

UCLA

UCLA Previously Published Works

Title

The influence of source-filter interaction on the voice source in a three-dimensional computational model of voice production.

Permalink

<https://escholarship.org/uc/item/2zv9m108>

Journal

Nature Energy, 154(4)

Author

Zhang, Zhaoyan

Publication Date


2023-10-01

DOI

10.1121/10.0021879

Peer reviewed

The influence of source-filter interaction on the voice source in a three-dimensional computational model of voice production

Zhaoyan Zhang^{a)} 

UCLA School of Medicine, 31-24 Rehabilitation Center, 1000 Veteran Avenue, Los Angeles, California 90095-1794, USA

ABSTRACT:

The goal of this computational study is to quantify global effects of vocal tract constriction at various locations (false vocal folds, aryepiglottic folds, pharynx, oral cavity, and lips) on the voice source across a large range of vocal fold conditions. The results showed that while inclusion of a uniform vocal tract had notable effects on the voice source, further constricting the vocal tract only had small effects except for conditions of extreme constriction, at which constrictions at any location along the vocal tract decreased the mean and peak-to-peak amplitude of the glottal flow waveform. Although narrowing in the epilarynx increased the normalized maximum flow declination rate, vocal tract constriction in general slightly reduced the source strength and high-frequency harmonic production at the glottis, except for a limited set of vocal fold conditions (e.g., soft, long vocal folds subject to relatively high pressure). This suggests that simultaneous laryngeal and vocal tract adjustments are required to maximize source-filter interaction. While vocal tract adjustments are often assumed to improve voice production, our results indicate that such improvements are mainly due to changes in vocal tract acoustic response rather than improved voice production at the glottis. © 2023 Acoustical Society of America. <https://doi.org/10.1121/10.0021879>

(Received 12 June 2023; revised 28 September 2023; accepted 30 September 2023; published online 19 October 2023)

[Editor: James F. Lynch]

Pages: 2462–2475

I. INTRODUCTION

In the source-filter theory of speech production, the voice source and vocal tract (filter) are assumed to act independently from each other. However, it has long been recognized that vocal tract adjustments may impact the voice source (e.g., [Ishizaka and Flanagan, 1972](#); [Rothenberg, 1981b,a](#); [Fant, 1982](#)). Such interaction can result from the fact that the vocal tract and larynx are anatomically connected and vocal tract adjustments may lead to changes in laryngeal configuration (e.g., tongue movement may affect larynx height). In this study, we focus on the source-filter interaction that results from acoustic and aerodynamic coupling between the voice source and vocal tract. While there have been many studies on this topic, most studies focused on conditions when the fundamental frequency (and sometimes also the second harmonic) is close to one of the vocal tract resonances ([Joliveau et al., 2004](#); [Titze, 2008](#); [Henrich et al., 2011](#); [Murtola et al., 2018](#); [Echternach et al., 2021](#)) or when vocal folds vibrate in a state near a bifurcation boundary ([Herzel, 1993](#); [Neubauer et al., 2001](#); [Tokuda et al., 2010](#); [Zañartu et al., 2011](#); [Zhang, 2018](#); [Herbst et al., 2023](#)). In this study, we aim to quantify the global effects of source-filter interaction on the voice source across a large range of vocal fold and vocal tract conditions that do not necessarily fall under the two special conditions mentioned above. Specifically, we focus on source-filter interaction in a

vocal tract constricted somewhere along the airway, as often occurs in singing (e.g., epilaryngeal narrowing) or vocal tract exercises in voice therapy (semi-occlusion at the lips).

Changes in vocal tract configuration impact the voice source both acoustically and aerodynamically. Acoustically, at frequencies below the first vocal tract resonant frequency, the vocal tract acts as an inertive load to the vocal folds, which facilitates establishing a favorable relationship between the intraglottal pressure and vocal fold vibration that is required to initiate and sustain vocal fold vibration ([Titze, 1988](#); [Zhang et al., 2006](#); [Zhang, 2016](#); [Zañartu et al., 2007](#)). This inertive effect is the strongest when the fundamental frequency of vocal fold vibration approaches one of the resonances (often the first resonance) of the vocal tract ([Ishizaka and Flanagan, 1972](#); [Zhang et al., 2006, 2009](#); [Titze, 2008](#)). Under such conditions, the vocal folds have the tendency to vibrate with a strong up-and-down, in-phase motion, at a frequency close to the vocal tract resonance frequency, often with a reduced phonation threshold pressure.

For speech and sometimes singing the fundamental frequency is often not adequately close to any of the subglottal or supraglottal resonances. Under such conditions, vocal tract inertance decreases rapidly as the fundamental frequency moves away from a vocal tract resonance, which reduces its impact on vocal fold vibration, as demonstrated in the experiments by [Zhang et al. \(2006, 2009\)](#). While source-filter interaction still has some effects on the voice source (e.g., the glottal flow waveform is skewed toward the

^{a)}Email: zyzhang@ucla.edu

closing phase with respect to the glottal area waveform; Rothenberg, 1981b; Fant, 1982), phonation frequency is no longer entrained to a vocal tract resonance, allowing relatively independent control of the voice source and articulation.

Recent studies showed that source-filter interaction may be enhanced by vocal tract constriction, even when the fundamental frequency is considerably lower than the first formant. For example, epilaryngeal narrowing has been shown to reduce the phonation threshold pressure (Titze and Story, 1997; Dollinger *et al.*, 2006; Kniesburges *et al.*, 2017), although the effect is not always consistent and the opposite has been observed (Montequin, 2003; Bailly *et al.*, 2008; Zhang, 2022b), and may also impact phonation frequency (Bailly *et al.*, 2008; Bailly *et al.*, 2014). Vocal tract constriction also increases aerodynamic coupling between the vocal tract and vocal folds. Extreme narrowing in the vocal tract causes considerable pressure drop across the location of vocal tract narrowing. This increases the supraglottal pressure and, for a given lung pressure, reduces the transglottal pressure, thus reducing both the vocal fold vibration amplitude and glottal flow amplitude (Bickley and Stevens, 1986; Titze, 2002; Dollinger *et al.*, 2006; Zhang, 2021). Titze (2006) argued that vocal tract constriction, either in the epilarynx or at the lips [as in semi-occluded vocal tract exercises (SOVTE)], improves impedance matching between the glottis and the vocal tract, which then improves vocal efficiency and vocal economy (ratio between measures of vocal output and vocal fold collision).

However, it remains unclear to what extent vocal tract constriction enhances source-filter interaction and impacts the voice source consistently across a large range of vocal conditions. Despite many previous studies, there have been few systematic, quantitative studies of how changes in vocal tract configurations affect the voice source. Due to limited access to the larynx, human subject studies often have to rely on inverse filtering to estimate the voice source (Holmberg *et al.*, 1988; Gauffin and Sundberg, 1989; Björkner *et al.*, 2006). Computational studies allow direct evaluation of the voice source, but most studies investigated only a small number of conditions.

More importantly, most previous studies focused on the effect of source-filter interaction on the phonation threshold pressure. Few studies investigated the effect of source-filter interaction on the glottal closure pattern and the acoustics and spectral characteristics of the voice source. As a result, when voice production (both efficiency and economy) improves after vocal tract exercises (e.g., after SOVTE), it is not always clear whether such improvements are due to improvement in the voice source as a result of source-filter interaction, or simply changes in the filter (e.g., improved sound amplification by the vocal tract at high frequencies or singer's formant; Sundberg, 1974), or active adjustments made by the speaker in the larynx or vocal tract in adaptation to the specific vocal exercises. Similarly, when singers modify timbre of their voice, it is often unclear to what degree source-filter interaction contributes to the observed changes in voice quality.

In this study, by systematically introducing constrictions along the vocal tract and performing voice production simulations, we aim to quantify changes in the voice source, including aerodynamics, glottal closure pattern, and acoustics, due to changes in vocal tract configuration. A quantitative understanding of how source-filter interaction impacts the voice source would provide insights into the therapeutic benefits of vocal exercises often used in voice therapy, and elucidate the nature of laryngeal and/or vocal tract adjustments that improve vocal efficiency and economy, and how much of such improvements can be carried over to a more natural vocal tract configuration as in speaking. Such a quantitative understanding of source-filter interaction would also allow us to develop parametric models of source-filter interaction, which is essential to many voice technology applications (e.g., inverse filtering, Alku, 1992; voice production inversion, Zhang, 2022a), particularly for voices produced with a relatively constricted vocal tract (e.g., epilaryngeal narrowing) as often in emotional speech and singing.

II. METHOD

A. Computational model

A three-dimensional model of voice production developed in our previous studies (Zhang, 2015, 2017, 2019) was used in this study. The model consists of a respiratory system, a three-dimensional vocal fold model, and a vocal tract. The vocal folds are modeled as a transversely isotropic, nearly incompressible, linear material with the plane of isotropy perpendicular to the anterior-posterior (AP) direction. A reduced-order formulation is used by projecting the governing equations of the vocal folds into the space spanned by the *in vacuo* eigenmodes of the vocal folds (Zhang, 2015), which significantly improves the computational efficiency. The glottal flow is modeled as a one-dimensional quasi-steady glottal flow model taking into consideration viscous loss up to the point of flow separation, with the flow separation point predicted by an ad hoc geometric model (Zhang, 2017). Vocal fold contact is modeled by applying a penalty pressure perpendicular to the vocal fold when the two vocal folds are in contact (Zhang, 2019). Despite simplifications made in our model to improve computational efficiency, our model has been shown to qualitatively and quantitatively reproduce observations from experiments and fully-resolved simulations (Zhang *et al.*, 2002; Zhang and Luu, 2012; Farahani and Zhang, 2016; Yoshinaga *et al.*, 2022). The general trends of voice production identified in our model have also been observed in other computational and experimental studies (e.g., Li *et al.*, 2018; Li *et al.*, 2020; Taylor and Thomson, 2022; McCollum *et al.*, 2023).

The vocal folds in our model (Fig. 1) are parameterized by five geometric measures and five mechanical control parameters. In this study, we systematically varied three geometric measures: the initial glottal angle α controlling the glottal gap in the horizontal plane, vertical thickness of the vocal fold medial surface T , and vocal fold length L

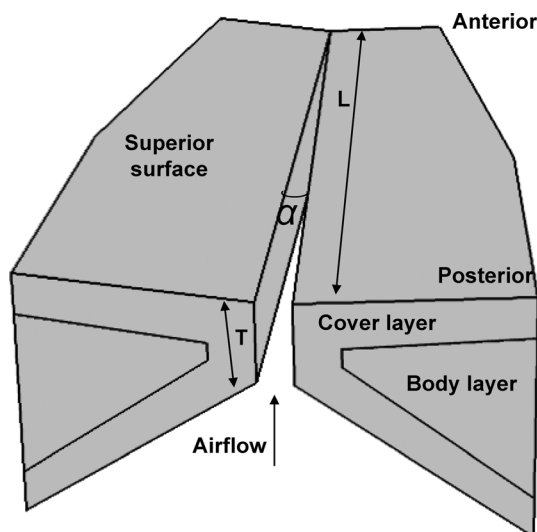


FIG. 1. The three-dimensional vocal fold model. The initial glottal angle α , medial surface vertical thickness T , and vocal fold length L , were systematically varied in the simulations.

along the anterior-posterior direction. These measures have been shown to have important effects on the voice source (Zhang, 2016, 2023). The specific ranges of variation for these three geometric measures are listed in Table I. These variations cover typical ranges of normal phonation and were shown in our previous studies to produce voices of different voice quality ranging from breathy, normal, and pressed voices as well as irregular vocal fold vibration (Zhang, 2018).

The mechanical control parameters of the model include the transverse Young's modulus E_t in the coronal plane, and the AP Young's modulus E_{ap} and the AP shear modulus G_{ap} in the body and cover layers. Our previous studies (Zhang, 2017) showed that the effects of vocal fold longitudinal stiffness on the glottal closure pattern and voice source acoustics are smaller than those of vocal fold geometric parameters, whereas the transverse stiffness has important impact on the glottal closure pattern (Zhang, 2017), excitement of irregular vocal fold vibration (Zhang, 2018), and vocal fold contact pressure (Zhang, 2019). In this study, the transverse stiffness was varied between 1 and 4 kPa (Table I). The vocal fold longitudinal stiffnesses were set to $G_{ap} = 10$ kPa and $E_{ap} = 40$ kPa, respectively, in both the body and cover layers. Despite these limited range of variations for the mechanical properties, parametric variations in the three geometric measures and transverse stiffness in this study were able to produce more than a 2-octave

TABLE I. Simulation conditions.

Initial glottal angle α ($^\circ$)	-1.6, 0, 1.6, 4, 8
Vertical thickness T (mm)	1, 2, 3, 4.5
Vocal fold length L (mm)	10, 17
Vocal fold transverse stiffness E_t (kPa)	1, 2, 4
Vocal tract constriction location VTC	VT1-VT6
Degree of constriction S	0.1, 0.2, 0.5, 1
Subglottal pressure P_s (kPa)	0.1-2 (16 steps)

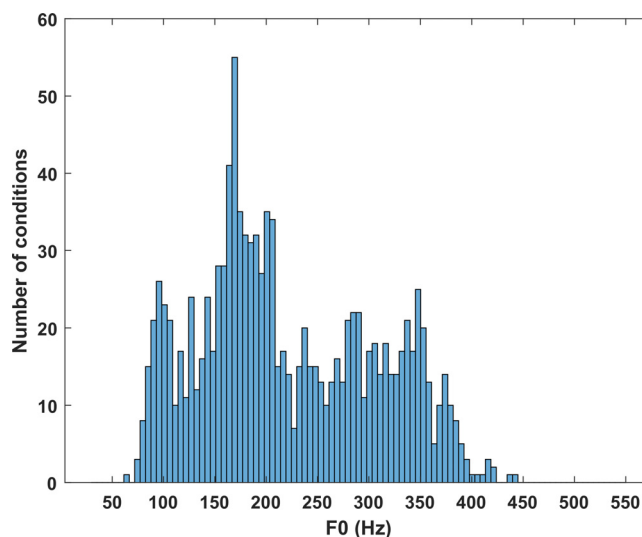


FIG. 2. (Color online) Histogram of the fundamental frequency (F0) in simulations without a vocal tract. The produced fundamental frequency spans the typical range of both male and female voices.

change in the fundamental frequency of vocal fold vibration (Fig. 2), an important parameter determining the strength of source-filter interaction. The effect of vocal fold stiffness on source-filter interaction will be further investigated in future studies.

B. Vocal tract configurations

In this study, two baseline vocal tract conditions and six series of parametric variations in vocal tract shape were considered. In the first baseline condition, voice simulations were performed without a vocal tract. If there were no source-filter interaction (i.e., the vocal system functions as a linear source-filter system), the voice source would remain unchanged no matter a vocal tract is absent or present, or whatever the vocal tract shape is. Thus, this first baseline condition simulates voice production without source-filter interaction, and any deviations in the voice source from this baseline condition are due to source-filter interaction. In the second baseline condition, a 17.1-cm long uniform vocal tract with a cross-sectional area of 2 cm^2 was included in voice simulations. This condition serves as a second reference condition, deviations from which would allow us to quantify changes in voice source due to source-filter interaction associated with further constricting an otherwise uniform vocal tract.

For the six additional series of parametric variations, a constriction of varying degree was introduced to an otherwise uniform vocal tract at six locations at which considerable vocal tract constriction is often observed (Story and Titze, 1998). These include constriction at the location of the false vocal folds (VT1), aryepiglottic folds (VT2), both locations of the false and aryepiglottic folds (VT3), pharynx (VT4), oral cavity (VT5), and lips (VT6), as shown in Fig. 3. For each of the six series, vocal tract constriction was realized by multiplying vocal tract area at the location of maximum constriction by a

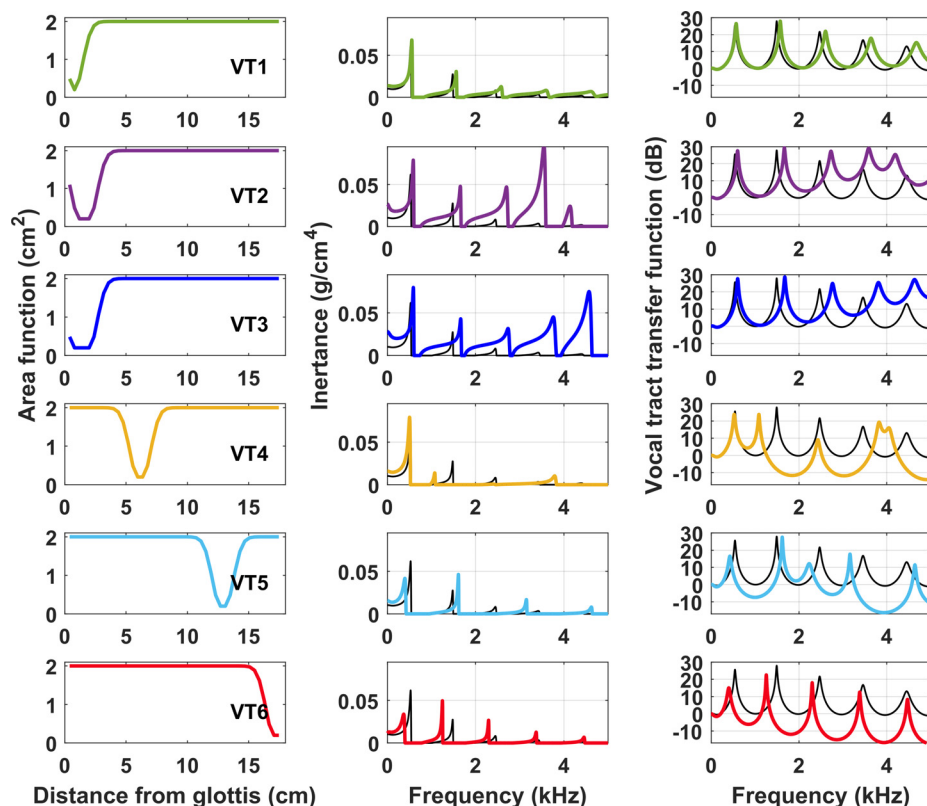


FIG. 3. (Color online) Six vocal tract configurations with constriction VT1-VT6. The corresponding area function (left), inertance (middle), and vocal tract transfer function (right) are shown from top (VT1) to bottom (VT6). For each vocal tract configuration, data are shown for the most constricted condition with the scaling factor $S=0.1$. Thin lines are data for the condition with a uniform vocal tract.

scaling factor. Four values of the scaling factor S were considered for each constriction location: 1, 0.5, 0.2, and 0.1, with the value $S=1$ corresponding to the uniform vocal tract condition and $S=0.1$ corresponding to the most constricted condition. To avoid abrupt changes in vocal tract area, a Gaussian function was applied to vocal tract segments immediately upstream and downstream of the constriction so that the vocal tract area smoothly transitioned from the minimal area back to the 2 cm^2 .

In this study, the vocal tract is modeled as a one-dimensional waveguide with a yielding vocal tract wall (Story, 1995; Zhang, 2022b). The effective mass, stiffness, and damping per unit area of the vocal tract wall were set to 16.3 kg/m^2 , 2187.0 kN/m^3 , and $13\,980\text{ Ns/m}^3$, respectively (Milenkovic and Mo, 1988). The vocal tract model also includes viscous loss and kinetic pressure loss, which are essential for modeling aerodynamic pressure drop across vocal tract constrictions. Our previous study (Zhang, 2022b) showed that the model was able to predict mean intraoral pressure values in a semi-occluded vocal tract that were comparable to human subjects experiments. Source-filter coupling is modeling by passing the glottal flow to the inlet of the vocal tract and imposing the pressure from the vocal tract at the glottal exit.

The effect of vocal tract constriction on vocal tract acoustics is evaluated by the vocal tract transfer function, or the ratio between the vocal tract output and input volume velocities. This was calculated from the acoustic response of the vocal tract to an impulse input at the vocal tract inlet. The impulse response also allows calculation of the vocal tract input impedance, from which the vocal tract input

inertance was calculated as the imaginary part of the input impedance divided by the angular frequency.

C. Simulation conditions and data analysis

The simulation conditions are listed in Table I. For each of the 25 vocal tract configurations, voice simulation of a half-second long sustained phonation was performed for each of the 120 vocal fold conditions and each of the 16 subglottal pressures. In total 48 000 conditions were simulated.

As discussed earlier, the effect of source-filter interaction on the voice source was quantified in this study by changes in selected voice source measures in a specific vocal tract configuration with respect to the corresponding reference condition without a vocal tract, which simulates voice production without source-filter interaction. In this study, we considered measures of vocal fold vibration, glottal flow, and source spectral characteristics. For vocal fold vibration, the mean glottal area (A_g0), peak-to-peak amplitude of the glottal area waveform (A_{gtamp}), closed quotient (CQ), maximum glottal area declination rate (MADR), and normalized MADRn (MADR normalized by $\pi F_0 A_{gtamp}$, where F_0 is the fundamental frequency of vibration) were extracted from the glottal area waveform. Note that by definition $MADRn = 1$ for a sinusoidal waveform and increases as the waveform becomes more skewed to the right. Similar measures were extracted for glottal aerodynamics, including the mean glottal flow (Q_{mean}), peak-to-peak amplitude of the glottal flow waveform (Q_{amp}), maximum flow declination rate (MFDR), and normalized MFDRn (MFDR normalized by $\pi F_0 Q_{amp}$). The peak vocal fold contact pressure

(P_{contact}) over the vocal fold medial surface was also extracted for each condition, as described in Zhang (2019). From the voice source term (the time derivative of the glottal flow waveform), we calculated the A-weighted sound pressure level at the glottis (SPL_{glottis}), the amplitude differences between the first harmonic and the second harmonic ($H1^*-H2^*$), the harmonic nearest 2 kHz ($H1^*-H2k^*$), and the harmonic nearest 5 kHz ($H1^*-H5k^*$), and the energy difference between the harmonics above 1 kHz and harmonics below 1 kHz ($\alpha 1k^*$), where the asterisks indicate that these measures were evaluated at the voice source. The fundamental frequency (F_0) was also extracted from the glottal flow. Changes in these measures due to source-filter interaction are denoted by the symbol Δ , which were calculated as the differences between a specific vocal tract condition and the corresponding condition without a vocal tract. For example, $\Delta MFDR$ denotes changes in the MFDR for a specific vocal tract configuration with respect to the condition without a vocal tract.

Analysis of variance (ANOVA) was performed to investigate the dependence of changes in voice source measures due to source-filter interaction on the laryngeal and vocal tract control parameters. The independent variables were the control parameters that were systematically varied in this study, as listed in Table I. Our initial analysis included both the main effects and two-way interactions. However, for some of the two-way interactions, sustained phonation was not achieved at all factorial combinations (Zhang, 2017). Also, different interaction terms reached significance for different voice source measures. Considering that the main effects of the seven control parameters on the 16 outcome measures are already complex enough, and to facilitate comparison across different voice source measures, the final ANOVA models included only the main effects and a two-way interaction between the vocal tract constriction location and degree of constriction ($S \times VTC$). Other interaction effects will be explored in future studies. Effect sizes of each control parameters on specific voice source measures were calculated as the percentage of total variance in the source measures that was explained by the individual control parameters. Multiple comparison was further performed to identify the trends of variation of the source measures with the control parameters.

III. RESULTS

A. Vocal tract acoustics

Figure 3 shows the vocal tract area function, vocal tract input inertance, and vocal tract transfer function for the most constricted condition (i.e., $S = 0.1$) in each of the six vocal tract constriction configurations. For comparison, each panel of the figure also shows the data for the condition with a uniform vocal tract (i.e., $S = 1$; thin black line). At low frequencies, constrictions in the epilarynx (VT1–VT3) increased the first formant, whereas the opposite was observed for constrictions in the oral cavity and at the lips (VT5–VT6). All of them increased the input inertance for

frequencies well below the first formant. At high frequencies, constrictions in the epilarynx, particularly at the level of the aryepiglottic folds, significantly increased the inertance in the high frequency range between 2 and 4 kHz, whereas the effects of constrictions in the pharynx, oral cavity, and at the lips were much smaller. As expected, the vocal tract transfer function had a significant boost in the 2–4 kHz range with constrictions in the aryepiglottic region, and was weakened by constrictions in the pharynx, oral cavity, or at the lips.

B. Laryngeal control of source-filter interaction

Figure 4 shows the effects sizes of the model control parameters on changes in voice source measures due to source-filter interaction. In general, the effect sizes of the subglottal pressure and laryngeal controls were larger than the vocal tract controls, particularly for changes in measures of vocal fold vibration and source spectral measures. Thus, while manipulation of vocal tract configuration is able to impact the voice source, the specific impact and its magnitude depend heavily on the laryngeal and respiratory configurations.

Figure 5 shows the mean changes in selected source measures due to source-filter interaction as estimated from the ANOVA analysis, as a function of the subglottal pressure and four laryngeal controls. The ordinate of each panel lists different levels of the subglottal pressure or specific laryngeal controls (also listed in Table I). The abscissa shows the estimated mean changes in selected measures

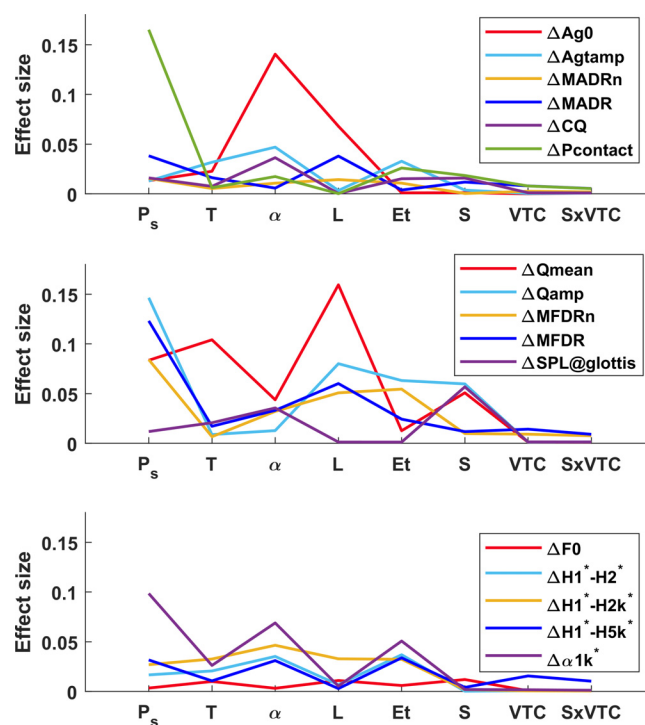


FIG. 4. (Color online) The effect sizes of the model control parameters on selected voice source measures. The control parameters include the subglottal pressure P_s , vocal fold vertical thickness T , initial glottal angle α , vocal fold length L , transverse stiffness E_t , vocal tract constricting factor S , location of vocal tract constriction VTC , and two-way interaction between S and VTC ($S \times VTC$).

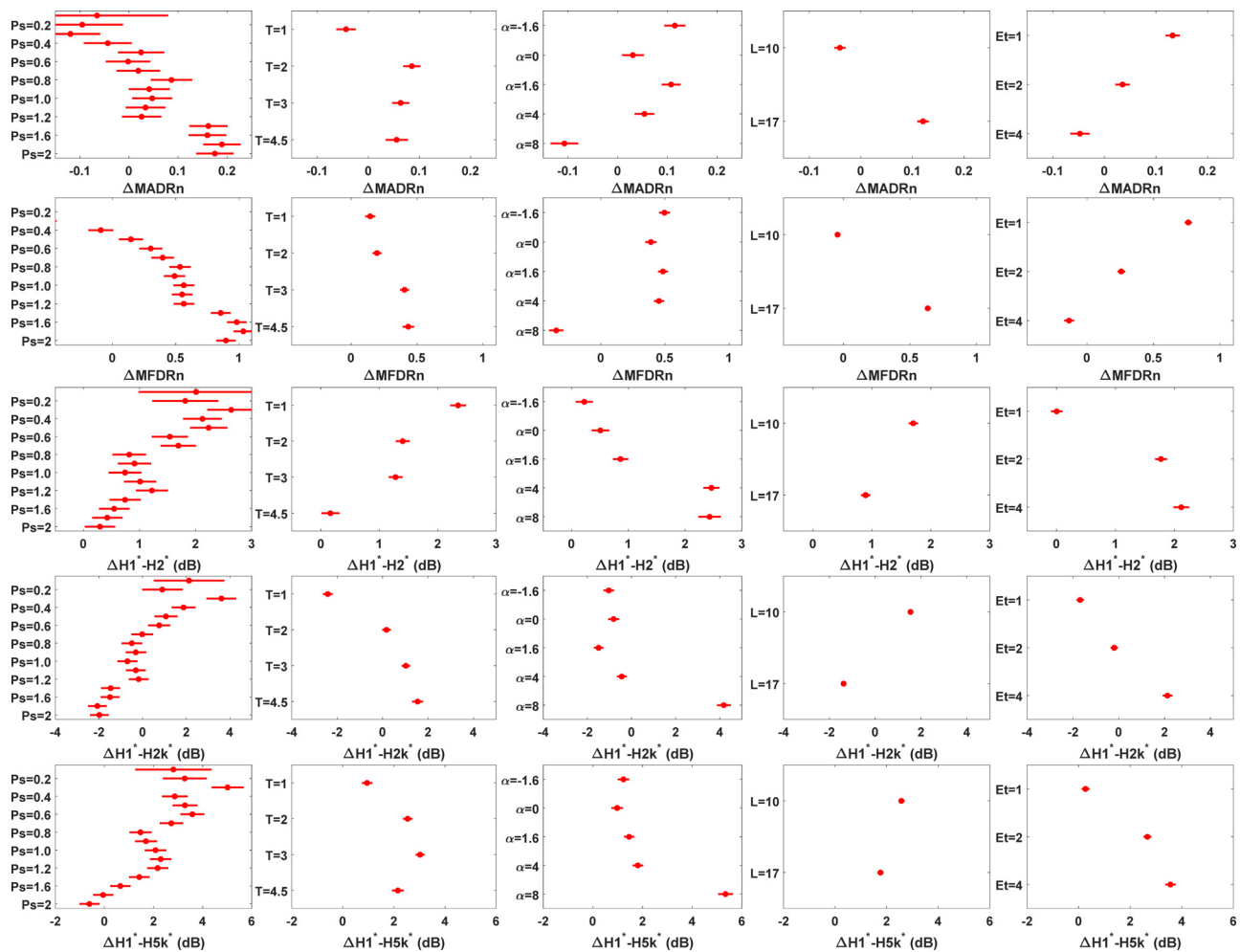


FIG. 5. (Color online) Laryngeal control of source-filter interaction. ANOVA-estimated mean changes (solid circles) in selected voice source measures with respect to the baseline condition without a vocal tract as a function of the subglottal pressure and four laryngeal controls. P_s : subglottal pressure (kPa); T : thickness (mm); α : initial glottal angle ($^\circ$); L : vocal fold length (mm); E_t : transverse vocal fold stiffness (kPa). The horizontal bars indicate the comparison intervals, with the interval widths calculated in a way so that the averaged changes are statistically significantly different ($p < 0.005$ with Bonferroni correction) when two conditions have non-overlapping bars.

(solid circles) at the specific levels. The horizontal bars are comparison intervals, with the interval widths calculated in a way so that the averaged changes are statistically significantly different ($p < 0.005$ with Bonferroni correction) when two conditions have non-overlapping bars (Hochberg and Tamhane, 1987).

In general, $\Delta MFDRn$ and $\Delta MADRn$ showed similar trends of variation with the subglottal pressure and laryngeal controls, although the magnitude of changes in $MADRn$ was about five times smaller than that in $MFDRn$. This indicates a much smaller impact of source-filter interaction on vocal fold vibration than the glottal flow. Note that both $MADRn$ and $MFDRn$ were normalized so that a sinusoidal waveform has a value of 1. Thus, a $\Delta MADRn$ on the order of 0.1 indicates a relatively small change in the normalized glottal closing speed. Although not shown in Fig. 5, a similarly small effect of source-filter interaction was observed for the closed quotient (CQ), with the estimated mean values of ΔCQ ranging from -0.04 to 0 , indicating a slight decrease in CQ with source-filter interaction.

The $\Delta MFDRn$ was in general significantly higher than zero, indicating that source-filter interaction generally increased MFDR. However, the magnitude of increase varied significantly with the specific laryngeal and respiratory conditions. The increase in $MFDRn$ due to source-filter interaction was the largest for vocal folds that were soft, long, thick, and sufficiently adducted when subject to high subglottal pressure. Note that these conditions also produced vocal fold vibration with complete glottal closure. The increase in $MFDRn$ was the smallest and even became negative for stiff, short, thin vocal folds minimally adducted at low subglottal pressures. Two examples are shown in Fig. 6. This observation is consistent with the theoretical findings in Rothenberg (1981a), which showed that source-filter interaction increased MFDR for vocal fold vibration with complete glottal closure, but the increase was much reduced in the presence of glottal leakage.

For most conditions, the estimated means of $\Delta H1^*-H2^*$ were significantly larger than zero, indicating that source-filter interaction increased $H1^*-H2^*$, which is

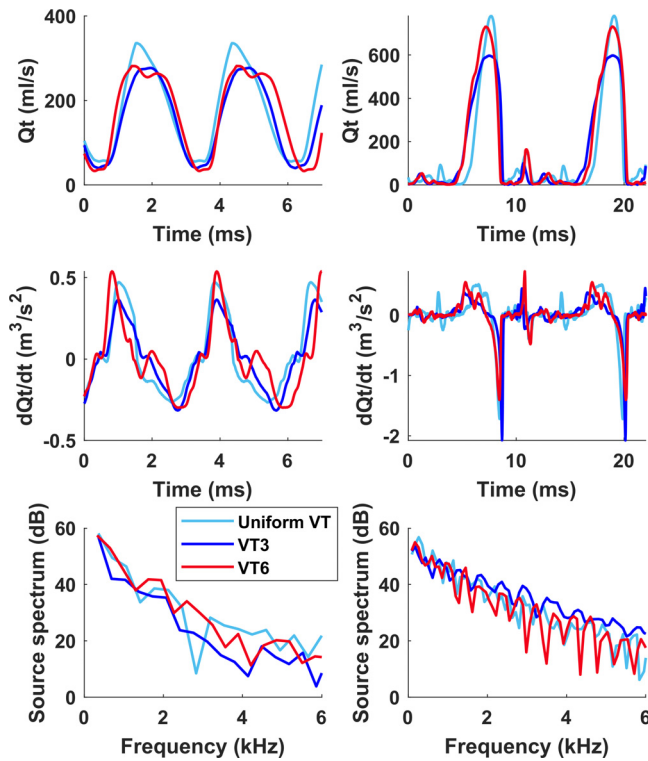


FIG. 6. (Color online) Effect of source-filter interaction on MFDR and high-frequency harmonic production varies depending on laryngeal configurations. Left: for vocal folds that are stiff and short ($E_f=4\text{ kPa}$; $L=10\text{ mm}$; $T=1\text{ mm}$; $\alpha=1.6^\circ$), epilaryngeal narrowing (VT3) only slightly increases MFDR with a notable reduction in high-frequency harmonic production. Right: for vocal folds that are soft and long ($E_f=1\text{ kPa}$; $L=17\text{ mm}$; $T=4.5\text{ mm}$; $\alpha=1.6^\circ$), epilaryngeal narrowing reduces the peak-to-peak amplitude of the glottal flow, but still significantly increases MFDR with slightly increased high-frequency harmonic production. In both conditions, constriction in the front part of the vocal tract (VT6) reduces the peak-to-peak amplitude but has only small effects on high-frequency harmonic production.

consistent with its decreasing effect on the CQ. The estimated mean values of $\Delta H1^* \text{--} H5k^*$ were also statistically significantly larger than zero, indicating that source-filter interaction in general reduced harmonic production in the frequency range around 5 kHz. This is expected considering that an inertive vocal tract tends to resist fast changes in the glottal flow and thus functions as a low-pass filter, particularly for conditions without complete glottal closure (Fig. 6, left panel). The trends of variation for $\Delta H1^* \text{--} H2k^*$ were similar to those for $\Delta H1^* \text{--} H5k^*$, although $\Delta H1^* \text{--} H2k^*$ became negative for conditions with large values of $\Delta MFDR_n$. Thus, while source-filter interaction generally reduced high-frequency harmonic production, at some laryngeal conditions (e.g., soft, long vocal folds tightly adducted) it may increase harmonic production at the mid- or even high-frequency range around 2–5 kHz, due to an increased rate of flow declination.

C. Vocal tract control of source-filter interaction

Compared to the subglottal pressure and laryngeal controls, the effect of vocal tract configuration was smaller.

This is particularly the case for measures of vocal fold vibration (top row in Fig. 4) and source acoustics (bottom row in Fig. 4). In comparison, vocal tract constriction had a relatively larger effects on aerodynamic measures, including the mean glottal flow (Q_{mean}), glottal flow amplitude (Q_{amp}), and SPL_{glottis} .

Figure 7 shows the averaged changes in selected voice source measures due to source-filter interaction at different vocal tract configurations. The ordinate of each panel lists the six configurations of vocal tract constriction (“VT1,” “VT2,” etc.; also see Fig. 3), each including four degrees of constriction, with $S=0.1$ corresponding to the most constricted condition for each vocal tract constriction location and $S=1$ corresponding to the uniform vocal tract configuration. The abscissa shows the mean changes in voice source measures (solid circles) for the specific vocal tract configuration as estimated from the ANOVA analysis. Again, the horizontal bars are comparison intervals, with the interval widths calculated in a way so that the averaged changes are statistically significantly different ($p < 0.005$ with Bonferroni correction) when two conditions have non-overlapping bars. Note that Fig. 7 shows two effects on the voice source: the effects of the inclusion of a uniform vocal tract are illustrated by conditions with $S=1$ in Fig. 7, whereas the effects of further constricting an otherwise uniform vocal tract can be quantified by comparing changes in voice measures of a specific condition with respect to the condition of a uniform vocal tract ($S=1$).

1. Vibratory effects of vocal tract constriction

Figure 7 shows that inclusion of a uniform vocal tract (conditions with $S=1$) slightly increased the mean (Ag_0) and peak-to-peak amplitude (Ag_{tamp}) of glottal opening area, slightly reduced the closed quotient (CQ), and increased peak vocal fold contact pressure (P_{contact}), with changes in these measures statistically significantly different from zero. Further introducing constrictions to an otherwise uniform vocal tract had negligible additional effect on the mean glottal area Ag_0 , but decreased the peak-to-peak amplitude of the glottal opening area Ag_{tamp} for conditions of extreme constriction ($S=0.1$). This differential effect of vocal tract constriction on the mean glottal area and glottal area amplitude may have contributed to the observed further decrease in the closed quotient with increasing degree of vocal tract constriction.

Constrictions in the epilarynx slightly increased the normalized maximum closing speed of the glottal area waveform $MADR_n$, whereas constrictions in other locations of the vocal tract generally decreased $MADR_n$. However, these changes were mostly statistically insignificant when compared to the condition with a uniform vocal tract. Epilaryngeal narrowing first increased then decreased the peak vocal fold contact pressure, although this effect was only borderline statistically significant. In contrast, extreme constrictions ($S \leq 0.2$) in the pharynx, oral cavity, and at the lips had a statistically significant effect of decreasing the peak vocal fold contact pressure.

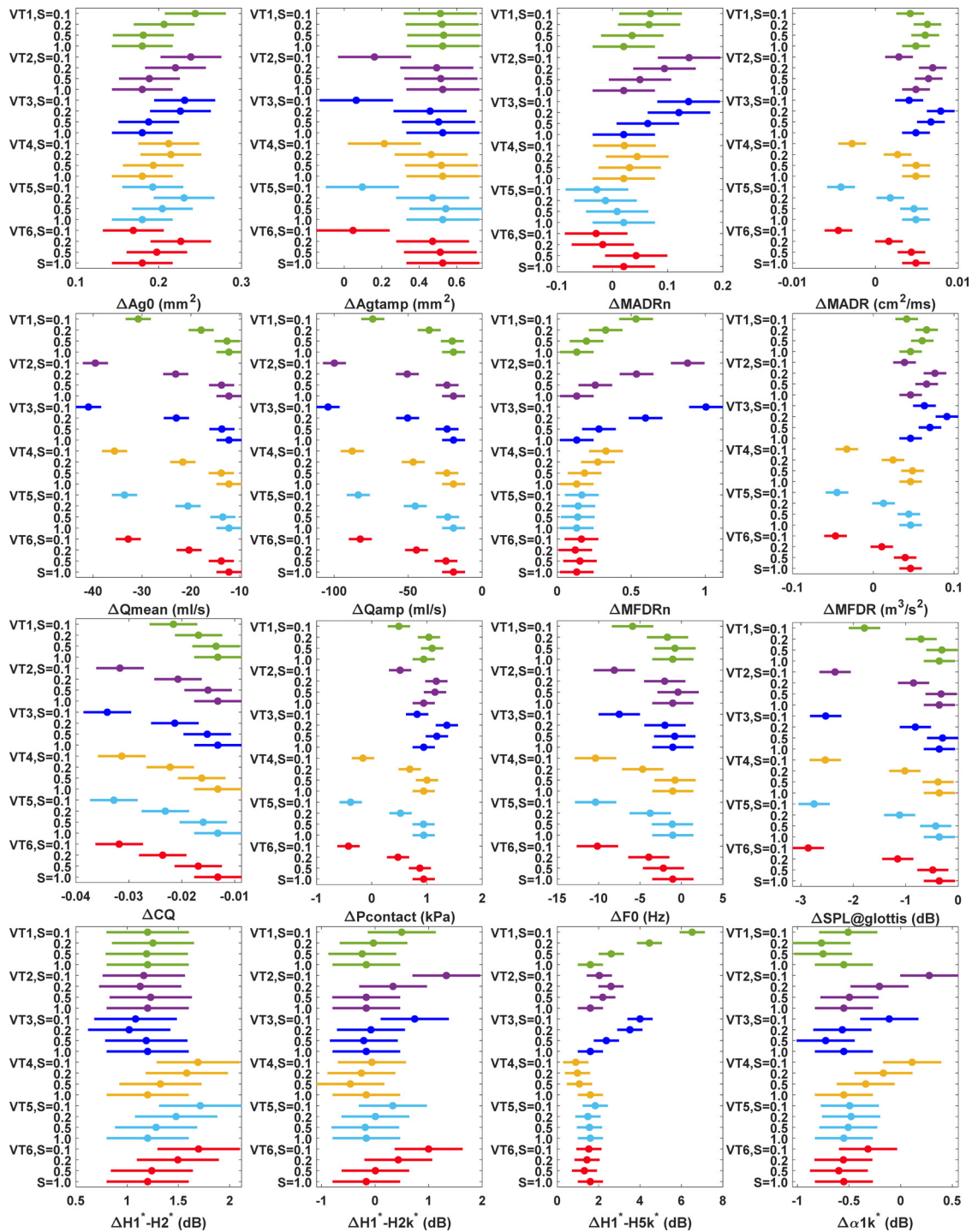


FIG. 7. (Color online) Vocal tract control of source-filter interaction. ANOVA-estimated mean changes (solid circles) in selected voice source measures with respect to the baseline condition without a vocal tract as a function of vocal tract constriction location *VT* and scaling factor *S*. The horizontal bars indicate the comparison intervals, with the interval widths calculated in a way so that the averaged changes are statistically significantly different ($p < 0.005$ with Bonferroni correction) when two conditions have non-overlapping bars.

Overall, the effects of source-filter interaction on vibratory measures were small and statistically insignificant except for the most constricted conditions, especially when compared to the conditions with a uniform vocal tract, consistent with Fig. 4. The averaged changes in the mean glottal opening area (ΔAg_0) due to source-filter

interaction were about 0.2 mm^2 . The averaged changes in the glottal area amplitude (ΔAg_{tamp}) were less than 0.6 mm^2 . Changes in the closed quotient (ΔCQ) varied between -0.01 and -0.04 , and changes in the normalized maximum glottal closing speed $MADR_n$ varied between -0.1 and 0.15 .

2. Aerodynamic effects of vocal tract constriction

In comparison, the effects of source-filter interaction on the glottal flow were relatively larger (second row, Fig. 7). On the one hand, both the mean and amplitude of the glottal flow decreased with increasing degree of vocal tract constriction. On the other hand, the normalized maximum flow declination rate in the closing phase MFDR_n increased with the inclusion of a uniform vocal tract and increased even more with further constriction in the epilarynx, although it did not change much (statistically insignificant) with constrictions in the pharynx, oral cavity, or at the lips. This indicates faster flow declination in the closing phase with the inclusion of a uniform vocal tract and additional epilaryngeal constriction, consistent with findings from previous studies (Rothenberg, 1981b; Fant, 1982).

The MFDR increased significantly with the inclusion of a uniform vocal tract. With additional epilaryngeal narrowing, MFDR first increased then decreased with increasing degree of constriction. In contrast, MFDR decreased monotonically with increasing degree of constriction in the pharynx, oral cavity, or at the lips. The non-monotonic trends of variation in MFDR with epilaryngeal narrowing were likely due to the opposite effects of source-filter interaction on the glottal flow amplitude Q_{amp} and MFDR_n, with the former gradually outweighing the latter as the vocal tract became increasingly constricted in the epilarynx. Note that changes in both the MADR and peak vocal fold contact pressure (P_{contact}) exhibited a similar pattern as ΔMFDR, indicating similar competing effects in play for the control of the MADR and peak contact pressure.

In summary, Fig. 7 shows that source-filter interaction in general had a smaller effect on the glottal area waveform than on the glottal flow waveform, consistent with the observation in Figs. 4 and 5. Figure 7 also shows that epilaryngeal constrictions impact the voice source differently from constrictions at other parts of the vocal tract. While constrictions at all locations decreased the glottal flow amplitude, only epilaryngeal constriction had notable effects on the normalized closing speed of the glottal flow (MFDR_n), which was minimally influenced by constrictions in the pharynx, oral cavity, or the lips.

3. Acoustic effects of vocal tract constriction

Figure 7 also shows the estimated mean changes in selected acoustic measures due to source-filter interaction. Changes in the voice source acoustic measures due to source-filter interaction were small in general, consistent with the observation in Figs. 4 and 5. The estimated mean changes were on the order of 2 dB for H1*–H2* and H1*–H2k*, and 1–7 dB for H1*–H5k*. These changes are smaller than their respective just noticeable differences (Garellek *et al.*, 2016), indicating minimal perceptual relevance. Changes in α1k* due to source-filter interaction were also on the order of 2 dB and mostly negative (i.e., reduced high-frequency harmonics). In general, changes due to the inclusion of a uniform vocal tract were statistically

significant, whereas changes due to vocal tract constriction were often statistically insignificant except for ΔH1*–H5k* with epilaryngeal constriction. These general trends of variations indicate slightly reduced harmonic production at mid- and high frequencies, consistent with the observations in Fig. 5.

Source-filter interaction also had an effect of decreasing the fundamental frequency, although this effect was statistically significant only for the most constricted vocal tract conditions, with the average decrease in F0 as large as about 10 Hz.

The estimated mean changes in the A-weighted sound pressure level evaluated at the glottis were consistently smaller than zero, and decreased further with increasing vocal tract constriction. This suggests that source-filter interaction in general reduced the source strength.

Since changes in vocal tract configuration significantly modify the formant structure of the vocal tract (Fig. 3), they are expected to have a large impact on the output acoustics. Our results showed that compared to a uniform vocal tract, epilaryngeal narrowing led to an increased sound pressure level (SPL) outside the mouth despite a reduced SPL_{glottis} at the glottis, whereas constriction in the front part of the vocal tract decreased the SPL outside the mouth. Similarly, epilaryngeal constriction increased high-order harmonic energy outside the mouth, whereas constrictions in other parts of the vocal tract decreased high-order harmonic energy outside the mouth.

D. Local effects of source-filter interaction on the voice source

The results above showed that epilaryngeal narrowing increased the rate of flow declination in the closing phase yet still reduced high-frequency harmonic production, particularly around 5 kHz. This is likely because that the increase in MFDR_n was not large enough to overcome the low-pass effect of an inertive vocal tract. It is possible that, for a small set of laryngeal and respiratory conditions (e.g., the condition shown on the right panel of Fig. 6), the increase in MFDR_n may be large enough to overcome the overall low-pass effect of an inertive vocal tract, resulting in an increased harmonic production at high frequencies around 2 kHz or even 5 kHz.

To test this hypothesis, we repeated the same analysis for a subset of data that satisfied the conditions $E_t = 1 \text{ kPa}$, $P_s \geq 800 \text{ Pa}$, and $\alpha \leq 4^\circ$. These conditions corresponded to those that vibrated with a considerable duration of glottal closure and for which source-filter interaction led to notable increase in MFDR_n, as shown in Fig. 5. We hypothesized that for this subset of condition, the relatively large increase in MFDR_n would result in increased high-order harmonic production, at least in the middle-frequency range around 2 kHz. This is confirmed in Fig. 8, which shows the estimated mean changes in the voice source measures for this subset of data as a function of the vocal tract configuration. While the general trends of variations in Fig. 8 are similar to those in Fig. 7 for most measures, notable differences can be

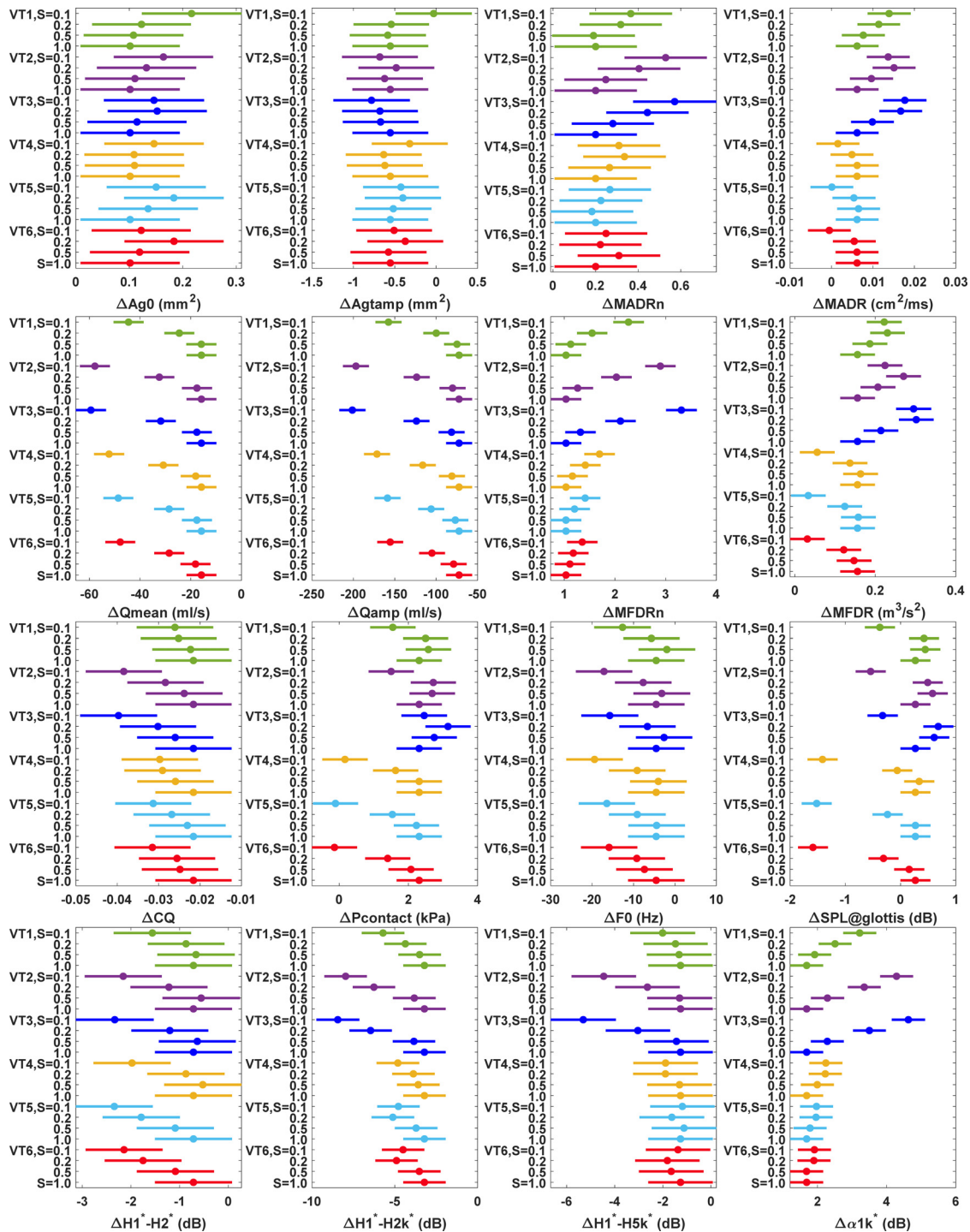


FIG. 8. (Color online) Local effect of source-filter interaction on the voice source for conditions with $E_t = 1$ kPa, $P_s \geq 800$ Pa, and $\alpha \leq 4^\circ$. ANOVA-estimated mean changes (solid circles) in selected source measures with respect to the baseline condition without a vocal tract as a function of vocal tract constriction location VT and scaling factor S . The horizontal bars indicate the comparison intervals, with the interval widths calculated in a way so that the averaged changes are statistically significantly different ($p < 0.005$ with Bonferroni correction) when two conditions have non-overlapping bars.

observed. First, changes in MFDRn due to source-filter interaction were much larger than those estimated from the entire dataset in Fig. 7. This significantly larger $\Delta MFDRn$ was able to overcome the low-pass effect of an inertive vocal tract, and reduce $\Delta H1^*-H2k^*$ and $\Delta H1^*-H5k^*$. As a result, unlike Fig. 7 in which epilaryngeal narrowing increased both $\Delta H1^*-H2k^*$ and $\Delta H1^*-H5k^*$, for this subset

of data epilaryngeal narrowing decreased both, indicating increased high-frequency harmonic production, and was able to even slightly increase the source strength $SPL_{glottis}$ (except the most constricted conditions). Note that, however, even in this subset of conditions, the magnitudes of changes in the $H1^*-H2k^*$ and $H1^*-H5k^*$ due to source-filter interaction are still smaller than or at most comparable

to the just noticeable differences measured in Garellek *et al.* (2016). It is worth noting that such reversal in the trends of $\Delta H1^*$ -H2k* and $\Delta H1^*$ -H5k* was not observed for a subset of data satisfying only one or two of the three conditions (i.e., $E_t = 1$ kPa, $P_s \geq 800$ Pa, and $\alpha \leq 4^\circ$), particularly for $\Delta H1^*$ -H5k*. This indicates that the effect of MFDRn increase on source spectra is mostly limited to low- and mid-frequencies.

E. Comparison to previous studies

Our findings are largely consistent with previous studies. The effect of source-filter interaction on MFDRn has been well documented (Rothenberg, 1981b,a; Fant, 1982; Titze and Palaparthi, 2016). The negative effects of extreme epilaryngeal narrowing on MFDR were observed in Samlan and Kreiman (2014) and Zhang (2021). Rothenberg (1981a) showed that an inertive vocal tract increases MFDR only in the absence of any glottal leakage, which is consistent with the observation in our study. The effect of vocal tract constriction on the mean glottal flow and flow amplitude was reported in both simulations (Titze, 2002; Samlan and Kreiman, 2014) and experiments (Dollinger *et al.*, 2006; Dollinger *et al.*, 2012). The experiments by Dollinger and colleagues also showed small effects of epilaryngeal narrowing on vocal fold vibration other than reduced vibration amplitude at the extreme narrowing condition, consistent with the observation in our study. Their experiments also showed that extreme epilarynx narrowing reduced the output SPL and F0, as observed in our study. The small effect of source-filter interaction on the voice source spectral shape was consistent with the findings in Sundberg *et al.* (2013) and Sundberg (2017). The findings of this study are also consistent with previous studies on voiced consonants and vocal tract constrictions in humans (e.g., Bickley and Stevens, 1986; Mittal *et al.*, 2014; Chong *et al.*, 2020), which showed reduced contact quotient and strength of excitation of the voice source as the degree of constriction increased.

IV. DISCUSSION AND CONCLUSIONS

Table II summarizes the statistically significant global effects of source-filter interaction on the voice source. While some notable effects can be observed with the inclusion of a uniform vocal tract, three major effects were observed with further constriction in the vocal tract. First, constrictions at any location along the vocal tract decreased the mean and peak-to-peak amplitude of the glottal flow waveform. Second, constriction in the epilarynx increased the normalized maximum flow declination rate MFDRn, which was only minimally affected by constrictions in the pharynx, oral cavity, or at the lips. However, the increase in MFDRn was outweighed by the decreased in the glottal flow amplitude at conditions of extreme vocal tract constriction. As a result, the source strength, evaluated by the sound pressure level at the glottis, generally decreased with increasing vocal tract constriction. Lastly, the increase in MFDRn due

TABLE II. Summary of statistically significant global effects of source-filter interaction on the voice source.

		Uniform vocal tract	Epilaryngeal narrowing	Narrowing in pharynx, oral cavity, and lips
Vibration	Ag0	↑		
	Agtamp	↑	↓	↓
	MADRn			
	CQ	↓	↓	↓
	Pcontact	↑		↓
Glottal flow	Qmean	↓	↓	↓
	Qamp	↓	↓	↓
	MFDRn	↑	↑	
Acoustics	SPL@glottis	↓	↓	↓
	High-order harmonics	↓	↓	

to epilaryngeal narrowing did not have much effect on high-frequency harmonic production, either, which decreased slightly with increasing vocal tract constriction, except for a limited set of laryngeal and respiratory conditions. In general, the mean changes in spectral shape measures of the voice source were smaller than their just noticeable differences, indicating minimal perceptual relevance.

In general, the effects of vocal tract constriction on the glottal area waveform, particularly MADRn, were smaller than similar measures of the glottal flow waveform. This is consistent with the large density ratio between the vocal folds and air and the small increase in the vocal tract input inductance with vocal tract constriction, which is generally small at frequencies well below the first vocal tract resonance (Fig. 3). However, due to the highly nonlinear nature of vocal fold contact mechanics, subtle changes in vocal fold vibration can lead to large changes in the peak vocal fold contact pressure, as observed in our study. It is interesting to note that the largest increase in the peak vocal fold contact pressure occurred with the inclusion of a uniform vocal tract, whereas further constricting the vocal tract had a much smaller effect on the peak contact pressure except at conditions of extreme constriction in the pharynx, oral cavity, and at the lips, which decreased the peak contact pressure.

It is generally assumed that source-filter interaction improves vocal efficiency through skewing the flow waveform to the right and increasing MFDR. However, our results showed that the improved vocal efficiency associated with vocal tract adjustments was mainly due to their effects on the formant structure rather than their effects on the voice source. In fact, constricting the vocal tract in this study slightly reduced both the SPL at the voice source and high-frequency harmonic production. Thus, source-filter interaction reduced the magnitude of improvement that would have been achieved had there been no source-filter interaction. Similar observations can be found in Sundberg (2017), which showed reduced source amplitude in the presence of source-filter interaction.

An interesting finding of this study is that the impact of source-filter interaction on the voice source depends more

on the laryngeal and respiratory configurations than the vocal tract configuration. Thus, while constricting the vocal tract increases the acoustic inertance of the vocal tract, simultaneous laryngeal and respiratory adjustments are needed to maximize source-filter interaction. Specifically, our results showed that vocal folds that are soft, long, and sufficiently adducted when subject to high subglottal pressure produced the largest improvement in MFDR and high-frequency harmonic production. In humans, such laryngeal adjustments can be achieved by the activation of the thyroarytenoid muscles, particularly at lower registers. At high registers, other mechanisms of source-filter interaction (e.g., formant tuning; Joliveau *et al.*, 2004) may become more effective.

Many voice therapy approaches often target enhancing source-filter interaction to improve vocal efficiency and vocal economy. For example, Titze (2006) showed that epilaryngeal narrowing in the epilarynx or at the lips increased the ratio between MFDR and MADR, an indirect measure of vocal economy. Titze also argued that vocal tract constriction raises the back pressure above the glottis and intraglottal pressure, which “tend to keep the vocal folds separated.” These changes would then minimize vocal fold collision while maximizing MFDR and vocal intensity. In our study, while we did observe an increase in the ratio between MFDR and MADR with epilaryngeal constriction, this increase was mainly due to an increase in MFDR. However, this MFDR did not translate to an increase in the SPL at the voice source. On the other hand, epilaryngeal constriction in our study did not lead to noticeable separation of the vocal folds nor a decrease in the peak vocal fold contact pressure. In fact, the largest changes in the mean glottal opening area and peak vocal fold contact pressure occurred with the inclusion of a uniform vocal tract, whereas only small (for the mean glottal opening area) or inconsistent (for peak contact pressure) changes were observed with further epilaryngeal constriction. Thus, the benefit of epilaryngeal narrowing toward reducing risk of vocal fold injury is achieved not by minimizing vocal fold collision within the glottis, but mostly through improvement in vocal efficiency, which again is achieved through modification of the formant structure of the vocal tract rather than strengthening the voice source. Specifically, epilaryngeal narrowing allows speakers to produce the same target output SPL with reduced subglottal pressure, the primary parameter controlling vocal fold contact pressure, thus reducing the peak vocal fold contact pressure, as demonstrated in our previous studies (Zhang, 2020, 2021).

Semi-occluded vocal tract exercises have been widely used in voice therapy to improve voice production and vocal efficiency. In this study, while constrictions at the lips did significantly reduce the peak vocal fold contact pressure, they also significantly reduced the output SPL. Lip constrictions also had minimal impact on the glottal area and glottal flow waveforms, except reducing their peak-to-peak amplitudes at conditions of extreme lip constriction. Thus, the therapeutic benefit of semi-occluded vocal tract exercises likely is not to directly improve vocal efficiency or economy. From a physical point of view, the main benefit of semi-occluded

vocal tract exercises is that it allows speakers to explore different vocal fold configurations without risking high vocal fold contact pressure or vocal fold injury, as suggested by Titze (2006). Vocal fold exercises without high contact pressure may also attenuate vocal fold inflammation and promote vocal fold wound healing (Verdolini-Abbott *et al.*, 2012). On the other hand, semi-occlusion at the lips also significantly increased both the mean and dynamic pressure inside the oral cavity (Zhang, 2022b), which may familiarize speakers with oral vibratory sensations and facilitate them to adopt favorable laryngeal and epilaryngeal configurations in a more open and natural vocal tract configuration. It is also possible that phonation with semi-occlusion at the lips may cause speakers to adjust their tongue position and shape, which may lead to favorable laryngeal and epilaryngeal configurations. How SOVTE facilitate the identification of such favorable configurations and how such favorable configurations are carried over to a more natural vocal tract configuration are still open research questions.

There are some limitations of this study that need to be mentioned. First, due to the large number of control parameters and source measures, in this study, we focused on global trends (main effects) observed across a large range of vocal conditions, which are thus readily available to be exploited by untrained speakers. It is possible that at some specific conditions, source-filter interaction may have local interaction effects that are larger than the global effects reported here, as shown in Sec. III D. Second, source-filter interaction is known to induce a qualitative mode change in vocal fold vibration when the fundamental frequency approaches the first formant (Titze, 2008; Echternach *et al.*, 2021), or when vocal fold vibration is near a bifurcation boundary between two qualitatively distinct modes of vibration (Herzel, 1993; Neubauer *et al.*, 2001; Tokuda *et al.*, 2010; Zañartu *et al.*, 2011; Zhang, 2018; Herbst *et al.*, 2023). These local effects of source-filter interaction will be explored in future studies. Finally, the glottal flow was simplified to be one-dimensional in our model, whereas complex flow behaviors have been reported. In particular, the interaction between the glottal jet and the false vocal folds is known to have an important impact on the overall glottal resistance (Bailly *et al.*, 2008; Zheng *et al.*, 2009; Kniesburges *et al.*, 2017). Thus, the findings of this study, particularly regarding false fold adduction, need to be verified in experiments or simulations in which the three-dimensional flow in the glottal and supraglottal region is adequately resolved.

ACKNOWLEDGMENTS

This study was supported by research grant R01 DC020240 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health.

AUTHOR DECLARATIONS

Conflict of Interest

The author has no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available from the author upon reasonable request.

- Alku, P. (1992). "Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering," *Speech Commun.* **11**(2–3), 109–118.
- Bailly, L., Bernardoni, N. H., Müller, F., Rohlf, A. K., and Hess, M. (2014). "Ventricular-fold dynamics in human phonation," *J. Speech. Lang. Hear. Res.* **57**(4), 1219–1242.
- Bailly, L., Pelorson, X., Henrich, N., and Rutu, N. (2008). "Influence of a constriction in the near field of the vocal folds: Physical modeling and experimental validation," *J. Acoust. Soc. Am.* **124**(5), 3296–3308.
- Bickley, C. A., and Stevens, K. N. (1986). "Effects of a vocal-tract constriction on the glottal source: Experimental and modelling studies," *J. Phon.* **14**(3–4), 373–382.
- Björkner, E., Sundberg, J., Cleveland, T., and Stone, E. (2006). "Voice source differences between registers in female musical theater singers," *J. Voice* **20**(2), 187–197.
- Chong, A. J., Risdal, M., Aly, A., Zymet, J., and Keating, P. (2020). "Effects of consonantal constrictions on voice quality," *J. Acoust. Soc. Am.* **148**(1), EL65–EL71.
- Dollinger, M., Berry, D., Luegmair, G., Huttner, B., and Bohr, C. (2012). "Effects of the epilarynx area on vocal fold dynamics and the primary voice signal," *J. Voice* **26**, 285–292.
- Dollinger, M., Berry, D., and Montequin, D. (2006). "The influence of epilarynx area on vocal fold dynamics," *Otolaryngol. Head. Neck Surg.* **135**, 724–729.
- Echternach, M., Herbst, C. T., Köberlein, M., Story, B., Döllinger, M., and Gellrich, D. (2021). "Are source-filter interactions detectable in classical singing during vowel glides?," *J. Acoust. Soc. Am.* **149**(6), 4565–4578.
- Fant, G. (1982). "Preliminaries to analysis of the human voice source," *STL-QPSR* **23**(4), 1–27.
- Farahani, M., and Zhang, Z. (2016). "Experimental validation of a three-dimensional reduced-order continuum model of phonation," *J. Acoust. Soc. Am.* **140**(2), EL172–EL177.
- Garellek, M., Samlan, R., Gerratt, B. R., and Kreiman, J. (2016). "Modeling the voice source in terms of spectral slopes," *J. Acoust. Soc. Am.* **139**(3), 1404–1410.
- Gauffin, J., and Sundberg, J. (1989). "Spectral correlates of glottal voice source waveform characteristics," *J. Speech. Lang. Hear. Res.* **32**(3), 556–565.
- Henrich, N., Smith, J., and Wolfe, J. (2011). "Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones," *J. Acoust. Soc. Am.* **129**(2), 1024–1035.
- Herbst, C. T., Elemans, C. P., Tokuda, I. T., Chatzioannou, V., and Švec, J. G. (2023). "Dynamic system coupling in voice production," *J. Voice* (published online).
- Herzel, H. (1993). "Bifurcations and chaos in voice signals," *Appl. Mech. Rev.* **46**(7), 399–413.
- Hochberg, Y., and Tamhane, A. C. (1987). *Multiple Comparison Procedures* (John Wiley & Sons, New York), Chap. 3, pp. 96–98.
- Holmberg, E., Hillman, R., and Perkell, J. (1988). "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," *J. Acoust. Soc. Am.* **84**(2), 511–529.
- Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Technical J.* **51**(6), 1233–1268.
- Joliveau, E., Smith, J., and Wolfe, J. (2004). "Tuning of vocal tract resonance by sopranos," *Nature* **427**(6970), 116.
- Kniesburges, S., Birk, V., Lodermeier, A., Schützenberger, A., Bohr, C., and Becker, S. (2017). "Effect of the ventricular folds in a synthetic larynx model," *J. Biomech.* **55**, 128–133.
- Li, S., Scherer, R. C., Fulcher, L. P., Wang, X., Qiu, L., Wan, M., and Wang, S. (2018). "Effects of vertical glottal duct length on intraglottal pressures and phonation threshold pressure in the uniform glottis," *J. Voice* **32**(1), 8–22.
- Li, Z., Chen, Y., Chang, S., and Luo, H. (2020). "A reduced-order flow model for fluid–structure interaction simulation of vocal fold vibration," *J. Biomech. Eng.* **142**(2), 0210051.
- McCullum, I., Throop, A., Badr, D., and Zakerzadeh, R. (2023). "Gender in human phonation: Fluid–structure interaction and vocal fold morphology," *Phys. Fluids* **35**(4), 041907.
- Milenkovic, P., and Mo, F. (1988). "Effect of the vocal tract yielding sidewall on inverse filter analysis of the glottal waveform," *J. Voice* **2**(4), 271–278.
- Mittal, V. K., Yegnanarayana, B., and Bhaskararao, P. (2014). "Study of the effects of vocal tract constriction on glottal vibration," *J. Acoust. Soc. Am.* **136**(4), 1932–1941.
- Montequin, D. A. (2003). "Developing a methodology to study the effect of the epilarynx tube on phonation threshold pressure and driving pressure," Ph.D. thesis, University of Iowa, Iowa City, IA.
- Murtola, T., Aalto, A., Malinen, J., Aalto, D., and Vainio, M. (2018). "Modal locking between vocal fold oscillations and vocal tract acoustics," *Acta Acust. united Ac.* **104**(2), 323–337.
- Neubauer, J., Mergell, P., Eysholdt, U., and Herzel, H. (2001). "Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes," *J. Acoust. Soc. Am.* **110**(6), 3179–3192.
- Rothenberg, M. (1981a). "Acoustic interaction between the glottal source and the vocal tract," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo), pp. 305–328.
- Rothenberg, M. (1981b). "An interactive model for the voice source," *STL-QPSR* **22**(4), 1–17.
- Samlan, R. A., and Kreiman, J. (2014). "Perceptual consequences of changes in epilaryngeal area and shape," *J. Acoust. Soc. Am.* **136**, 2798–2806.
- Story, B. H. (1995). "Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract," Ph.D. thesis, University of Iowa, Iowa City, IA.
- Story, B. H., and Titze, I. R. (1998). "Parameterization of vocal tract area functions by empirical orthogonal modes," *J. Phon.* **26**(3), 223–260.
- Sundberg, J. (1974). "Articulatory interpretation of the 'singing formant,'" *J. Acoust. Soc. Am.* **55**(4), 838–844.
- Sundberg, J. (2017). "An effect of source-filter interaction on amplitudes of source spectrum partials," in *Proceedings of the 10th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, MAVEBA 2017*, December 13–15, Firenze, Italy, pp. 95–98.
- Sundberg, J., Lã, F. M. B., and Gill, B. P. (2013). "Formant tuning strategies in professional male opera singers," *J. Voice* **27**(3), 278–288.
- Taylor, C. J., and Thomson, S. L. (2022). "Optimization of synthetic vocal fold models for glottal closure," *J. Eng. Sci. Med. Diagn. Ther.* **5**(3), 031106.
- Titze, I. R. (1988). "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.* **83**, 1536–1552.
- Titze, I. R. (2002). "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," *J. Acoust. Soc. Am.* **111**(1), 367–376.
- Titze, I. (2006). "Voice training and therapy with a semi-occluded vocal tract: Rationale and scientific underpinnings," *J. Speech. Lang. Hear. Res.* **49**, 448–459.
- Titze, I. R. (2008). "Nonlinear source–filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**(4), 1902–1915.
- Titze, I. R., and Palaparthi, A. (2016). "Sensitivity of source–filter interaction to specific vocal tract shapes," *IEEE/ACM Trans. Audio. Speech. Lang. Process.* **24**(12), 2507–2515.
- Titze, I. R., and Story, B. H. (1997). "Acoustic interactions of the voice source with the lower vocal tract," *J. Acoust. Soc. Am.* **101**(4), 2234–2243.
- Tokuda, I. T., Zemke, M., Kob, M., and Herzel, H. (2010). "Biomechanical modeling of register transitions and the role of vocal tract resonators," *J. Acoust. Soc. Am.* **127**, 1528–1536.
- Verdolini-Abbott, K., Li, N. Y., Branski, R. C., Rosen, C. A., Grillo, E., Steinhauer, K., and Hebda, P. A. (2012). "Vocal exercise may attenuate acute vocal fold inflammation," *J. Voice* **26**(6), 814.e1.
- Yoshinaga, T., Zhang, Z., and Iida, A. (2022). "Comparison of one-dimensional and three-dimensional glottal flow models in left-right asymmetric vocal fold conditions," *J. Acoust. Soc. Am.* **152**(5), 2557–2569.
- Zañartu, 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**(1), 326–339.
- Zañartu, M., Mongeau, L., and Wodicka, G. R. (2007). "Influence of acoustic loading on an effective single mass model of the vocal folds," *J. Acoust. Soc. Am.* **121**, 1119–1129.
- 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**(2), 898–910.
- Zhang, Z. (2016). "Mechanics of human voice production and control," *J. Acoust. Soc. Am.* **140**(4), 2614–2635.
- Zhang, Z. (2017). "Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model," *J. Acoust. Soc. Am.* **142**(4), 2311–2321.
- Zhang, Z. (2018). "Vocal instabilities in a three-dimensional body-cover phonation model," *J. Acoust. Soc. Am.* **144**(3), 1216–1230.
- Zhang, Z. (2019). "Vocal fold contact pressure in a three-dimensional body-cover phonation model," *J. Acoust. Soc. Am.* **146**(1), 256–265.
- Zhang, Z. (2020). "Laryngeal strategies to minimize vocal fold contact pressure and their effect on voice production," *J. Acoust. Soc. Am.* **148**, 1039–1050.
- Zhang, Z. (2021). "Interaction between epilaryngeal and laryngeal adjustments in regulating vocal fold contact pressure," *JASA Express Lett.* **1**(2), 025201.
- Zhang, Z. (2022a). "Estimating subglottal pressure and vocal fold adduction from the produced voice in a single subject study," *J. Acoust. Soc. Am.* **151**, 1337–1340.
- Zhang, Z. (2022b). "Oral vibratory sensations during voice production at different laryngeal and semi-occluded vocal tract configurations," *J. Acoust. Soc. Am.* **152**, 302–312.
- Zhang, Z. (2023). "Vocal fold vertical thickness in human voice production and control: A review," *J. Voice*.
- Zhang, Z., and Luu, T. H. (2012). "Asymmetric vibration in a two-layer vocal fold model with left-right stiffness asymmetry: Experiment and simulation," *J. Acoust. Soc. Am.* **132**, 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., Neubauer, J., and Berry, D. A. (2009). "Influence of vocal fold stiffness and acoustic loading on flow-induced vibration of a single-layer vocal fold model," *J. Sound Vib.* **322**, 299–313.
- Zheng, X., Bielałowicz, S., Luo, H., and Mittal, R. (2009). "A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation," *Ann. Biomed. Eng.* **37**, 625–642.