JASA ARTICLE

Oral vibratory sensations during voice production at different laryngeal and semi-occluded vocal tract configurations

Zhaoyan Zhang[a)]
Department of Head and Neck Surgery, University of California, Los Angeles, 31-24 Rehabilitation Center, 1000 Veteran Ave., Los Angeles, California 90095-1794, USA

ABSTRACT:
Voice therapy often emphasizes vibratory sensations in the front part of the vocal tract during phonation to improve vocal efficiency. It remains unclear what laryngeal and vocal tract adjustments are elicited in speakers by this emphasis on oral vibratory sensations. Using a three-dimensional phonation model, this study aims to identify laryngeal and epilaryngeal adjustments that might produce maximal oral vibratory sensations during phonation, as quantified by the oral sound pressure level (SPL), and thus are likely to be elicited in voice therapy at different semi-occluded vocal tract configurations. Results show that maximum oral SPL occurs at intermediate vocal fold adduction configurations characterized by a trade-off between glottal gap and vocal fold vertical thickness. Epilaryngeal tube narrowing further increases the oral SPL in an open vocal tract, but has little effect on oral SPL in semi-occluded vocal tracts. Laryngeal and epilaryngeal configurations producing the maximum oral SPL generally have lower peak vocal fold contact pressure when producing a target output SPL. These favorable configurations are more easily identified in open vocal tracts than semi-occluded vocal tracts. However, semi-occlusion increases both the mean and dynamic oral pressure, which may familiarize speakers with oral vibratory sensations and facilitate adoption of favorable laryngeal configurations.

https://doi.org/10.1121/10.0012365

(Received 6 May 2022; revised 16 June 2022; accepted 20 June 2022; published online 8 July 2022)

[Editor: James F. Lynch] Pages: 302–312

I. INTRODUCTION

Voice therapy or vocal training often emphasizes vibratory sensations in the front part of the vocal tract, particularly when aimed at improving vocal efficiency [e.g., vocal function exercises (Stemple *et al.*, 1994), resonant voice therapy (Verdolini-Marston *et al.*, 1995), and semi-occluded vocal tract exercises (SOVTE) (Titze, 2006)]. While it is generally assumed that such emphasis on oral vibratory sensations leads to changes in voice production that maximize vocal efficiency and minimize vocal fold injury, it remains unclear what laryngeal and vocal tract adjustments are elicited by such emphasis and whether or not these adjustments actually improve voice production and reduce risk of vocal fold injury. In this study we investigate oral vibratory sensations, as indirectly quantified by sound pressure level (SPL) in the oral cavity behind the lips, at different laryngeal and vocal tract configurations in a three-dimensional voice production model. The goal is to identify laryngeal and vocal tract adjustments that might result in maximal oral vibratory sensations and thus are likely to be elicited by an emphasis on vibratory sensations in voice therapy. Identification of such adjustments would provide insight into the scientific rationale for voice therapy in improving voice production, and may allow clinicians to better target patient-specific vocal behavior to improve treatment outcomes in voice therapy.

Theoretically, the oral SPL and thus vibratory sensations can be increased by either increasing the power supply of voice production (i.e., the subglottal pressure), improving vocal efficiency, or semi-occlusion at the lips. Increasing the subglottal pressure often significantly increases both vocal effort and risk of vocal fold injury, and patients in voice therapy are often instructed to maintain ease of phonation to avoid excessively high subglottal pressure. Vocal efficiency can be improved by laryngeal or vocal tract adjustments. In this paper, we quantify vocal efficiency as the ratio between the output sound pressure level radiated from the mouth and the corresponding subglottal pressure. Previous studies showed that at the laryngeal level, vocal efficiency is the highest at a barely abducted laryngeal configuration [i.e., the vocal folds almost touch each other but without compression (Berry *et al.*, 2001; Zhang, 2020)] and in vocal folds with an intermediate vertical thickness (Zhang, 2021a). Vocal efficiency can be further increased by improving the impedance matching between the sound source at the glottis and the free space outside the mouth, by either narrowing the epilaryngeal tube (Sundberg, 1974; Titze and Story, 1997; Zhang, 2021a), increasing mouth opening (Titze and Worley, 2009; Zhang, 2021b), lengthening the vocal tract (Story *et al.*, 2000), or a combination of them. While these adjustments may increase vocal fold contact pressure for a given subglottal pressure (Zhang, 2020, 2021a, 2021b), the improved vocal efficiency often significantly reduces the subglottal pressure required to produce a target output SPL, thus

---

[a)]Electronic mail: zyzhang@ucla.edu

reducing the overall vocal fold contact pressure and risk of vocal fold injury in voice tasks targeting a specific output SPL. Thus, focusing on the oral vibratory sensations while maintaining ease of phonation is expected to facilitate speakers to adopt at least some of these favorable laryngeal and/or vocal tract configurations that maximize vocal efficiency and reduce vocal fold contact pressure.

Oral vibratory sensations during phonation can also be increased by significantly reducing the mouth opening, which reduces output SPL but allows sound pressure to build up better inside the oral cavity. This is particularly the case for SOVTE, in which the mouth opening is either reduced or constricted by a narrow tube. While SOVTE have become widely used in voice clinics, it remains unclear whether or not the favorable laryngeal and vocal tract adjustments discussed above would still increase oral SPL in a semi-occluded vocal tract configuration, and thus be elicited by an emphasis on oral vibratory sensations in SOVTE.

While there have been many studies investigating changes in laryngeal and vocal tract configurations during and after SOVTE (Vampola *et al.*, 2011; Laukkanen *et al.*, 2012; Guzman *et al.*, 2013a,b; Guzman *et al.*, 2017; Hampala *et al.*, 2015; Patel *et al.*, 2019; Lulich and Patel, 2021), the findings are not always consistent with theoretical predictions. These imaging studies generally reported a lower vertical laryngeal position after the exercises, indicating an elongated vocal tract. A low vertical laryngeal position is often associated with reduced vocal fold adduction, which was observed in Guzman *et al.* (2013b). However, the opposite (i.e., increased adduction) was observed in Laukkanen (1992), and no prominent trends in vocal fold adjustments were observed in a computed tomography study (Hampala *et al.*, 2015). Within the vocal tract, these studies often reported improved velopharyngeal closure and increased tongue dorsum height. However, vocal tract changes in the epilarynx are inconsistent. While Guzman *et al.* (2013a) reported increased anterior-posterior aryepiglottic narrowing in a videolaryngoscopic study, most other studies reported widening of the epilarynx and hypopharynx.

These conflicting results could be due to the fact that vocal exercises may induce simultaneous adjustments in the respiratory and laryngeal subsystems as well as the vocal tract, and different speakers may adopt different adjustments. Speakers may also differ in acuity to oral vibratory sensations, which may affect their ability to identify favorable laryngeal and vocal tract configurations based on oral vibratory sensations. To circumvent these difficulties, in this study we use a computational model to identify laryngeal and vocal tract adjustments that maximize oral SPL at different semi-occluded vocal tract configurations. Specifically, this study aims to answer three questions. First, what laryngeal and epilaryngeal configurations can be elicited by an emphasis on maximizing oral SPL in an open or semi-occluded vocal tract? Second, are these configurations associated with reduced vocal fold contact pressure when producing a target output SPL? Finally, does semi-occlusion of the vocal tract make it easier for speakers to identify these favorable configurations compared to vocal tracts without semi-occlusion?

II. METHOD

A. Computational model and simulation conditions

The same three-dimensional body-cover vocal fold model as in our previous studies (Zhang, 2017, 2019, 2020) was used in this study. The reader is referred to these previous studies for details of the model (Zhang, 2017, 2019, 2020). Briefly, each vocal fold layer is modeled as a transversely isotropic, nearly incompressible, linear material with a plane of isotropy perpendicular to the anterior-posterior (AP) direction. The glottal flow is modeled as a one-dimensional quasi-steady glottal flow model taking into consideration viscous loss, as described in detail in Zhang (2017). Vocal fold contact occurs when portions of the vocal fold cross the glottal midline, in which case a penalty pressure perpendicular into the vocal fold is applied to the contact surface of the vocal fold (Zhang, 2019). The parameters of the penalty pressure model are selected to ensure a small penetration depth of the vocal folds crossing the glottal midline so that the corresponding penalty pressure would approximate the true contact pressure (Zhang, 2019). This voice production model has been able to qualitatively and quantitatively reproduce experimental observations (Zhang *et al.*, 2002; Zhang and Luu, 2012; Farahani and Zhang, 2016).

The vocal fold geometry is parameterized by various geometric and mechanical properties of the vocal folds (Fig. 1). In this study, two important geometric variables, the initial glottal angle α and the medial surface vertical thickness T, were systematically varied to produce different conditions of vocal fold adduction. These two geometric controls were shown in our previous studies to have a large effect on voice production. Specifically, the initial glottal angle α was varied from –1.6°, 0°, 1.6°, 4°, to 8°. The negative initial glottal angle corresponds to a pressed vocal fold configuration,

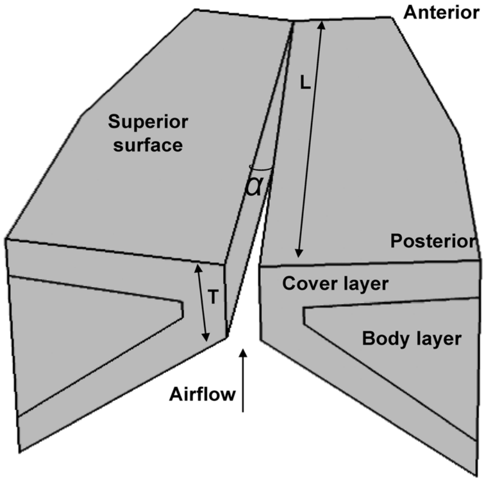

FIG. 1. The three-dimensional vocal fold model. The initial glottal angle α and medial surface vertical thickness T were systematically varied to produce different conditions of vocal fold adduction.

whereas the 0°–1.6° and 4°–8° initial glottal angles roughly correspond to a barely abducted and breathy vocal fold configuration, respectively. The medial surface vertical thickness was varied from 1, 2, 3, to 4.5 mm. Our previous studies showed that this range of vertical thicknesses was able to produce voices of different quality ranging from breathy, normal, and pressed voices as well as irregular vocal fold vibrations (Zhang, 2016, 2018).

Mechanically, each vocal fold layer is characterized by its transverse Young's modulus E_t, the AP Young's modulus E_{ap}, the AP shear modulus G_{ap}, and density. The effect of these mechanical properties on voice production has been investigated in detail in our previous studies (Zhang, 2017, 2020). To reduce the number of conditions to be simulated, in this study the body and cover layers were set to have identical mechanical properties, with E_t, G_{ap}, and E_{ap} set to 2, 10, and 40 kPa, respectively. Simulations with systematic variations in vocal fold stiffness, particularly the transverse stiffness in the coronal plane, will be performed in future studies to investigate potential effects of stiffness on oral SPL and vibratory sensations.

B. Vocal tract configurations

Three vocal tract configurations were considered in this study (Fig. 2). The first two were vocal tract shapes corresponding to the open vowel /ɑ/ and close vowel /u/ produced by a male subject, as measured from magnetic resonance imaging (Story *et al.*, 1996). These two configurations were compared to investigate the differences in voice production between open and close vowels. The third vocal tract configuration was included to investigate the effect of semi-occlusion at the lips, by extending the /u/ vocal tract with an 8-cm long tube with an inner diameter of 5 mm (Fig. 2, third column; simulating a drinking straw as often used in semi-occluded vocal tract exercises).

For each vocal tract configuration, the area function in the epilaryngeal region was further manipulated to simulate epilaryngeal expansion or narrowing. As in Zhang (2021a), the area function in the epilaryngeal region was multiplied by a scaling factor S_e ranging from 0.5, 1, 2, to 3. This manipulation was applied to the first 3.2 and 2.4 cm segment of the vocal tract closest to the glottis for the /ɑ/- and /u/-based vocal tract configurations, respectively.

The vocal tract is modeled as a one-dimensional waveguide coupled with a yielding vocal tract wall (Story, 1995). 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 N s/m^3, respectively (Milenkovic and Mo, 1988). For each vocal tract configuration, the impulse response was obtained by exciting the vocal tract with an impulse input to the vocal tract entrance. The vocal tract transfer function, defined as the ratio between the vocal tract output and input acoustic volume velocities, and the vocal tract input impedance were then calculated from the impulse response. The imaginary part of the vocal tract input impedance was further divided by the angular frequency to obtain the vocal tract input inertance. Note that the vocal tract inertance is defined only if the input reactance is positive.

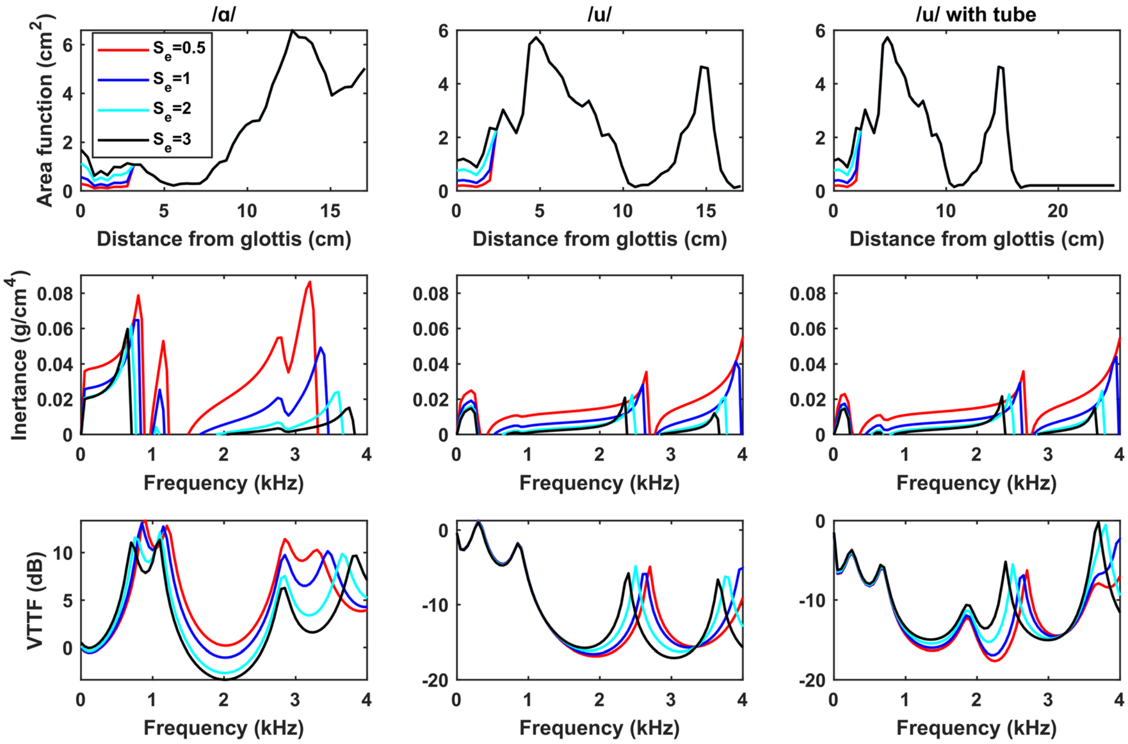

FIG. 2. (Color online) The vocal tract area function (top), inertance (middle), and vocal tract transfer function (bottom) of the /ɑ/ vocal tract, /u/ vocal tract, and /u/ vocal tract extended by a tube.

C. Simulations and data analysis

For each vocal fold and vocal tract condition, a half-second of voice production was simulated for a constant subglottal pressure ranging from 100 Pa to 2.4 kPa, similar to Zhang (2020, 2021a, 2021b). From each simulation, the output SPL was extracted from the output acoustics as described in Zhang (2016). An oral SPL was also calculated using the acoustic pressure inside the oral cavity at a location 8 mm from the lips. Laryngeal and epilaryngeal conditions that produce maximum oral SPL were then identified for each of the three vocal tract configurations.

The impact of different laryngeal, epilaryngeal, and vocal tract semi-occlusion configurations (i.e., open vs semi-occluded vocal tract) on voice production was evaluated by comparing the subglottal pressure required to produce a target output SPL and the corresponding peak vocal fold contact pressure P_c. The subglottal pressure required to produce a target output SPL is an indirect measure of vocal efficiency, whereas the peak vocal fold contact pressure quantifies the risk of vocal fold injury. In this study, the target output SPL was set to 70, 50, and 40 dB for the /ɑ/ vocal tract, /u/ vocal tract, and /u/ vocal tract with a tube, respectively, based on the SPL range produced with each vocal tract configuration. The impact on ease of phonation was evaluated by calculating the phonation threshold pressure, or the minimum pressure required to initiate and sustain phonation, at different vocal fold and vocal tract configurations. The closed quotient (CQ) was also calculated as the fraction of the cycle in which the glottal flow falls within the lower 10% between the minimum and maximum glottal flow rates.

III. RESULTS

A. Effect of semi-occlusion on vocal tract transfer function and input inertance

Figure 2 shows the vocal tract area function and the corresponding input inertance and vocal tract transfer function for the three vocal tract configurations at different degrees of epilaryngeal manipulations. As expected, the vocal tract transfer function is higher for the /ɑ/ vocal tract than the /u/ vocal tract, partially due to the difference in lip opening. In all three configurations, epilaryngeal tube narrowing increases vocal tract inertance, particularly in the 2–4 kHz range.

For the /ɑ/ vocal tract, epilaryngeal narrowing also increases the first and second formants, and brings them closer in frequency. As a result, the vocal tract transfer function at low frequencies becomes stronger with increasing epilaryngeal narrowing. In contrast, for the /u/ vocal tract, with or without a tube, manipulations of the epilaryngeal area function have little effect on the low-frequency portion, particularly around the first formant, of the vocal tract transfer function. This reduced sensitivity is because for a vocal tract constricted in both ends, as in the case of the last two vocal tract configurations in Fig. 2, the first formant is related to vocal tract wall impedance and the acoustic compliance associated with the overall vocal tract volume (Stevens, 1998; Patel *et al.*, 2019). Under such conditions, epilaryngeal manipulation is expected to have minimal effect on the first formant, as shown in Fig. 2.

B. Phonation threshold pressure and mean oral pressure

Figure 3 shows the phonation threshold pressure at different laryngeal and vocal tract conditions. The general trends of variation of the phonation threshold pressure with the initial glottal angle and thickness are similar to findings from our previous study (Zhang, 2017) in the absence of a vocal tract. In general, the phonation threshold pressure is the lowest at conditions with an initial glottal angle of 0° and 1.6°, particularly for thick vocal folds ($T = 3$ and 4.5 mm).

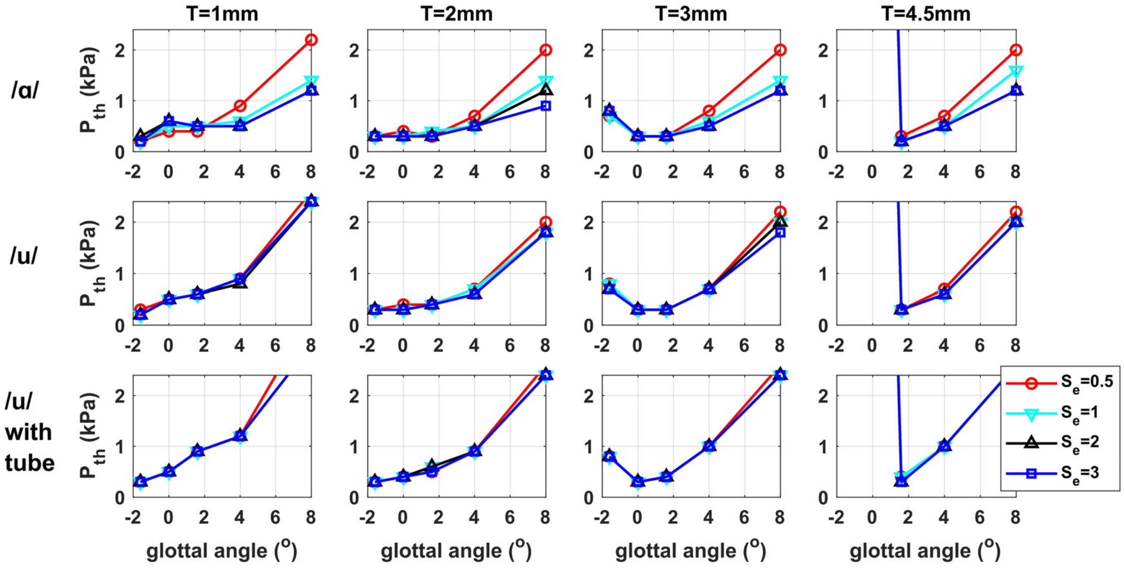

FIG. 3. (Color online) Phonation threshold pressure P_{th} as a function of the initial glottal angle, vertical thickness, and epilaryngeal manipulation for conditions with a vocal tract shape of the /ɑ/ (top), /u/ (middle), and /u/ extended with a tube (bottom).

The effect of epilaryngeal narrowing is generally small and inconsistent at small initial glottal angles. This is consistent with our experimental observation that the effect of vocal tract acoustics on phonation onset is generally small when the fundamental frequency is far away from vocal tract resonances (Zhang *et al.*, 2006). At large initial glottal angles ($\alpha > 2°$), extreme epilaryngeal narrowing reduces the glottal flow peak-to-peak amplitude (Zhang, 2021a) and thus a higher subglottal pressure is required to initiate phonation, which is consistent with the observation in Titze (2002). This increase in phonation threshold pressure is particularly large for the /ɑ/ vocal tract, but much smaller in vocal tracts that are already semi-occluded.

In general, the phonation threshold pressure is the highest in the /u/ vocal tract with a tube, followed by the /u/ vocal tract, and is the lowest in the /ɑ/ vocal tract. This is particularly the case at large initial glottal angles ($\alpha \geq 0°$). This increase may be related to the buildup of pressure within the semi-occluded vocal tracts. In our simulations, for a subglottal pressure of 1 kPa, the mean oral pressure behind the lips ranges from 30 to 280 Pa for the /u/ vocal tract and from 90 to 500 Pa for the /u/ vocal tract with a tube, in contrast to being almost zero for the /ɑ/ vocal tract. These mean oral pressure values for the /u/ vocal tract with a tube are comparable to those measured in humans by Guzman *et al.* (2013b) and Maxfield *et al.* (2015). In general, the mean oral pressure increases with increasing initial glottal angle and decreasing vertical thickness, both of which reduce the laryngeal impedance to the airflow and thus increase the mean oral pressure.

C. Oral SPL in an open vocal tract

The first row of Fig. 4 shows the oral SPL in the /ɑ/ vocal tract as a function of the initial glottal angle and epilaryngeal manipulation for conditions of different medial surface vertical thickness and a subglottal pressure of 1 kPa. Overall, the oral SPL is lower in the thinnest ($T = 1$ mm) and thickest ($T = 4.5$ mm) conditions than the two intermediate thickness conditions. For each thickness, the maximum oral SPL occurs at an intermediate glottal angle. In general, the initial glottal angle at which the maximal oral SPL occurs increases with increasing thickness (Fig. 5). For example, maximal oral SPL occurs at the negative initial glottal angle of –1.6° for $T = 1$ mm, 0° for $T = 2$ mm, and 1.6° for $T = 3$ and 4.5 mm. In this vocal tract configuration and a subglottal pressure of 1 kPa, the overall peak oral SPL occurs at the condition $T = 3$ mm and an initial glottal angle of 1.6°. Our previous studies (Zhang, 2016, 2017) showed that the thickness and the initial glottal angle are the two most important parameters in controlling the mean glottal flow, mean glottal opening area, and closed quotient of vocal fold vibration. A large thickness coupled with a small or negative initial glottal angle often produces irregular vibrations that are typical of pressed phonation, whereas the other extreme often results in breathy phonation. The trade-off between the thickness and initial glottal angle shown in Fig. 5 thus represents a series of intermediate adduction conditions that are neither too tight nor too open. A similar trade-off was observed in Titze (2006) and Laukkanen *et al.* (2008).

Figure 4 (first row) also shows that in general the oral SPL increases with increasing epilaryngeal tube narrowing. This effect is the largest at the thinnest conditions, with the oral SPL increasing by as much as 10 dB. For some conditions, the oral SPL is maximum at an intermediate degree of epilaryngeal narrowing and decreases with further epilaryngeal narrowing. This reduction of the oral SPL is likely because extreme epilaryngeal narrowing increases the vocal tract impedance and reduces the peak-to-peak amplitude of the glottal flow (Titze, 2002; Zhang, 2021a).

The bottom three rows of Fig. 4 shows the impact of different laryngeal and epilaryngeal adjustments on voice production. Note that this impact is evaluated at conditions that produce a target output SPL, 70 dB in this case, in order to focus on the effect on vocal efficiency. Such focus is of practical importance as effective communication often requires voice production with a certain output SPL.

Figure 4 shows that when targeting a specific output SPL, conditions producing the highest oral SPL generally require the least subglottal pressure to produce the target output SPL and thus have a lower peak vocal fold contact pressure. However, the lowest peak contact pressure occurs at the thinnest vocal fold conditions, at the cost of an increased subglottal pressure required to produce the target 70 dB output SPL. The bottom row of Fig. 4 also shows a general trend of decreasing closed quotient with decreasing vocal fold thickness T, which is also accompanied by reduced higher-order harmonic excitation in the voice source spectra (not shown in the figure).

Thus, for an open vocal tract such as the /ɑ/, aiming to maximize the oral SPL (thus oral vibratory sensations) facilitates the speaker to adopt an intermediate vocal fold adduction configuration that is neither too tight (large thickness and small or negative initial glottal angle) nor too open (small thickness and large initial glottal angle), and epilaryngeal tube narrowing. Both adjustments significantly reduce the peak vocal fold contact pressure and the required subglottal pressure when targeting a specific output SPL.

D. Oral SPL in semi-occluded vocal tracts

Figure 6 shows the results for the /u/ vocal tract. The general trends with regard to how the oral SPL varies with the initial glottal angle and vertical thickness remain the same as those in /ɑ/ vocal tract. The oral SPL is much lower in the thickest and the thinnest vocal fold condition, and in conditions with a very large initial glottal angle. The initial glottal angle at which the maximum oral SPL occurs increases with increasing thickness (Fig. 5, second column). Thus, emphasis on maximizing the oral SPL likely will result in vocal fold configurations that are neither too tight nor too open, with a trade-off between the thickness and initial glottal angle, as shown in Fig. 5. In general, conditions

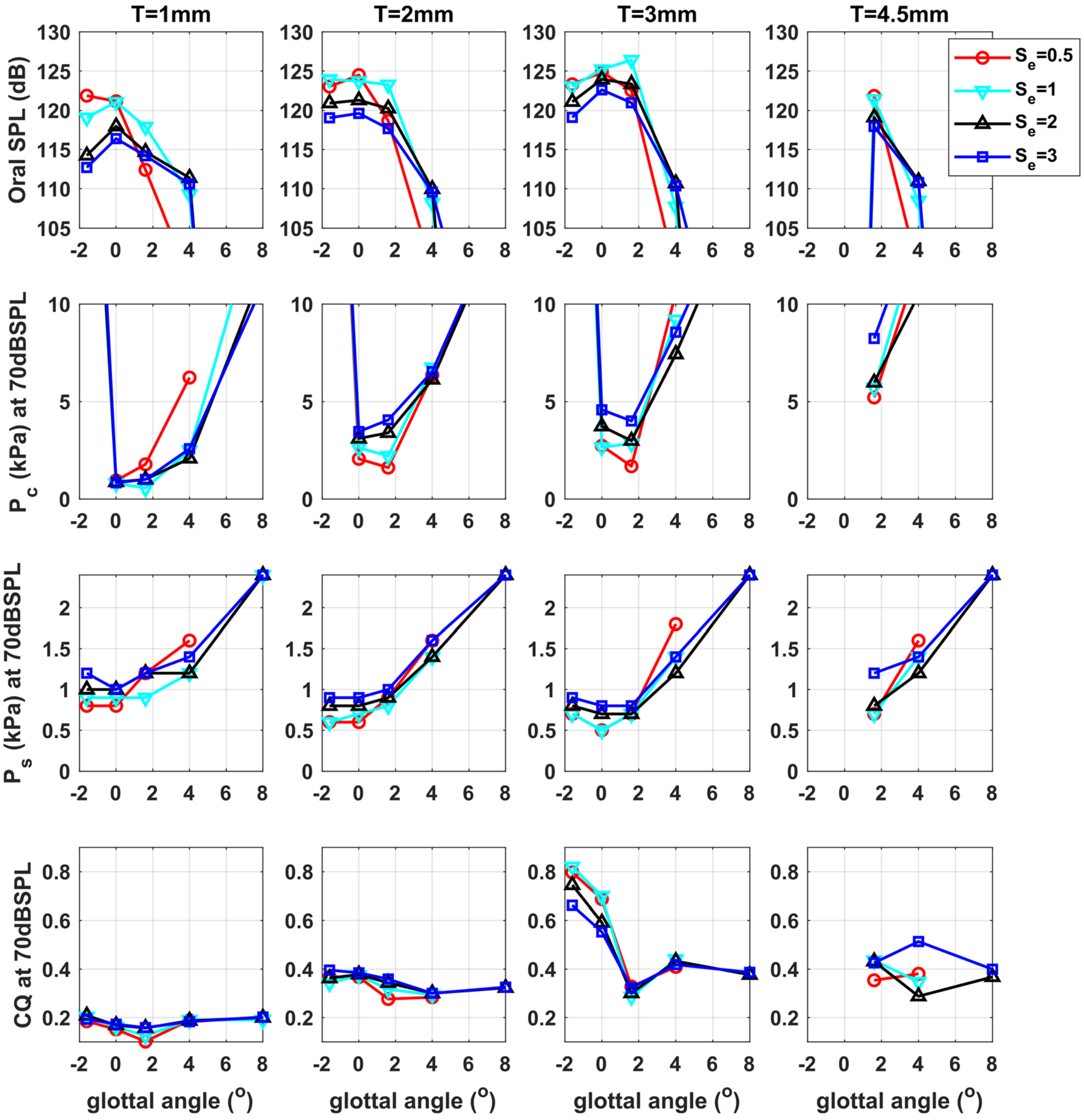

FIG. 4. (Color online) /ɑ/ vocal tract. From top to bottom: the oral SPL produced at a 1 kPa subglottal pressure, and the peak contact pressure P_c and subglottal pressure P_s required to produce a 70 dB SPL outside and the corresponding CQ values as a function of the initial glottal angle, epilaryngeal scaling factor, and medial surface vertical thickness T.

producing the highest oral SPL have a lower peak vocal fold contact pressure, although the lowest peak contact pressure occurs in the thinnest vocal fold conditions.

In contrast to the /ɑ/ vocal tract in which epilaryngeal narrowing generally results in a significant increase in the oral SPL, the effect of epilaryngeal narrowing on the oral SPL is almost negligible in the /u/ vocal tract. This small effect is consistent with the small effect of epilaryngeal manipulations on the vocal tract transfer function (Fig. 2). Thus, emphasis on maximizing the oral SPL in the /u/ vocal tract likely will not elicit epilaryngeal narrowing, and may even suppress extreme epilaryngeal narrowing.

Similar trends are also observed for the /u/ vocal tract semi-occluded with a 5 mm-diameter tube in Fig. 7. The overall maximum oral SPL occurs at the condition $T = 2$ mm and an initial glottal angle of 0°. Epilaryngeal tube narrowing slightly decreases the oral SPL, as in the /u/ vocal tract. A notable difference between this vocal tract condition (/u/ extended with a tube) and the /u/ vocal tract is that the phonation range in terms of the initial glottal angle is more restricted when the vocal tract is constricted with a tube. In both vocal tract configurations, variations in oral SPL across different laryngeal conditions are generally smaller than that in the /ɑ/ vocal tract, which can also be observed in Fig. 5.

E. Laryngeal and epilaryngeal conditions elicited by an emphasis on oral SPL

One research question of this study is whether or not semi-occlusion at the lips makes it easier to identify favorable laryngeal and epilaryngeal conditions. It is reasonable

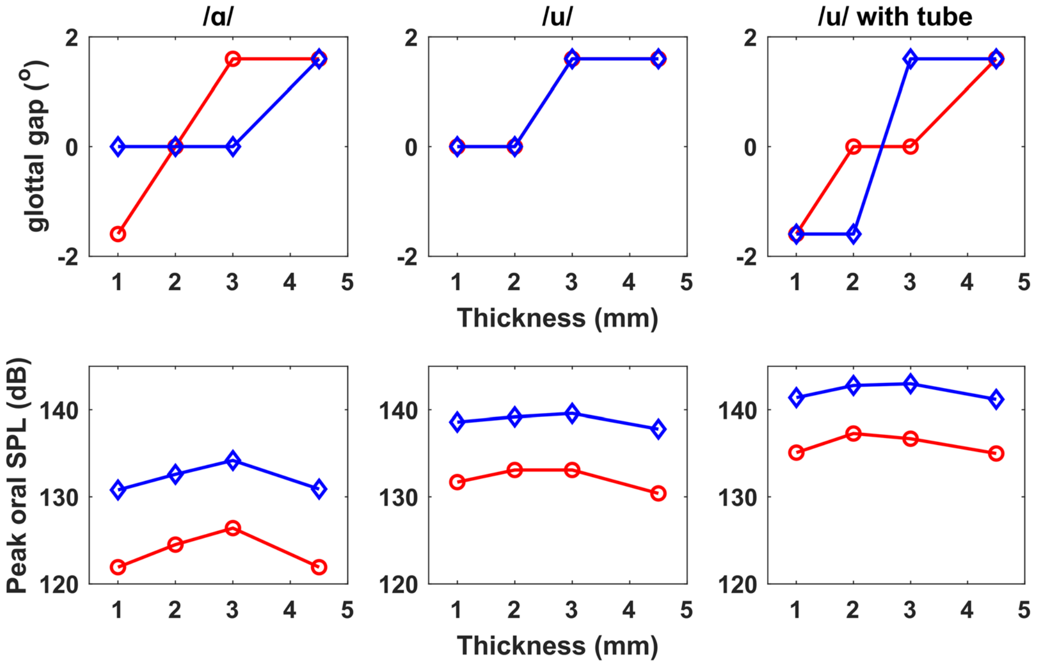

FIG. 5. (Color online) Laryngeal configurations that maximize the oral SPL (top) show a trade-off between the initial glottal angle and vocal fold vertical thickness. The bottom panels show the corresponding peak oral SPL. Circles: $P_s = 1$ kPa; diamonds: $P_s = 1.8$ kPa. The /ɑ/ vocal tract has the largest range of variation in peak oral SPL across different laryngeal conditions.

to assume that large variations in oral SPL across different laryngeal conditions (thus high sensitivity of oral SPL to laryngeal and epilaryngeal adjustments) would allow speakers to better identify and adopt favorable conditions that maximize the oral SPL. In this study, we evaluate this by identifying the laryngeal and epilaryngeal conditions that produce an oral SPL within a given decibel range of the overall maximum oral SPL for each of the three vocal tract configurations. Our rationale is that the fewer favorable conditions identified for a vocal tract configuration, the easier it would be for speakers to identify and adopt the favorable laryngeal and epilaryngeal conditions. Three decibel ranges (1, 3, and 5 dB) are considered, simulating varying degrees of speaker acuity to oral vibratory sensations. The results are shown in Fig. 8.

In general, none of three vocal tract configurations has identified conditions that are either too tight (large thickness and negative glottal angles) or too open (large glottal angles and small thickness), indicating that all three vocal tracts are able to avoid these inefficient conditions with an emphasis on maximizing oral SPL. The /ɑ/ vocal tract has the fewest identified favorable conditions, implying that it might be easier to identify favorable conditions in the /ɑ/ vocal tract than the other two vocal tracts. Both the /u/ vocal tract and /u/ vocal tract with a tube have considerably more identified conditions, including some conditions with the smallest thickness ($T = 1$ mm) and some relatively tight adduction conditions (e.g., $T = 3$ mm and $\alpha = -1.6°$). The /u/ vocal tract with a tube has slightly fewer conditions than the /u/ vocal tract, indicating a slight advantage of the /u/ vocal tract with a tube in identifying favorable conditions. As expected, for all three vocal tracts, the number of identified conditions increases with increasing decibel range or decreasing speaker acuity to vibratory sensations, indicating increasing difficulty in identifying favorable conditions with decreasing speaker acuity.

IV. DISCUSSION

The main research questions of this study are what laryngeal and epilaryngeal configurations can be elicited by an emphasis on maximizing oral SPL during phonation, and whether or not these configurations produce the lowest peak vocal fold contact pressure when producing a target output SPL. Our results show that at the laryngeal level, maximizing oral vibratory sensations at a given subglottal pressure could lead to intermediate adduction conditions that are neither too tight nor too open, which generally have a lower peak vocal fold contact pressure when producing a target output SPL and require less subglottal pressure to do so. This correspondence between maximum oral SPL and intermediate adduction conditions is consistently observed whether or not the vocal tract is open or semi-occluded at the lips.

Previous studies (Berry *et al.*, 2001; Verdolini *et al.*, 1998) have shown that maximum vocal economy is obtained at barely adducted vocal fold conditions, or a gap width around 0.5–0.6 mm between the vocal processes. In contrast, our results identify a range of favorable laryngeal conditions resulting from a trade-off between the vertical thickness and initial glottal angle (Fig. 5). A similar trade-off between the glottal gap width and vocal fold thickness in favorable laryngeal configurations was also observed in Titze (2006) and Laukkanen *et al.* (2008). The existence of a range of favorable laryngeal configurations implies that speakers

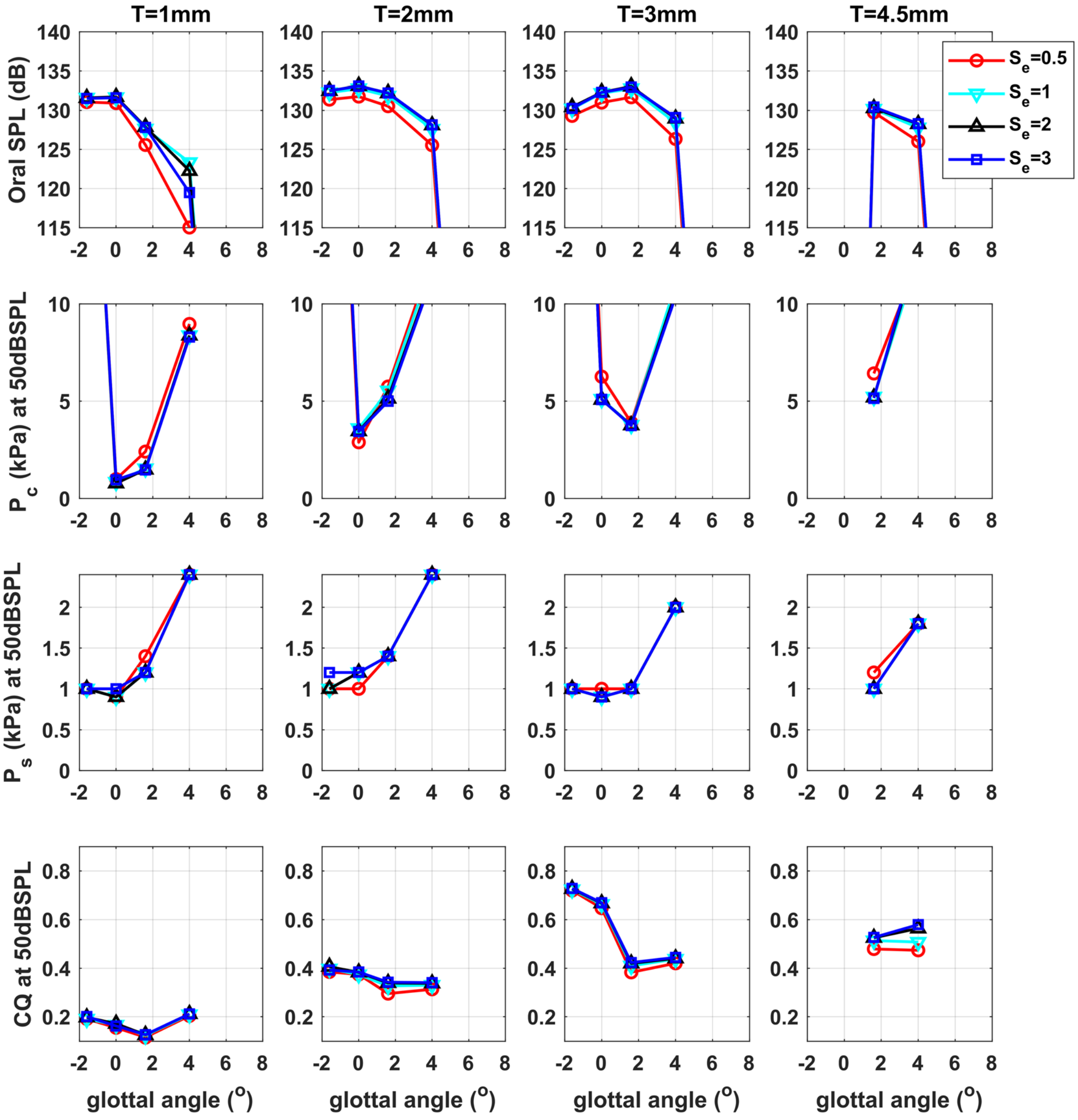

FIG. 6. (Color online) /u/ vocal tract. From top to bottom: the oral SPL produced at a 1 kPa subglottal pressure, and the peak contact pressure P_c and subglottal pressure P_s required to produce a 50 dB SPL outside and the corresponding CQ values as a function of the initial glottal angle, epilaryngeal scaling factor, and medial surface vertical thickness T.

may maximize oral SPL by adopting different laryngeal configurations, which makes it possible to produce desired voice quality while still maintaining vocal efficiency and relatively low risk of vocal fold injury. This may also partially explain the relatively large variations in voice outcome measures observed in recent experimental studies.

While epilaryngeal tube narrowing is generally considered to enhance source-tract interaction and thus expected to increase oral SPL, our results show that this is only true for an open vocal tract, such as the /ɑ/ vocal tract investigated in this study. With an open vocal tract, an emphasis on maximizing oral vibratory sensations facilitates epilaryngeal narrowing, which also reduces the peak vocal fold contact pressure when producing a target output SPL. In contrast, our study also shows that the effect of epilaryngeal tube manipulations on oral SPL is generally small in a semi-occluded vocal tract, with the oral SPL decreasing only slightly with epilaryngeal narrowing. Thus, speakers in general will not experience noticeable changes in the oral SPL whether they expand or narrow the epilarynx. In other words, maximizing oral SPL with semi-occlusion at the lips does not facilitate any consistent vocal tract changes in the epilarynx in most people, and may lead to slight expansion in the epilarynx in speakers with high acuity to vibratory sensations. This is consistent with the observation in recent imaging studies that SOVTE often lead to widening of the epilarynx [e.g., Vampola *et al.* (2011), Laukkanen *et al.* (2012), Guzman *et al.* (2013b), Guzman *et al.* (2017), Hampala *et al.* (2015), Patel *et al.* (2019), and Lulich and Patel (2021)].

This insensitivity of oral SPL to epilaryngeal tube manipulations in semi-occluded vocal tracts is due to the small effect of epilaryngeal manipulations on the first formant in semi-occluded vocal tracts, in which the first formant is determined mainly by the acoustic compliance of

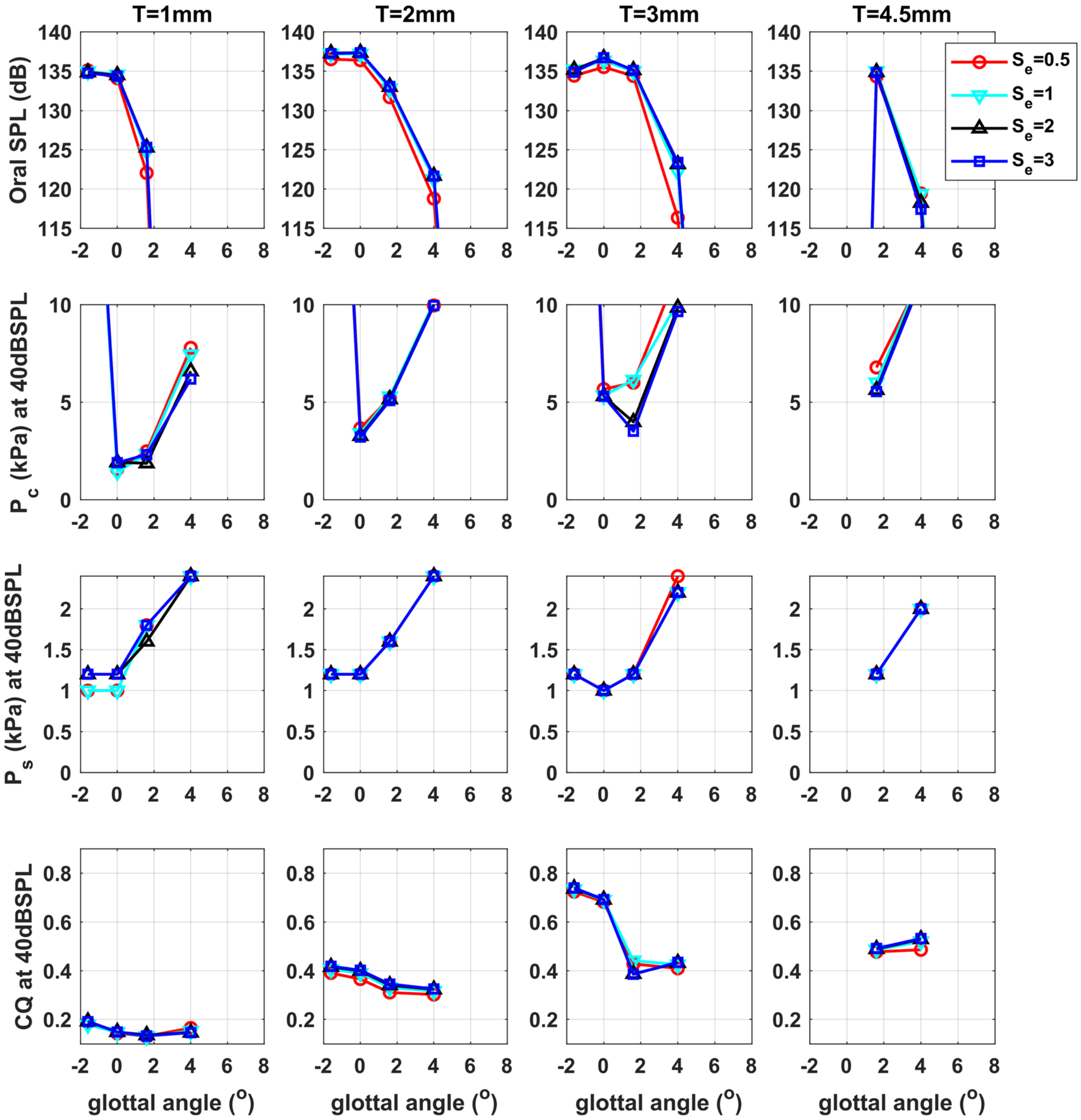

FIG. 7. (Color online) /u/ vocal tract semi-occluded by a 5-mm diameter tube. From top to bottom: the oral SPL produced at a 1 kPa subglottal pressure, and the peak contact pressure P_c and subglottal pressure P_s required to produce a 40 dB SPL outside and the corresponding CQ values as a function of the initial glottal angle, epilaryngeal scaling factor, and medial surface vertical thickness T.

the air inside the vocal tract volume and the acoustic mass of the vocal tract wall, as shown in previous studies (Stevens, 1998; Patel *et al.*, 2019). The contribution of vocal tract wall inertance to the first formant becomes important when the lip constriction provides a higher impedance to the oscillating airflow than the vocal tract wall, as shown in Patel *et al.* (2019). Thus, this insensitivity of first formant and oral SPL to epilaryngeal manipulations is expected to vary depending on both the vocal tract wall properties (e.g., relaxed vs tensed cheeks) and degree of semi-occlusion at the lips (mouth opening area, or the diameter and length of a constricting tube if present), which is worth further investigation.

Our results suggest that semi-occlusion at the lips does not facilitate identification of favorable laryngeal conditions any more than open vocal tracts, at least not in the sense of increased contrast in oral SPL across different laryngeal conditions (Figs. 5 and 8). In fact, Fig. 8 indicates that it is easier to identify favorable laryngeal conditions in the /ɑ/ vocal tract. This is probably because the large vocal tract impedance associated with semi-occlusion may have reduced the effect of laryngeal adjustments on oral SPL. However, semi-occlusion of the vocal tract at the lips does significantly increase both the mean oral pressure and oral SPL (by about 10 dB in this study) compared to an open vocal tract. Thus, it is possible that one main benefit of SOVTE is to, as stated by Titze (2006), familiarize speakers with oral vibratory sensations, increase their acuity to sensation, thus allowing them to better adopt favorable laryngeal configurations.

In this study the oral SPL is used as an indirect measure of oral vibratory sensations. Because the voice spectrum is generally dominated by low-frequency harmonic energy, the

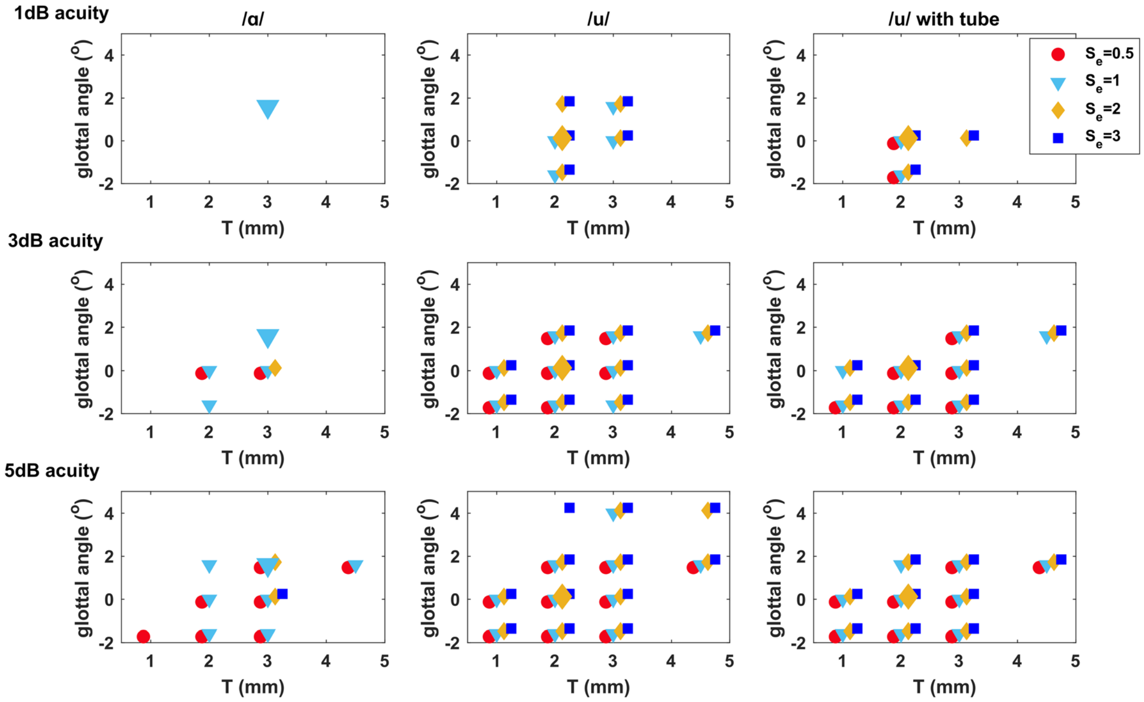

FIG. 8. (Color online) Laryngeal and epilaryngeal conditions that produce an oral SPL within 1 dB (top), 3 dB (middle), and 5 dB (bottom) of the maximal oral SPL for the three vocal tract configurations (/ɑ/ vocal tract, /u/ vocal tract, /u/ vocal tract with a tube). Symbols denote epilarynx with different scaling factors: ●: Sepi = 0.5; ▼: Sepi = 1; ◆: Sepi = 2; ■: Sepi = 3. The symbol with the largest size in each panel denotes the condition producing the maximal oral SPL for the corresponding vocal tract configuration.

oral SPL is largely determined by energy in the low frequency range, particularly around the fundamental frequency, which is also more effective in exciting vocal tract wall vibration due to its low resonance frequency (Sundberg, 1992). In this sense, the oral SPL is a reasonable indirect measure of vocal tract vibration and oral vibratory sensations. Nevertheless, the findings of this study need to be verified in future experiments.

V. CONCLUSIONS

Our results show that emphasis on maximizing oral vibratory sensations in voice therapy would facilitate speakers to adopt an intermediate adduction configuration at the larynx. Maximizing oral vibratory sensations may also lead to epilaryngeal tube narrowing in an open vocal tract, but does not facilitate any consistent vocal tract changes in the epilarynx in a semi-occluded vocal tract, in which the effect of epilaryngeal manipulations on vocal tract resonances (particularly the first formant) is small. In general, changes in oral SPL due to laryngeal and epilaryngeal adjustments are larger in an open vocal tract than a semi-occluded vocal tract. This suggests that an open vocal tract configuration may allow speakers to better adopt favorable laryngeal configurations. However, semi-occlusion does significantly increase the mean and dynamic pressure within the oral cavity, which may familiarize speakers with oral vibratory sensations, increase their acuity to vibratory sensations, and thus prepare them to better adopt favorable laryngeal configurations in a more open and natural vocal tract configuration.

ACKNOWLEDGMENTS

This study was supported by research Grant No. R01DC001797 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health.

Berry, D., Verdolini, K., Montequin, D. W., Hess, M. M., Chan, R. W., and Titze, I. R. (**2001**). “A quantitative output-cost ratio in voice production,” J. Speech. Lang. Hear. Res. **44**, 29–37.

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.

Guzman, M., Castro, C., Testart, A., Muñoz, D., and Gerhard, J. (**2013a**). “Laryngeal and pharyngeal activity during semioccluded vocal tract postures in subjects diagnosed with hyperfunctional dysphonia,” J. Voice **27**(6), 709–716.

Guzman, M., Laukkanen, A. M., Krupa, P., Horáček, J., Švec, J. G., and Geneid, A. (**2013b**). “Vocal tract and glottal function during and after vocal exercising with resonance tube and straw,” J. Voice **27**(4), 523.e19–523.e34.

Guzman, M., Miranda, G., Olavarria, C., Madrid, S., Muñoz, D., Leiva, M., Lopez, L., and Bortnem, C. (**2017**). “Computerized tomography measures during and after artificial lengthening of the vocal tract in subjects with voice disorders,” J. voice **31**(1), 124.e1–124.e10.

Hampala, V., Laukkanen, A., Guzman, M. A., Horacek, J., and Svec, J. G. (**2015**). “Vocal fold adjustment caused by phonation into a tube: A double-case study using computed tomography,” J. Voice **29**(6), 733–742.

Laukkanen, A. M. (**1992**). “About the so called ‘resonance tubes’ used in Finnish voice training practice: An electroglottographic and acoustic investigation on the effects of this method on the voice quality of subjects with normal voice,” Scand. J. Log. Phon. **17**(3-4), 151–161.

Laukkanen, A. M., Horáček, J., Krupa, P., and Švec, J. G. (**2012**). “The effect of phonation into a straw on the vocal tract adjustments and formant frequencies. A preliminary MRI study on a single subject completed with acoustic results,” Biomed. Sign. Process. Control **7**(1), 50–57.

Laukkanen, A. M., Titze, I. R., Hoffman, H., and Finnegan, E. (**2008**). "Effects of a semioccluded vocal tract on laryngeal muscle activity and glottal adduction in a single female subject," Folia Phoniatr. Logop. **60**(6), 298–311.

Lulich, S. M., and Patel, R. R. (**2021**). "Semi-occluded vocal tract exercises in healthy young adults: Articulatory, acoustic, and aerodynamic measurements during phonation at threshold," J. Acoust. Soc. Am. **149**(5), 3213–3227.

Maxfield, L., Titze, I., Hunter, E., and Kapsner-Smith, M. (**2015**). "Intraoral pressures produced by thirteen semi-occluded vocal tract gestures," Log. Phon. Vocol. **40**(2), 86–92.

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.

Patel, R. R., Lulich, S. M., and Verdi, A. (**2019**). "Vocal tract shape and acoustic adjustments of children during phonation into narrow flow-resistant tubes," J. Acoust. Soc. Am. **146**(1), 352–368.

Stemple, J. C., Lee, L., D'Amico, B., and Pickup, B. (**1994**). "Efficacy of vocal function exercises as a method of improving voice production," J. Voice **8**(3), 271–278.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (The MIT Press, Cambridge, MA), Chap. 3.

Story, B. H. (**1995**). "Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract," Ph.D. dissertation, University of Iowa, Iowa City, IA, Chap. 2.

Story, B. H., Laukkanen, A. M., and Titze, I. R. (**2000**). "Acoustic impedance of an artificially lengthened and constricted vocal tract," J. Voice **14**(4), 455–469.

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.

Sundberg, J. (**1974**). "Articulatory interpretation of the singing formant," J. Acoust. Soc. Am. **55**, 838–844.

Sundberg, J. (**1992**). "Phonatory vibrations in singers: A critical review," Music Percept. **9**(3), 361–381.

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., and Story, B. H. (**1997**). "Acoustic interactions of the voice source with the lower vocal tract," J. Acoust. Soc. Am. **101**(4), 2234–2243.

Titze, I. R., and Worley, A. S. (**2009**). "Modeling source-filter interaction in belting and high-pitched operatic male singing," J. Acoust. Soc. Am. **126**(3), 1530–1540.

Vampola, T., Laukkanen, A. M., Horáček, J., and Švec, J. G. (**2011**). "Vocal tract changes caused by phonation into a tube: A case study using computer tomography and finite-element modeling," J. Acoust. Soc. Am. **129**(1), 310–315.

Verdolini, K., Druker, D., Palmer, P., and Samawi, H. (**1998**). "Laryngeal adduction in resonant voice," J. Voice **12**, 315–327.

Verdolini-Marston, K., Burke, M. D., Lessac, A., Glaze, L., and Caldwell, E. (**1995**). "A preliminary study on two methods of treatment for laryngeal nodules," J. Voice **9**, 74–85.

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. (**2017**). "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. (**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**(2), 1039–1050.

Zhang, Z. (**2021a**). "Interaction between epilaryngeal and laryngeal adjustments in regulating vocal fold contact pressure," JASA Express Lett. **1**(2), 025201.

Zhang, Z. (**2021b**). "Vocal tract adjustments to minimize vocal fold contact pressure during phonation," J. Acoust. Soc. Am. **150**, 1609–1619.

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.