Vocal Fold Vertical Thickness in Human Voice Production and Control: A Review

Zhaoyan Zhang, *Los Angeles, CA*

Summary: While current voice research often focuses on laryngeal adjustments in a two-dimensional plane from a superior endoscopic view, recent computational simulations showed that vocal control is three-dimensional and the medial surface vertical thickness plays an important role in regulating the glottal closure pattern and the spectral shape of the produced voice. In contrast, while a small glottal gap is required to initiate and sustain phonation, further changes in the glottal gap within this small range have only small effects on glottal closure and spectral shape. Vocal fold stiffness, particularly along the anterior-posterior direction, plays an important role in pitch control but has only a small effect on glottal closure and spectral shape. These results suggest that voice research should pay more attention to medial surface shape in the vertical dimension. Future studies in a large population of both normal speakers and patients are needed to better characterize the three-dimensional medial surface shape, its variability between speakers, changes throughout the life span, and how it is impacted by voice disorders and clinical interventions. The implications for voice pedagogy and clinical intervention are discussed.

Key Words: Cause-effect relationship−Medial surface vertical thickness−Glottal closure−Voice quality−Vocal register−Clinical intervention−Voice pedagogy.

INTRODUCTION

Humans are able to produce a great variety of voice types differing in pitch, loudness, and voice quality. This is achieved through muscular control of the geometry and mechanical properties of the vocal folds which, together with the respiratory support and adjustments in the vocal tract, regulates vocal fold vibration and the produced voice. An important goal of voice research is to establish a cause-effect understanding of voice production that links the biomechanical properties of the vocal system (vocal fold geometry, stiffness, position, and subglottal pressure) and the produced voice outcomes.

While ideally the cause-effect relationship should be established in humans or animal models with comparable vocal physiology, there are challenges in establishing a causal understanding of voice production in these models. It is currently difficult to reliably control and measure vocal fold geometry and stiffness in human or animal models, thus making it difficult to establish cause-effect relationship between geometry and stiffness and the produced voice. More importantly, vocal fold geometry, including length, depth, and thickness, often co-vary with each other and with vocal fold stiffness in these models so that it is almost impossible to isolate effects of individual controls. As a result, when voice changes occur in these models, often we are not sure whether the voice changes are due to changes in vocal fold geometry, changes in stiffness, or a combination of both.

Because of these limitations, voice research in human and animal models is often limited to a study of correlation instead of causation. Correlation is often made between either different voice outcome measures (acoustics, aerodynamics, or vibration) in different conditions of pitch, loudness, or voice quality, or between voice outcome measures and a few control parameters of the vocal system that can be relatively easily measured in humans (eg glottal gap, subglottal pressure). The effects of control parameters of the vocal system that cannot be directly measured in humans are often neglected.

One such neglected control parameter in voice research and clinical intervention is the medial surface shape of the vocal folds. Because it is difficult to measure medial surface shape in humans, currently voice research and clinical intervention often focus on vocal fold vibration and glottal closure in a two-dimensional plane as viewed from above (see eg[1,2] and a more recent review[3]). While it is generally assumed that the degree of vocal fold adduction determines the duration of glottal closure during phonation, in this two-dimensional view, the degree of vocal fold adduction is often evaluated based on the glottal gap in the horizontal plane alone. The effect of changes in medial surface shape in the vertical dimension is often neglected.

Considering that the glottal airflow-vocal fold interaction occurs primarily on the medial surface,[4,5,6] one would expect the medial surface to play an important role in voice production and vocal control. Indeed, the effect of medial surface shape on vocal registers is well appreciated in the literature on vocal pedagogy (see eg[7-15]). van den Berg[7] argued that voices in the chest register are often produced with thick vocal folds, whereas voices in falsetto are often produced with thin vocal folds. However, the effect of medial

Accepted for publication February 16, 2023.
Presented at the 51st Voice Foundation Annual Symposium: Care of the Professional Voice, Philadelphia, PA.
This study was supported by research grants R01 DC001797 and R01 DC020240 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health.
From the UCLA School of Medicine, Los Angeles, CA.
Address correspondence and reprint requests to Zhaoyan Zhang, UCLA School of Medicine, 31-24 Rehabilitation Center, 1000 Veteran Ave, Los Angeles, CA 90095-1794 E-mail: zyzhang@ucla.edu
Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997

https://doi.org/10.1016/j.jvoice.2023.02.021

surface thickness on voice production has yet to be systematically investigated in humans, due largely to experimental challenges as mentioned above.

In his seminal paper on the body-cover theory of phonation, Hirano[16] discussed four voice types differentiated by the stiffness conditions in the body and cover layers. Although he did not explicitly mention medial surface shape, it is quite clear that the four voice types also differ in the medial surface shape in the illustrations he used to describe the four voice types (Figure 2 in[16]). Thus, one may question whether the differences between the four voice types are due to differences in vocal fold stiffness alone or whether the differences in vocal fold thickness also play some role. As mentioned above, it is almost impossible to answer this question in human or animal models due to the co-variations of stiffness and geometry in these models.

For this reason, we turned to computational modeling in order to isolate the effects of individual vocal control parameters. Computational modeling allows systematic manipulation of control parameters and is thus better suited for cause-effect investigations. In particular, it allows us to change one control at a time so that we can be certain that the observed voice changes are due to changes in one specific control only, not changes in other controls.

In the past few years, we have conducted a series of large-scale, three-dimensional simulations of voice production,[17-21] with systematic parametric variations in vocal fold geometry (anterior-posterior length, medial-lateral depth, and inferior-superior thickness), vocal fold approximation or the prephonatory glottal gap in the horizontal plane, vocal fold stiffness in both the body and cover layers and along both the longitudinal and transverse directions, subglottal pressure, and vocal tract shape, for a total of about 300,000 vocal fold/vocal tract conditions. With this many conditions, we were able to systematically investigate the individual effects of vocal fold geometry, stiffness, vocal fold approximation, and subglottal pressure on vocal fold vibration, aerodynamics, acoustics (F0, intensity, and spectral shape), voice types (modal and non-modal voices), and vocal fold contact pressure during vocal fold collision.

In this paper the major cause-effect relationships identified in these computational studies are summarized. We focus on the regulation of the glottal closure pattern (duration of glottal closure and flow declination in the closing phase) and spectral shape, both of which are highly correlated with the perception of voice quality and as a result are of clinical and pedagogical interest. In particular, the simulations showed a dominant effect of vocal fold medial surface shape in the vertical dimension on glottal closure and spectral shape. The implications of these findings for voice research, voice pedagogy, and clinical intervention will be discussed.

Major findings from computational simulations

Figure 1 shows the vocal fold model used in our simulation studies. The model is three-dimensional and based on continuum mechanics, which allows realistic representation of the physics involved in phonation. Our model also includes a respiratory system and a vocal tract so that interaction between the three subsystems of voice production is inherently included. Details of the model formulation can be found in our previous studies.[22,17,18] The vocal fold geometry in Figure 1 is simplified from realistic human larynges. This simplification is necessary in order to parameterize vocal fold geometry using a small number of geometric controls, thus making it practical for parametric voice simulations. However, our recent studies showed that the major findings from this simplified geometry remain the same in models with realistic vocal fold geometry.[23]

Simulations were performed with parametric variations in nine physiological control parameters of the vocal system as well as vocal tract shape.[17-21] These include four geometric measures of the vocal folds (length L, thickness T, and

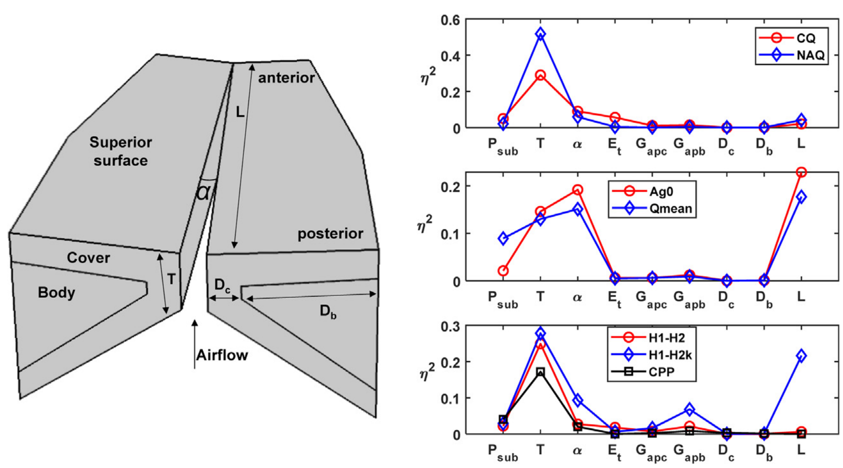

FIGURE 1. Left: The three-dimensional vocal fold model used in the computational studies. Right: the effect sizes (η^2) of the nine physiological control parameters on the glottal closure pattern (CQ and NAQ), mean glottal opening area (Ag0) and mean glottal flow (Qmean), and voice acoustic measures (H1-H2, H1-H2k, CPP). See text for the symbols for the control parameters. Data from.[21]

depths of the body layer D_b and cover layer D_c), three stiffness measures of the vocal folds (transverse stiffness E_t in the coronal plane, anterior-posterior (AP) stiffness in the body layer G_{apb} and cover layer G_{apc}), initial (prephonatory) glottal angle α quantifying the degree of vocal fold approximation in the horizontal plane, and subglottal pressure P_{sub}. Note that in this model the glottal gap is zero at the anterior end and, for a positive glottal angle, increases along the anterior-posterior direction toward the vocal processes. In order to focus on cause-effect relationships at the laryngeal level, in the following results will be presented from simulations without a vocal tract, followed by a brief summary of the results from simulations involving source-tract interaction.

Figure 1 shows the effect sizes of the nine physiological control parameters on selected outcome measures.[21] The effect sizes were calculated as the percentage of total variance in the outcome measure of interest that was explained by individual control parameters. The selected outcome measures include the closed quotient (CQ), normalized amplitude quotient (NAQ, a normalized measure of the maximum flow declination rate or closing quotient of the glottal flow;[24]), mean glottal area opening (Ag0), mean glottal flow (Qmean), amplitude differences between the first harmonic and the second harmonic (H1-H2) and the harmonic nearest 2kHz (H1-H2k), and cepstral peak prominence (CPP,[25]) of the produced voice. Figure 1 shows that the vertical thickness T has the largest, dominant effect, among all physiological controls, on both CQ and NAQ, two important measures of glottal closure during phonation. The prephonatory glottal gap α, stiffness E_t, length L, and subglottal pressure P_{sub} have some effects, but these effects are smaller and less consistent than that of vertical thickness.

Thus, in contrast to the general assumption that the glottal gap plays an important role in regulating the glottal closure pattern, our simulations showed that the glottal closure pattern is determined primarily by the vertical thickness of the medial surface (Figure 2), instead of the glottal gap. While the prephonatory glottal gap has to be sufficiently small to initiate and sustain phonation, further changes within this small-gap range have only small effects on the closure pattern except near phonation onset.[17] Reducing the glottal gap does bring the vocal folds closer to the glottal midline, but it also reduces the vibration amplitude of the vocal folds. These two effects often cancel out each other, reducing the overall effect of changes in glottal gap on the glottal closure pattern. Similarly, increasing AP stiffness in very soft vocal folds improves glottal closure, but further increase in stiffness often does not produce any further improvement[6,17].

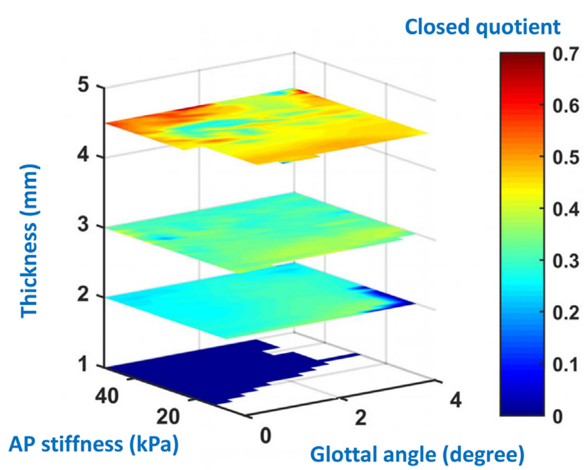

FIGURE 2. The closed quotient is primarily controlled by the vertical thickness. The thicker the vocal folds, the larger the closed quotient. The effect of vocal fold anterior-posterior (AP) stiffness and glottal angle is small. Adapted from.[17]

Our simulations also showed that a small glottal gap does not always guarantee complete glottal closure during phonation. In addition to sufficient vocal fold approximation, the vocal fold medial surface has to be sufficiently thick in order to achieve complete glottal closure during phonation.[17] If the vocal folds are too thin, they are pushed open by airflow, even if the glottis is completely closed at rest. They still vibrate, but are not able to come back to completely close the glottis, as sometimes observed in excised larynx experiments.[26]

Because of this importance of vertical thickness in maintaining prephonatory adductory position of the vocal folds against airflow, vertical thickness T also has the second largest effect size on the mean glottal opening area and mean glottal flow (Figure 1), only slightly smaller than that of the initial glottal angle. Thus, increasing vertical thickness is an important strategy in maintaining glottal closure and conserving airflow consumption against increasing subglottal pressure.[27] This is particularly the case at high subglottal pressures, at which increasing vertical thickness becomes the most effective means of conserving airflow.[27,28] This indicates the importance of vocal fold thickening in loud voice production and its potential role in hyperadduction (see more discussion in[27]).

The medial surface thickness also determines the duration of glottal closure during phonation. The thicker the folds, the larger the closed quotient (Figure 2;[17,21]). Extremely thin vocal folds often vibrate in an in-phase, up-and-down motion, similar to a reed in some musical instruments. With sufficient thickness, the vocal folds are able to sustain a wave propagating along the medial surface, resulting in a vertical phase difference along the medial surface. This vertical phase difference means that when the inferior glottis starts to open, the superior glottis would still remain closed for a while. Thus, the thicker the vocal folds, the larger the vertical phase difference, and the longer the glottis remains closed during phonation.[17]

This dominant effect of vertical thickness on glottal closure leads to similar dominant effects of the vertical thickness on spectral measures (eg H1-H2, H1-H2k) as well as CPP. All three measures are primarily controlled by medial surface vertical thickness, with thicker vocal folds generally producing a lower H1-H2 (eg a weaker first harmonic) and stronger harmonic excitation at high frequencies (2k-5kHz).

The simulations also showed that when the vocal folds are too thick and too soft (low transverse stiffness), they are more likely to exhibit irregular vibration such as subharmonics or chaos, particularly when tightly approximated (small glottal gap) and at high subglottal pressures.[19] When the pressure is low, vocal folds that are thick and soft often produce excessively long glottal closure at low frequencies, or vocal fry. In general, thicker vocal folds experience higher vocal fold contact pressure and thus higher risk of vocal fold injury.[20,29] Thickness also has a moderate effect on the phonation threshold pressure, which is generally the lowest at intermediate vertical thickness, and increases when the vocal folds are either too thin or too thick (Figure 2 in[18]; Table V in[21]).

Vocal fold stiffness plays a large role in controlling the fundamental frequency of phonation and phonation threshold pressure.[21] However, its effects on the closed quotient (CQ) and normalized amplitude quotient (NAQ) are small,[17,18] except for the transverse stiffness in the coronal plane which has a moderate effect on both CQ and H1-H2.[18,21] Increasing body-cover ratio in the transverse stiffness also tends to increase the vertical phase difference and facilitate a more wave-like motion along the medial surface.[30]

Vocal tract adjustments in general have small effects on the voice source. Our recent simulations (to be published) showed that narrowing the vocal tract tends to lower the CQ and increase H1-H2, but the effect sizes are small (changes in CQ in the order of 0.05, and about 1-2 dB change in H1-H2). However, vocal tract adjustments do affect the formant structure and thus may have significant effect on the intensity and spectral shape of the produced voice outside the mouth. Vocal tract adjustments that increase the output vocal intensity often increase the peak vocal fold contact pressure. Extreme constrictions of the vocal tract also reduce the mean glottal flow and glottal flow amplitude.[29,31]

To summarize, our simulations have shown that the vertical dimension of the vocal fold medial surface has a dominant role in the control of glottal closure pattern and spectral shape. The vocal fold thickness has to be just right, not too thin or too thick (Figure 3). If too thin, the vocal folds will vibrate with incomplete closure, high airflow, and a smooth airflow waveform, which leads to high H1-H2 and weak harmonics at high frequencies. If too thick, the vocal folds will vibrate with long glottal closure, low airflow, skewed airflow waveform, low H1-H2, strong high-frequency harmonics, and sometimes irregular vibration and high risk of vocal fold injury.

Similar findings on the important role of vocal fold thickness in vocal control have been reported in recent studies from other research groups (eg[32,33,34]). In other studies in which the thickness was kept constant (eg[35]), variations of other control parameters (eg stiffness) led to only small variations in the closed quotient, consistent with the findings just reviewed. The importance of vertical thickness in regulating the glottal flow has been reported in.[28,36]

Sensitivity of voice production to subtle, realistic changes in medial surface shape

The vocal fold geometry in Figure 1 is simplified compared to realistic vocal folds. In particular, realistic larynges do not have a well-defined medial surface or vertical thickness. The medial surface in humans also goes through subtle, complex shape changes due to laryngeal muscle activation.

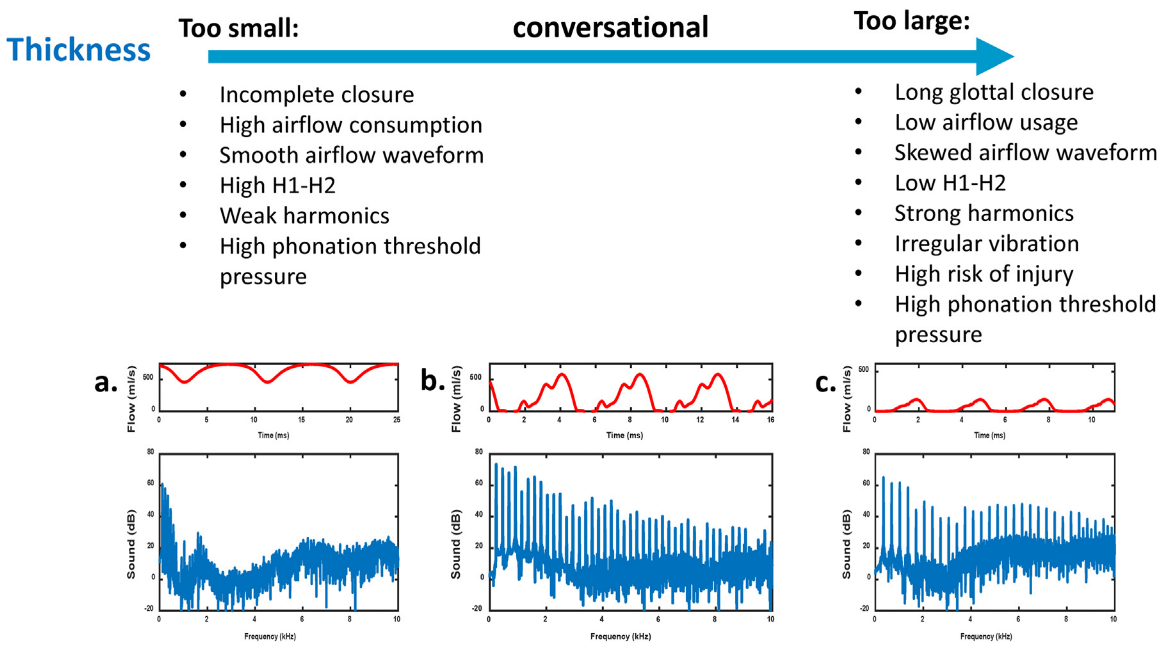

FIGURE 3. Medial surface vertical thickness plays an important role in regulating glottal closure pattern and spectral shape of the produced voice. The bottom row shows typical waveforms of the glottal flow and sound source spectra when the vocal fold thickness is too small **A**. just right as in conversational voice **B**. and too large **C**.

To investigate whether voice production is sensitive to such realistic, subtle changes in medial surface shape, we repeated the computational studies in vocal fold models with geometry taken from magnetic resonance imaging (MRI) of human larynges.[23] Our goal was to systematically manipulate medial surface shape and observe its impact on voice production. If the same observation applies as in our simulations with simplified vocal fold geometry, we would expect subtle changes in medial surface shape to have large impact on voice outcomes, particularly the glottal closure pattern. Based on observations from MRI studies and *in vivo* canine experiments,[37,38] we developed empirical rules that allow us to parametrically vary medial surface shape, by introducing medial bulging at the superior and inferior portions of the medial surface.

Figures 4b-4c show the effect of medial surface shape manipulation on voice production.[23] Each data point or square in the figure corresponds to a unique medial surface shape, some of which are shown in Figure 4a. The middle figure shows the phonation threshold pressure for each unique medial surface shape. The figure on the right shows the closed quotient for each medial surface shape at four subglottal pressures. Although the changes in medial surface shape are small (in the order of 0.5 mm), the impact on voice is large. For example, changes in medial surface shape can increase phonation threshold pressure from 200 Pa to 2000 Pa, and CQ can vary from 0.1 to 0.9. This confirms that voice production is highly sensitive to subtle changes in medial surface shape.

It is interesting to note that the impact of medial surface shape manipulation is larynx-specific. When the same simulations were repeated in a different larynx that has a thinner medial surface than the larynx used in Figure 4, the same medial surface shape manipulations produced a much smaller range of variation in the closed quotient. The variation pattern was also slightly different.[23] This suggests that the resting medial surface shape at least partially determines individual speakers' vocal capabilities, as discussed further below.

The results from this study[23] also provide insight on how to determine the effective vertical thickness for a smoothly varying medial surface. It appears that the effective thickness is primarily determined by the thickness of the most medial portion of the medial surface that forms a straight or almost-straight glottis in the coronal plane.[23] This observation is partially supported by findings in recent computational fluid dynamics studies in[28] and,[36] which showed that the medial surface vertical thickness had a large effect on the glottal flow only when the medial surface forms a straight, non-convergent, glottal shape in the vertical plane.

Manipulation of medial surface shape in excised larynges

In an attempt to validate the findings from our simulations, we manipulated medial surface shape in excised human larynges and observed its impact on voice production.[39] During the experiment, we opened a thyroplasty window on the thyroid cartilage and used a wooden stick to manipulate medial surface shape while maintaining similar amount of vocal fold approximation as viewed from above. Depending on whether the wooden stick was directed toward the superior or inferior portion of the medial surface, we could either decrease or increase medial surface vertical thickness while keeping the glottal gap constant. Figure 5 shows images of vocal fold vibration as viewed from above during one cycle of vocal fold vibration for four thickness conditions. The vocal folds vibrated with incomplete glottal closure when the wooden stick was directly toward the superior medial surface, creating a thin medial surface. The glottal closure significantly improved as the wooden stick was gradually directly toward the inferior portion of the medial surface, and vocal folds were able to vibrate with a long period of glottal closure. This result indicates that even if the vocal folds appear to be sufficiently medialized when viewed from above, the glottal closure pattern can vary significantly, depending on the medial surface shape.

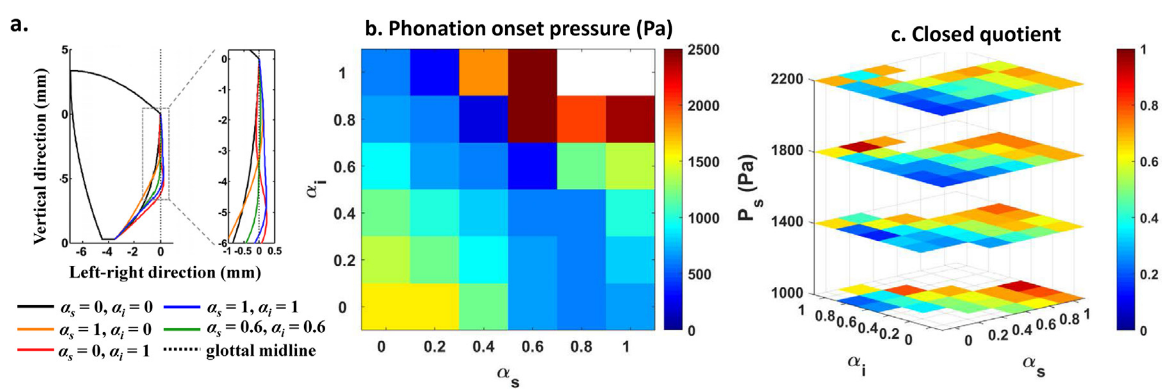

FIGURE 4. Voice production is sensitive to subtle changes in medial surface shape. In an MRI-based vocal fold model, the medial surface shape is parametrically manipulated to simulate superior-medial bulging and inferior-medial bulging according to observations from *in vivo* canine experiments. The degrees of superior and inferior medial bulging are controlled by parameters α_s and α_i, respectively **A**. Although changes in medial surface shape are small, they have significant influence on both **B** the phonation threshold pressure and **C** closed quotient. (Adapted from[23]).

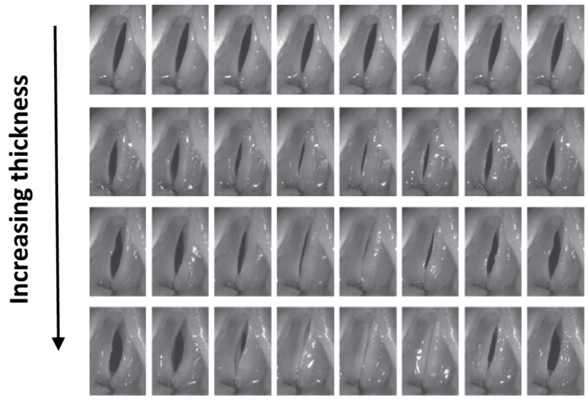

FIGURE 5. Superior view of vocal fold vibration at different conditions of vocal fold thickness. Each row shows a sequence of superior images of vocal fold vibration over one cycle of vocal fold vibration. Thin vocal folds vibrate with incomplete closure despite being fully medialized at rest. Increasing thickness improves glottal closure. (Adapted from[39]).

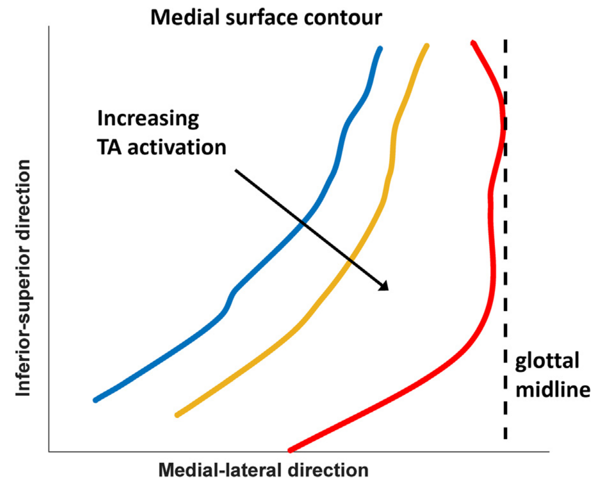

FIGURE 6. Increasing activation of the thyroarytenoid muscle not only medializes the vocal fold, which reduces membranous glottal gap, but also increases vocal fold vertical thickness.

The experimental procedure to manipulate medial surface in this study is similar to the procedure surgeons often apply during medialization surgery, in which surgeons often poke around in the thyroplasty pocket to identify the location that gives the best voice outcomes. It is possible that this poking-around process allows surgeons to find the best location for implant insertion that would restore optimal medial surface shape.

Muscular control of medial surface shape

Regulation of medial surface shape in humans involves both the intrinsic and extrinsic laryngeal muscles as well as the respiratory system (ie tracheal pull). In a series of experiments, Hirano and colleagues electrically stimulated excised canine larynges and investigated the effect on vocal fold shape.[40,41,42] They showed that activation of the thyroarytenoid muscles caused the inferior portion of the medial surface to bulge toward the glottal midline, thus increasing the medial surface vertical thickness, whereas activation of the cricothyroid muscles reduced the vertical thickness. A small effect on vocal fold thickness by actions of the lateral cricoarytenoid muscles was also observed. Similar observations were reported in more recent studies.[37,38]

Since the same set of laryngeal intrinsic muscles also regulates vocal fold stiffness, changes in medial surface shape in humans are often accompanied by changes in stiffness. Vocal fold thickening is often accompanied by reduced stiffness and tension in the vocal fold cover layer, whereas vocal fold thinning is often accompanied by increased stiffness and tension in the cover layer.

These *in vivo* canine experiments also showed that thickness and glottal gap often co-vary, particularly at mid-membranous locations,[43,44,45] as shown in Figure 6. These studies showed that while the lateral cricoarytenoid and interarytenoid muscles are responsible for gross control of the membranous glottal gap (from breathy to modal), the thyroarytenoid muscle plays the role of a fine controller of the membranous glottal gap, particularly in voice quality variations from modal to pressed. Thus, changes in the medial surface thickness at membranous locations are often accompanied by changes in the membranous glottal gap (Figure 6). This co-variation may partially explain why the glottal gap is often observed to co-vary with the glottal closure pattern and voice quality despite a small effect observed in our simulation studies.

Actions of the extrinsic laryngeal muscles, together with tracheal pull, often change the relative positions of the thyroid, cricoid, and arytenoid cartilages, and may also change vocal fold medial surface shape. In general, muscle activities that shorten the vocal folds are likely to increase vocal fold thickness. Thus, vocal fold thickness may also be increased by tilting the thyroid cartilage upward and backward or tilting the cricoid forward and downward, and decreased by tilting the thyroid or cricoid cartilage in the opposite direction.[46,47,15] Rise of the larynx increases the folding of the vocal folds and supraglottal structures and often leads to tighter adduction of the vocal folds and likely increases vocal fold thickness.[47,48] Lowering the larynx on the other hand tends to open the glottis and reduce the vertical thickness.

However, this is not always the case, and the opposite has been observed on the effect of extrinsic muscles, particularly in trained singers. Currently there is little experimental data on how actions of the extrinsic muscles impact vocal fold posturing, especially the medial surface shape. This lack of understanding is not surprising considering that the many extrinsic muscles often work together to posture the larynx, thus making it difficult to isolate the effects of individual muscles in human subjects studies. Future experimental studies that allow improved control of individual muscles are required to provide a more complete picture of extrinsic muscular control of medial surface shape.

Summary: vocal control is three-dimensional

The results above show that vocal control at the laryngeal level is three-dimensional. Specifically, the medial surface shape in the vertical dimension plays a dominant role in regulating the glottal closure pattern and spectral shape of the produced voice, and voice is sensitive to subtle changes in medial surface shape. Thus, voice quality control is mainly achieved by thickness control at the laryngeal level and complemented by vocal tract adjustments.

This important role of vertical thickness in vocal control may have evolutionary origins. The main function of the larynx is to close the glottis and protect the lungs. While the glottis can be closed by medializing a pair of thin vocal folds, thicker vocal folds would provide more effective airway protection, particularly when the lung pressure is high as in emotional situations. Indeed, protection of the lungs often involves closure of the vocal folds as well as other supraglottal structures (false folds, aryepiglottic fold, and epiglottis). It is likely that this thickness-regulating mechanism (for both vocal folds and supraglottal structures) was inherited during the emergence and evolution of speech for the regulation of voice quality. This view is supported by the finding from a recent study that many acoustic measures that are controlled by thickness (eg H1-H2, H1-H2k, CPP) consistently appear in the acoustic space of natural human voices, independent of speakers and languages.[49]

Implications for voice research

The importance of vocal fold thickness suggests that in addition to glottal gap from the superior view, voice research should also pay attention to changes in medial surface shape in the vertical dimension, particularly when changes in voice quality are observed. Because of the covariations between thickness, glottal gap, and vocal fold stiffness in human or animal models, caution should be taken in interpreting results from these models, particularly with respect to causal relationships in human voice production. For example, as discussed above (Figure 6), vocal fold adduction not only medializes the vocal folds but also induces simultaneous, subtle changes in the medial surface shape along the vertical dimension. Thus, while changes in voice production may appear to be correlated with changes in the glottal gap from a superior view, the observed changes in voice production may actually be caused by accompanying changes in vocal fold thickness.

Similarly, previously we raised the question whether the differences between the four voice conditions (soft, loud, modal, falsetto) discussed in Hirano's body-cover theory are due to differences in the stiffness condition or differences in medial surface shape. While Hirano's discussion focused more on the stiffness conditions, the results from our simulations indicate that the differences in the glottal closure pattern and voice quality between these four voice conditions likely result from differences in medial surface shape. Although changes in the stiffness conditions are observed to correlate with production of different voice types, there is no direct cause-effect relation between vocal fold stiffness and voice types. The correlation may simply result from the fact that both vocal fold stiffness and geometry are regulated by the same set of laryngeal muscles.[18]

Implications for vocal registers and voice pedagogy

The large effect of vertical thickness observed in our simulations supports van den Berg's hypothesis that thickness plays an important role in regulating voice quality at different vocal registers. Vocal fry and chest-like voices are likely produced with thicker vocal folds, whereas head- or falsetto-like voices are produced by thinner folds. Because thickness co-varies with stiffness in humans, it is not surprising that different voice qualities may be more easily produced at different pitch ranges. For example, vocal fry and chest-like voices are more easily produced at lower pitches whereas head- or falsetto-like voices are more easily produced at higher pitches.

Male vocal folds are often thicker than female vocal folds. Thus, male voices, particularly the speaking voice, are often produced with a longer glottal closure phase and stronger high-frequency harmonics than female voices,[50,51,21] and may be more easily produced in the chest register than female voices.

By gradually varying the vocal folds from thick to thin, one should be able to smoothly transition from a chest-like voice (long closure phase and strong high-frequency harmonics) to a head-like voice (incomplete glottal closure and weak high-frequency harmonics), thus avoiding the perception of a sudden change in voice quality or register change. However, a sudden change in voice quality or register change may occur if actions of the intrinsic and extrinsic laryngeal muscles result in a sudden change in thickness, as one attempts to achieve desired pitch or intensity goals, particularly in untrained singers.

Theoretically, a speaker should be able to produce a voice with strong high-frequency harmonics at any pitch in the range if sufficient thickness can be maintained. Thus, it is reasonable to expect that all else being equal, speakers with thicker vocal folds would be better able to maintain sufficient thickness and produce strong high-frequency harmonics at high pitches than speakers with thinner vocal folds. Whether male or female, one may be endowed with naturally thick vocal folds and will be able to more easily produce a fuller voice at high notes. In contrast, speakers with naturally thin vocal folds will be able to produce a lighter clear voice more easily at low notes. On the other hand, one could also nurture versatility of the vocal folds through exercises targeting the thyroarytenoid muscles that strengthen isolated control of the vocal folds from thick to thin. Such an approach has been shown to be successful in training singers to produce a fuller voice at high pitches, particularly in female singers[10], and a myriad of voice qualities at all pitches in male and female singers.[15]

At higher pitches when it is difficult to maintain sufficient vocal fold thickness, one may compensate for the weak harmonic production at the laryngeal level with vocal tract adjustments, for example epilaryngeal narrowing, which boosts harmonics in the 2-3 kHz range (eg[52,53]).

Implications for clinical intervention

The importance of medial surface thickness in regulating glottal closure and the mean glottal flow during phonation suggests that hypo- and hyper-adduction should be evaluated considering changes in both the glottal gap in the horizontal plane and medial surface thickness in the vertical dimension. In our simulations, vocal fold vibration and voice production typical of hyperadduction (eg low airflow, long closure, and possibly irregular vibration) often are observed only in thick vocal folds.

The excised larynx experiment in[39] showed that vocal folds that are not sufficiently thick may vibrate with incomplete closure, even if the vocal folds appear to be sufficiently medialized when viewed from above (Figure 5). While medial surface shape often co-varies with glottal gap and stiffness in healthy speakers, these covariations may be weakened due to pathology, as for example in glottal insufficiency due to either vocal fold paralysis, paresis, or presbylaryngis. Under such situations, clinical intervention needs to restore both the desired glottal gap and optimal medial surface shape. If the medial surface shape is not properly controlled during intervention, the same degree of medialization can result in highly variable voice outcomes, depending on medial surface shape. Restoring sufficient vertical thickness of the medial surface further enhances interaction between the two folds and improves glottal closure, and thus is able to compensate for the negative effect of left-right stiffness asymmetry, a condition that often accompanies glottal insufficiency.[54]

CONCLUSIONS

While current voice research often focuses on laryngeal adjustments in a two-dimensional plane from a superior endoscopic view, our computational simulations showed that vocal control is three-dimensional. In particular, the medial surface shape in the vertical dimension plays an important role in regulating the glottal closure pattern and spectral shape of the produced voice. In contrast, while a small glottal gap is required to initiate and sustain phonation, further changes in the glottal gap within this small range have only small effects on glottal closure and spectral shape (and thus on voice quality). Vocal fold stiffness, particularly stiffness along the anterior-posterior direction, has a large effect on pitch control but only a small effect on glottal closure and spectral shape. These results suggest that voice research should pay more attention to changes in medial surface shape in addition to glottal gap from a superior view.

Early imaging studies often focused on medial surface shape at a few coronal sections (eg[7,55,56]). More recent imaging studies show that the medial surface shape varies significantly along the anterior-posterior direction.[23,57,58] Such anterior-posterior variation in medial surface shape has been shown to have a considerable effect on the resulting glottal closure pattern, with the closed quotient likely determined by the thinnest cross-sections of the vocal folds.[23] Future studies in a large population of both normal speakers and patients are needed to better characterize the three-dimensional medial surface shape, its variability between speakers, changes throughout the life span, and how it is impacted by voice disorders and clinical interventions.

Although the medial surface is often hidden from the superior view, its vertical thickness may be inferred from the phase difference between the upper and lower margins of the medial surface, which can be measured from endoscopic recordings from a superior view, particularly during the closing phase of vocal fold vibration. Slightly changing the angle of the distal tip of the endoscope may also provide a better view of the medial surface.

A more quantitative evaluation of the three-dimensional medial surface can be achieved using computed tomography or magnetic resonance imaging. Currently these imaging methods do not provide sufficient temporal solution to visualize changes in medial surface shape within one cycle of vocal fold vibration. However, it provides information about the medial surface shape averaged over many phonation cycles, and thus can generate insights into the gross muscular adjustments of medial surface shape made to modulate voice production. More recently, Fisher et al.[59] were able to measure changes in medial surface within one cycle of vocal fold vibration, by applying a rapidly switched phase encoding gradient along the direction of motion. Another method of potential is optical coherence tomography,[60,61] which has high temporal and spatial resolutions but is limited by depth of penetration. Further technological developments will surely deepen our insights into the cause-effect relationships in human voice production and control.

Acknowledgments

The author would like to thank Jody Kreiman, Kimberly Steinhauer, Lisa Popeil, and Ronald Scherer for their comments on an earlier draft of this paper.

REFERENCES

1. Laver J. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press; 1980. Chap. 3.
2. Klatt DH, Klatt LC. Analysis, synthesis and perception of voice quality variations among male and female talkers. *J Acoust Soc Am*. 1990;87:820–856.
3. Garellek M. The phonetics of voice. *The Routledge handbook of phonetics*. Routledge; 2019:75–106.
4. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am*. 1988;83:1536–1552.
5. Berry DA, Montequin DW, Tayama N. High-speed digital imaging of the medial surface of the vocal folds. *J Acoust Soc Am*. 2001;110:2539–2547.
6. Zhang Z. Mechanics of human voice production and control. *J Acoust Soc Am*. 2016;140(4):2614–2635.

7. van den Berg JW. Register problems. *Ann N Y Acad Sci*. 1968;155:129–134.
8. Hollien H. On vocal registers. *J Phon*. 1974;2:125–143.
9. Sundberg J, Hogset C. Voice source differences between falsetto and modal registers in counter tenors, tenors and baritones. *Logopedics Phoniatrics Vocol*. 2001;26:26–36.
10. Austin SF. Treasure "chest"-a physiological and pedagogical review of the low mechanism. *J Singing*. 2005;61:241–251.
11. Henrich DN. Mirroring the voice from Garcia to the present day: Some insights into singing voice registers. *Logop Phoniatr Vocol*. 2006;31:3–14.
12. Roubeau B, Henrich N, Castellengo M. Laryngeal vibratory mechanisms: the notion of vocal register revisited. *J Voice*. 2009;23:425–438.
13. Herbst CT. Registers-The Snake Pit of Voice Pedagogy Part 1: Proprioception, perception, and laryngeal mechanisms. *J Sing*. 2020;77:175–190.
14. Herbst CT. Registers−The Snake Pit of Voice Pedagogy: PART 2: Mixed voice, vocal tract influences, individual teaching systems. *J Sing*. 2021;77:345–359.
15. Steinhauer K, Klimek MM, Estill J. *The Estill Voice Model: Theory and Translation*. Pittsburgh, Pennsylvania: Estill Voice International; 2017.
16. Hirano M. Morphological structure of the vocal fold and its variations. *Folia Phoniatr*. 1974;26:89–94.
17. Zhang Z. Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *J Acoust Soc Am*. 2016;139:1493–1507.
18. Zhang Z. Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model. *J Acoust Soc Am*. 2017;142:2311–2321.
19. Zhang Z. Vocal instabilities in a three-dimensional body-cover phonation model. *J Acoust Soc Am*. 2018;144:1216–1230.
20. Zhang Z. Laryngeal strategies to minimize vocal fold contact pressure and their effect on voice production. *J Acoust Soc Am*. 2020;148:1039–1050.
21. Zhang Z. Contribution of laryngeal size to differences between male and female voice production. *J Acoust Soc Am*. 2021;150:4511–4521.
22. Zhang Z. Regulation of glottal closure and airflow in a three-dimensional phonation model: Implications for vocal intensity control. *J Acoust Soc Am*. 2015;137(2):898–910.
23. Wu L, Zhang Z. Voice production in a MRI-based subject-specific vocal fold model with parametrically controlled medial surface shape. *J Acoust Soc Am*. 2019;146:4190–4198.
24. Alku P, Backstrom T, Vilkman E. Normalized amplitude quotient for parametrization of the glottal flow. *J Acoust Soc Am*. 2002;112:701–710.
25. Hillenbrand JM, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. *J Speech Hear Res*. 1994;37:769–778.
26. Zhang Z. Restraining mechanisms in regulating glottal closure during phonation. *J Acoust Soc Am*. 2011;130:4010–4019.
27. Desjardins M, Verdolini Abbott K, Zhang Z. Computational simulations of respiratory-laryngeal interactions and their effects on lung volume termination during phonation: considerations for hyperfunctional voice disorders. *J Acoust Soc Am*. 2021;149:3988–3999.
28. Li S, Scherer RC, Fulcher LP, et al. Effects of vertical glottal duct length on intraglottal pressures and phonation threshold pressure in the uniform glottis. *J Voice*. 2018;32:8–22.
29. Zhang Z. Interaction between epilaryngeal and laryngeal adjustments in regulating vocal fold contact pressure. *JASA Express Letters*. 2021;1: 025201.
30. Zhang Z. Characteristics of phonation onset in a two-layer vocal fold model. *J Acoust Soc Am*. 2009;125:1091–1102.
31. Titze IR. Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *J Acoust Soc Am*. 2002;111:367–376.
32. Chang, S. Computational fluid-structure interaction for vocal fold modeling (Doctoral dissertation). 2016.
33. Li Z, Chen Y, Chang S, Luo H. A reduced-order flow model for fluid−structure interaction simulation of vocal fold vibration. *J Biomech Engi*. 2020;142.
34. Taylor CJ, Thomson SL. Optimization of Synthetic Vocal Fold Models for Glottal Closure. *J Eng Sci Med Diagn Ther*. 2022;5: 031106.
35. Wang X, Jiang W, Zheng X, et al. A computational study of the effects of vocal fold stiffness parameters on voice production. *J Voice*. 2021;35:327–3e1.
36. Li S, Scherer RC, Wan M. Effects of Vertical Glottal Duct Length on Intraglottal Pressures in the Convergent Glottis. *Appl Sci*. 2021;11:4535.
37. Vahabzadeh-Hagh A, Zhang Z, Chhetri D. Quantitative evaluation of the in vivo vocal fold medial surface shape. *J Voice*. 2017;31:513.e15–513.e23.
38. Vahabzadeh-Hagh A, Zhang Z, Chhetri D. Three-dimensional posture changes of the vocal fold from paired intrinsic laryngeal muscles. *Laryngoscope*. 2017;127:656–664.
39. Zhang Z, Chhetri D. Effect of changes in medial surface shape on voice production in excised human larynges. *J Acoust Soc Am*. 2019;146: EL412–EL417.
40. Hirano M. Phonosurgery. Basic and clinical investigations. *Otologia Fukuoka*. 1975;21:239–442.
41. Hirano M. Vocal mechanisms in singing: Laryngological and phoniatric aspects. *J Voice*. 1988;2:51–69.
42. Hirano M, Kakita Y. Cover-body theory of vocal fold vibration. In: Daniloff RG, ed. *Speech Science: Recent Advances*. San Diego: College-Hill Press; 1985:1–46. edited by.
43. Choi H, Berke G, Ye M, et al. Function of the thyroarytenoid muscle in a canine laryngeal model. *Ann Otol Rhinol Laryngol*. 1993;102:769–776.
44. Chhetri D, Neubauer J, Berry D. Neuromuscular control of fundamental frequency and glottal posture at phonation onset. *J Acoust Soc Am*. 2012;131:1401–1412.
45. Luegmair G, Chhetri D, Zhang Z. The role of thyroarytenoid muscles in regulating glottal closure in an in vivo canine larynx model. *Proc Meet Acoust*. 2014;22: 060007.
46. Sonninen A. The external frame function in the control of pitch in the human voice. *Ann New York Acad Sci*. 1968;155:68–90.
47. Vilkman E, Sonninen A, Hurme P, et al. External laryngeal frame function in voice production revisited: a review. *J Voice*. 1996;10:78–92.
48. Vilkman E, Karma P. Vertical hyoid bone displacement and fundamental frequency of phonation. *Acta oto-laryngologica*. 1989;108:142–151.
49. Lee Y, Keating P, Kreiman J. Acoustic voice variation within and between speakers. *J Acoust Soc Am*. 2019;146:1568–1579.
50. Holmberg E, Hillman R, Perkell J. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am*. 1988;84:511–529.
51. Hanson HM, Stevens KN, Kuo HKJ, et al. Towards models of phonation. *J Phon*. 2001;29:451–480.
52. Sundberg J. Articulatory interpretation of the singing formant. *J Acoust Soc Am*. 1974;55:838–844.
53. Yanagisawa E, Estill J, Kmucha ST, et al. The contribution of aryepiglottic constriction to "ringing" voice quality—a videolaryngoscopic study with acoustic analysis. *J Voice*. 1989;3:342–350.
54. Zhang, Z Contribution of undesired medial surface shape to suboptimal voice outcome after medialization laryngoplasty, *J Voice.* 2022. In press. https://doi.org/10.1016/j.jvoice.2022.03.010.
55. Hollien H. Vocal fold thickness and fundamental frequency of phonation. *J Speech Hear Res*. 1962;5:237–243.
56. Hollien H, Colton RH. Four laminagraphic studies of vocal fold thickness. *Folia Phoniatrica et Logopaedica*. 1969;21:179–198.
57. Jun BC, Kim HT, Kim HS, et al. Clinical feasibility of the new technique of functional 3D laryngeal CT. *Acta oto-laryngologica*. 2005;125:774–778.
58. Bakhshaee H, Moro C, Kost K, et al. Three-dimensional reconstruction of human vocal folds and standard laryngeal cartilages using computed tomography scan data. *J Voice*. 2013;27:769–777.
59. Fischer J, Özen AC, Ilbey S, et al. Sub-millisecond 2D MRI of the vocal fold oscillation using single-point imaging with rapid encoding. Magnetic Resonance Materials in Physics. *Biol Med*. 2022;35:301–310.
60. Kobler JB, Chang EW, Zeitels SM, et al. Dynamic imaging of vocal fold oscillation with four-dimensional optical coherence tomography. *The Laryngoscope*. 2010;120:1354–1362.
61. Burns JA. Optical coherence tomography: imaging the larynx. *Curr Opin Otolaryngol Head Neck Surg*. 2012;20:477–481.