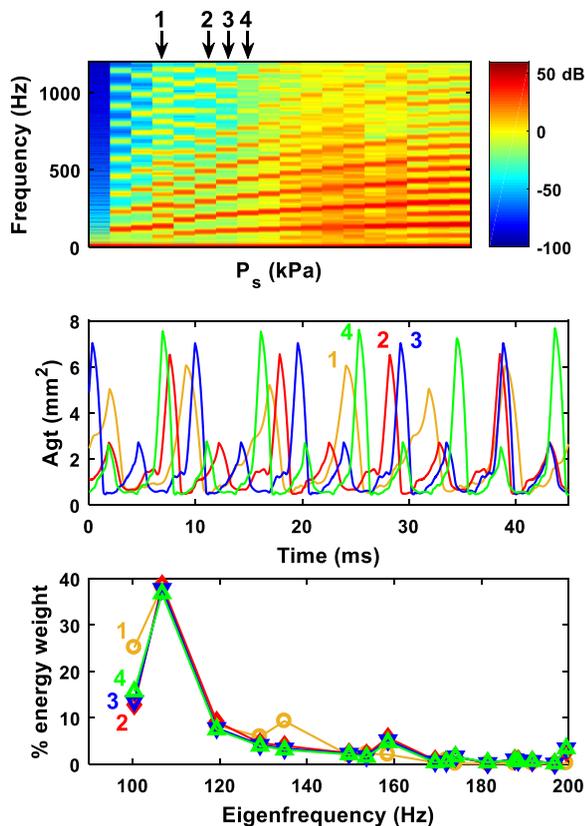


FIG. 13. (Color online) Spectrogram (top), glottal area functions (middle), and percentage energy weights of the first few eigenmodes (bottom) for the same vocal fold conditions shown in Fig. 10 but with the /a/ vocal tract. The presence of the vocal tract suppresses vocal instabilities that are otherwise present in conditions 3 and 4 in Fig. 10, but also induces subharmonic vibration in condition 1 which otherwise exhibits regular vibration without a vocal tract.

the observation in this study that thick vocal folds under tight approximation are more likely to produce type 3 voices.

This study also shows that transverse stiffness plays an important role in facilitating or suppressing vocal instabilities. This is consistent with the observation in [Berry et al. \(1994\)](#), in which irregular vocal fold vibration (e.g., subharmonic or chaotic vibration) was observed when the transverse stiffness in the cover layer was significantly reduced. In general, our study shows that the AP stiffness, which is closely related to AP tension, has a much smaller role. However, very low AP stiffness in the cover layer, when coupled with low transverse stiffness, large thickness, and an intermediate initial glottal angle, does facilitate the occurrence of type 2 voices at low subglottal pressures (Figs. 4 and 12), which appears to be consistent with the description of creaky voice by [Laver \(1980\)](#). In reality, however, the AP stiffness is often positively correlated with the transverse stiffness, because both are regulated by vocal fold shortening ([Zhang et al., 2017](#)). Because an increase in either the AP or transverse stiffness increases F_0 , our results also suggest that vocal instabilities (e.g., creaky voice) are less likely to occur with increasing F_0 ([Kuang, 2017](#)).

[Slifka \(2006\)](#) reported irregular vocal fold vibration (primarily type 3) at the end of utterances when the vocal folds are moving apart and the subglottal pressure is decreasing (around or below 500 Pa). In other words, such end of

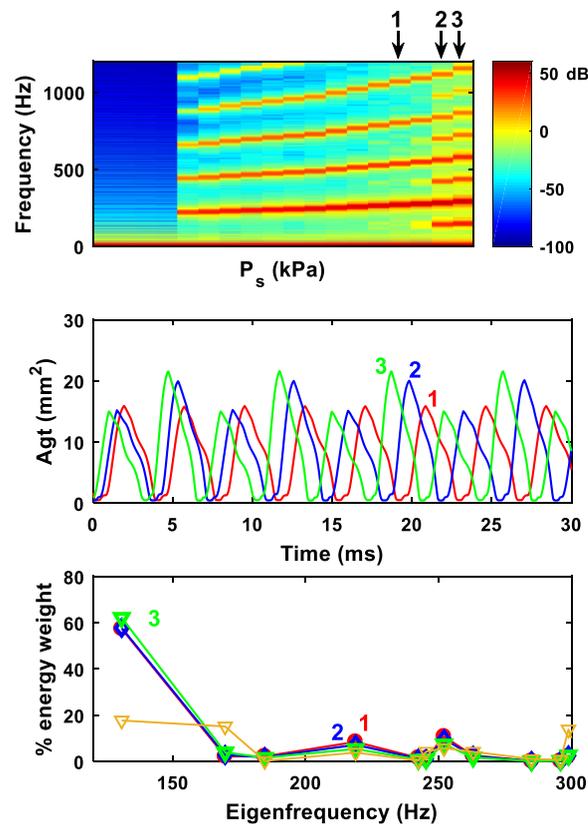


FIG. 14. (Color online) Source-tract interaction induced vocal instability occurs as the F_0 approaches the first formant of the /i/ vocal tract. Top panel shows the spectrogram as the subglottal pressure is increased. The middle and bottom panels show the glottal area functions and the percentage energy weights of the first few vocal fold eigenmodes for the three conditions indicated in the top panel. The thin line in the bottom row is the percentage energy weights for condition 3 in the case without a vocal tract.

utterance vocal instabilities generally occur at a non-constricted glottal configuration with low subglottal pressure. Although vocal instabilities in this study mostly occur at higher subglottal pressures, some vocal instabilities (mostly type 2 voices) are observed at low subglottal pressures for thick vocal folds with low overall stiffness (Fig. 4). Our study also shows the tendency of vocal folds to exhibit instabilities near phonation onset. On the other hand, it is possible such end-of-utterance vocal instabilities may be related to the unsteady nature at the end of utterance, in which “many parameters that control phonation are changing in the same time span” ([Slifka, 2006](#)). The effect of such changes in vocal control may be better captured in a dynamic voice simulation with time-varying laryngeal and respiratory controls as in [Zhang \(2017b\)](#) rather than a sustained phonation as in the present study ([Keating et al., 2015](#)), which needs further investigation.

This study reveals different mechanisms of vocal instability. One mechanism originates solely at the laryngeal level, with the transition to vocal instabilities involving energy redistribution among the first few vocal fold eigenmodes. Similar changes in vocal fold vibration patterns due to energy redistribution among vocal fold eigenmodes have been observed in previous studies (e.g., [Tokuda et al., 2007](#); [Zhang, 2009](#)). However, it is still unclear why and at what conditions such energy redistribution occurs. Considering

that vocal instabilities are more likely to occur for vocal fold conditions that produce tight glottal closure (a small mean glottal opening during vibration; Fig. 3), it is reasonable to assume that vocal fold contact plays an important role in facilitating energy redistribution among eigenmodes and inducing vocal instabilities. This may particularly be the case for thick vocal folds, which vibrate with a long period of vocal fold closure. Vocal fold contact, due to its inherent nonlinearity, enhances cross-mode interaction, and thus may suppress some vocal fold eigenmodes but more strongly excite other eigenmodes. Changes in the degree of vocal fold contact (e.g., due to increasing subglottal pressure) can lead to significant F0 changes (Zhang, 2016), which may favor energy redistribution to eigenmodes of nearby eigenfrequencies. With the same reasoning, a reduced degree in vocal fold approximation would reduce the extent of vocal fold contact along the anterior-posterior direction, and thus reduce cross-mode interaction and potential energy redistribution among vocal fold eigenmodes. There is no clear explanation as to why increased transverse stiffness suppresses vocal instabilities. This could be due to the reduced vocal fold contact with increasing transverse stiffness (Zhang, 2017a), or increased frequency spacing between eigenmodes which makes it more difficult for cross-mode interaction and thus energy redistribution among eigenmodes. Future studies are required to better understand this mechanism.

Previous studies have shown that vocal instabilities arise when the F0 or lower harmonics of vocal fold vibration approach one of the resonances of the subglottal or supraglottal tracts (Ishizaka and Flanagan, 1972; Zhang *et al.*, 2006; Titze, 2008; Zanartu *et al.*, 2011; Wade *et al.*, 2017). This is confirmed in our study, in which vocal instabilities are often observed when the F0 of vocal fold vibration approaches a nearby vocal tract resonance with increasing subglottal pressure. A new finding in our study is that at certain vocal fold conditions (thick folds, low transverse stiffness, and tight vocal fold approximation), the mere presence of a vocal tract is sufficient to induce vocal instabilities by destabilizing an otherwise regular vocal fold contact pattern, without the F0 or harmonics being close to an obvious vocal tract resonance. In our study, the presence of a vocal tract significantly increases the number of vocal fold conditions that produce type 2 or 3 vocal instabilities, even for the /a/ sound in which the first formant is much higher than the F0 range investigated. Our preliminary study demonstrates the same trend when the vocal tract damping factor is doubled. However, it is possible that with a more realistic vocal tract loss model and data (Hanna *et al.*, 2016), the effect of the vocal tract on vocal instability may be reduced, which merits further investigation.

Although our model captures many important geometric features of human vocal folds, it is still simplified compared with realistic human vocal fold geometry (Wu and Zhang, 2016). The findings of this study thus need to be verified in experiments or simulations using realistic human vocal fold geometry. Similarly, a better experimental characterization of the anisotropic mechanical properties of the vocal folds at different voice conditions would allow future simulations to

focus on mechanical conditions that are more likely to occur in human phonation. Finally, the vocal fold surface condition is likely to play an important role in vocal fold contact and the contact-induced vocal instabilities. These issues will be addressed in future studies.

ACKNOWLEDGMENTS

This study was supported by research Grant Nos. R01 DC001797, R01 DC009229, and R01 DC011299 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health.

- Alipour, F., Berry, D. A., and Titze, I. R. (2000). "A finite-element model of vocal-fold vibration," *J. Acoust. Soc. Am.* **108**, 3003–3012.
- Alipour-Haghighi, F., and Titze, I. R. (1991). "Elastic models of vocal fold tissues," *J. Acoust. Soc. Am.* **90**, 1326–1331.
- Berry, D. A. (2001). "Mechanisms of modal and nonmodal phonation," *J. Phon.* **29**, 431–450.
- Berry, D. A., Herzel, H., Titze, I. R., and Krischer, K. (1994). "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions," *J. Acoust. Soc. Am.* **95**, 3595–3604.
- Berry, D. A., Zhang, Z., and Neubauer, J. (2006). "Mechanisms of irregular vibration in a physical model of the vocal folds," *J. Acoust. Soc. Am.* **120**, EL36–EL42.
- Brockmann, M., Storck, C., Carding, P. N., and Drinnan, M. J. (2008). "Voice loudness and gender effects on jitter and shimmer in healthy adults," *J. Speech Lang. Hear. Res.* **51**, 1152–1160.
- Farahani, M., and Zhang, Z. (2016). "Experimental validation of a three-dimensional reduced-order continuum model of phonation," *J. Acoust. Soc. Am.* **140**, EL172–EL177.
- Gerratt, B., and Kreiman, J. (2001). "Toward a taxonomy of nonmodal phonation," *J. Phon.* **29**, 365–381.
- Gordon, M., and Ladefoged, P. (2001). "Phonation types: A cross-linguistic overview," *J. Phon.* **29**, 383–406.
- Hanna, N., Smith, J., and Wolfe, J. (2016). "Frequencies, bandwidths and magnitudes of vocal tract and surrounding tissue resonances, measured through the lips during phonation," *J. Acoust. Soc. Am.* **139**(5), 2924–2936.
- Herzel, H. (1993). "Bifurcations and chaos in voice signals," *Appl. Mech. Rev.* **46**(7), 399–413.
- Herzel, H., Berry, D. A., Titze, I. R., and Saleh, M. (1994). "Analysis of vocal disorders with methods from nonlinear dynamics," *J. Speech. Hear. Res.* **37**, 1008–1019.
- Hirano, M. (1988). "Vocal mechanisms in singing: Laryngological and phoniatric aspects," *J. Voice* **2**, 51–69.
- Hirano, M., and Kakita, Y. (1985). "Cover-body theory of vocal fold vibration," in *Speech Science: Recent Advances*, edited by R. G. Daniloff (College-Hill Press, San Diego), pp. 1–46.
- Hollien, H., and Curtis, F. (1960). "A laminagraphic study of vocal pitch," *J. Speech Hear. Res.* **3**, 361–371.
- Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Tech. J.* **51**, 1233–1267.
- Isshiki, N. (1989). *Phonosurgery: Theory and Practice* (Springer-Verlag, Tokyo), Chap. 3.
- Isshiki, N. (1998). "Mechanical and dynamical aspects of voice production as related to voice therapy and phonosurgery," *J. Voice* **12**, 125–137.
- Keating, P., Garellek, M., and Kreiman, J. (2015). "Acoustic properties of different kinds of creaky voice," in *Proceedings of the 18th International Congress of Phonetic Sciences*, Paper No. 0821.1-5.
- Kuang, J. (2017). "Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice," *J. Acoust. Soc. Am.* **142**, 1693–1706.
- Laver, J. (1980). *The Phonetic Description of Voice Quality* (Cambridge University Press, Cambridge), Chap. 3.
- Scherer, R. (1989). "Physiology of creaky voice and vocal fry," *J. Acoust. Soc. Am.* **86**, S25.
- Sidlof, P., Svec, J., Horacek, J., Vesely, J., Klepacek, I., and Havlik, R. (2008). "Geometry of human vocal folds and glottal channel for mathematical and biomechanical modeling of voice production," *J. Biomech.* **41**, 985–995.

- Slifka, J. (2006). "Some physiological correlates to regular and irregular phonation at the end of an utterance," *J. Voice* **20**(2), 171–186.
- Story, B. H. (1995). "Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract," Ph.D. dissertation, University of Iowa, Chap. 2.
- Story, B. H., Titze, I. R., and Hoffman, E. A. (1996). "Vocal tract area functions from magnetic resonance imaging," *J. Acoust. Soc. Am.* **100**, 537–554.
- Titze, I. (2008). "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**, 2733–2749.
- Titze, I., and Talkin, D. (1979). "A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation," *J. Acoust. Soc. Am.* **66**, 60–74.
- Titze, I. R. (1995). *Workshop on Acoustic Voice Analysis: Summary Statement*, National Center for Voice and Speech, Iowa City, IA, 1–36.
- Tokuda, I. T., Horacek, J., Svec, J. G., and Herzel, H. (2007). "Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments," *J. Acoust. Soc. Am.* **122**, 519–531.
- Vahabzadeh-Hagh, A., Zhang, Z., and Chhetri, D. (2017). "Quantitative evaluation of the *in vivo* vocal fold medial surface shape," *J. Voice* **31**, 513.e15–513.e23.
- van den Berg, J. W., and Tan, T. S. (1959). "Results of experiments with human larynxes," *Pract. Otorhinolaryngol.* **21**, 425–450.
- Wade, L., Hanna, N., Smith, J., and Wolfe, J. (2017). "The role of vocal tract and subglottal resonances in producing vocal instabilities," *J. Acoust. Soc. Am.* **141**, 1546–1559.
- Wu, L., and Zhang, Z. (2016). "A parametric vocal fold model based on magnetic resonance imaging," *J. Acoust. Soc. Am.* **140**, EL159–EL165.
- Zanartu, M., Mehta, D. D., Ho, J. C., Wodicka, G. R., and Hillman, R. E. (2011). "Observation and analysis of *in vivo* vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study," *J. Acoust. Soc. Am.* **129**, 326–339.
- Zhang, Z. (2009). "Characteristics of phonation onset in a two-layer vocal fold model," *J. Acoust. Soc. Am.* **125**, 1091–1102.
- Zhang, Z. (2015). "Regulation of glottal closure and airflow in a three-dimensional phonation model: Implications for vocal intensity control," *J. Acoust. Soc. Am.* **137**, 898–910.
- Zhang, Z. (2016). "Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model," *J. Acoust. Soc. Am.* **139**, 1493–1507.
- Zhang, Z. (2017a). "Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model," *J. Acoust. Soc. Am.* **142**, 2311–2321.
- Zhang, Z. (2017b). "Toward real-time physically-based voice simulation: An Eigenmode-based approach," *Proc. Mtgs. Acoust.* **30**, 060002.
- Zhang, Z., and Luu, T. (2012). "Asymmetric vibration in a two-layer vocal fold model with left-right stiffness asymmetry: Experiment and simulation," *J. Acoust. Soc. Am.* **132**(3), 1626–1635.
- Zhang, Z., Mongeau, L., and Frankel, S. H. (2002). "Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes," *J. Acoust. Soc. Am.* **112**(4), 1652–1663.
- Zhang, Z., Neubauer, J., and Berry, D. A. (2006). "The influence of subglottal acoustics on laboratory models of phonation," *J. Acoust. Soc. Am.* **120**(3), 1558–1569.
- Zhang, Z., Samajder, H., and Long, J. (2017). "Biaxial mechanical properties of human vocal fold cover under vocal fold elongation," *J. Acoust. Soc. Am.* **142**, EL356–EL361.