

# The Physical Aspects of Vocal Health

Zhaoyan Zhang

For most people, not much conscious thought or effort is needed to produce a voice with the desired pitch, loudness, and voice quality. However, voice disorders are quite common. When disorders occur, the voice may require more effort to produce, be too weak to be heard, or have undesired quality changes that draw unwanted attention. Such changes can affect a speaker's personal identity and the ability to effectively communicate, thus limiting the ability to participate in educational, occupational, or social activities.

Most people have experienced difficulty with their voice after screaming at a sports event or after an upper respiratory infection such as the cold or flu. For teachers, singers, and other professional voice users, voice problems occur more often and the symptoms are often severe. For these people, the voice may get tired toward the end of the day. Sometimes the voice is no longer able to meet the higher expectations and greater demands of one's profession and those individuals have to make career changes.

This article focuses on voice disorders that are related to the production of sound by vocal fold vibration. Voice disorders are often grouped into three major categories based on their etiology. The first category includes organic voice disorders arising from structural changes to the larynx (e.g., inflammation due to an infection or voice overuse) that interfere with the vocal mechanisms.

The second category, neurogenic voice disorders, is related to neurological dysfunctions due to either paralysis, paresis, or neurological disease (e.g., Parkinson's disease) that impact neurological control of the vocal system.

The third category has been characterized in many ways, including as "functional" voice disorders. This category includes voice disorders with no known underlying organic or neurological origins that are presumably related to the improper use of vocal mechanisms and are thus "functional" in some aspect. A widely held

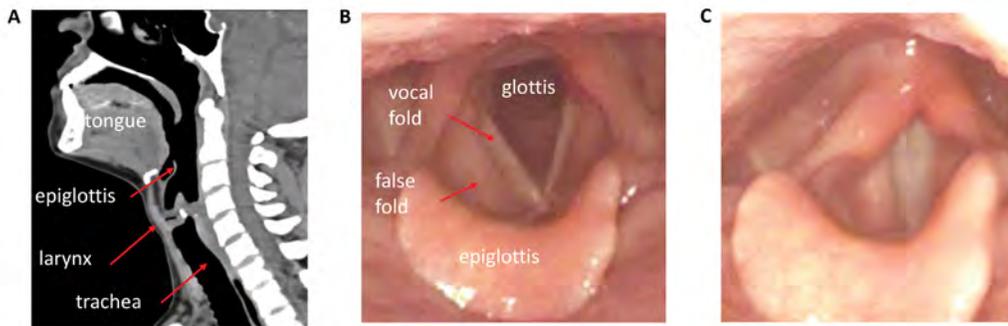
assumption is that these disorders may have psychological origins, but more often they are adaptations to transient tissue changes (e.g., laryngitis) or compromised vocal mechanisms (e.g., paresis or paralysis).

The purpose of this article is not to discuss every voice disorder or category of disorders (but for more information, see Boone et al., 2010; Colton et al., 2011). Instead, it provides an updated review of the physical aspects of vocal health. The focus is on the physical components involved in healthy voice production, the major pathophysiology of voice disorders, and clinical care of common voice problems. The article ends by briefly discussing the existing knowledge gaps between current scientific understanding and the practice of clinical voice care.

## Physiology of Voice Production

The human voice is produced in the larynx (**Figure 1A**), which houses the two opposing vocal folds. Each vocal fold consists of a soft membranous cover layer folded around an inner muscular layer. The vocal folds are connected together anteriorly but slightly separated posteriorly, forming a triangular-shaped airway (the glottis) (**Figure 1B**). At rest, the glottis remains open and allows airflow in and out of the lungs during breathing. During voice production (also known as phonation), the two vocal folds are brought together to close the glottis (**Figure 1C**). When the lung pressure is high enough (about 200 Pa), the vocal folds will be excited into a self-sustained vibration, which periodically opens and closes the glottis. This modulates airflow through the glottis and produces sound, which then propagates through the vocal tract and radiates from the mouth and nasal opening into the voice we hear.

An important feature of normal voice production is that the glottis remains closed for an extended duration within each cycle of vocal fold vibration (see **Multimedia 1** at [acousticstoday.org/zzhangmedia](https://acousticstoday.org/zzhangmedia)), which interrupts the glottal flow. The rapid decline of the glottal flow



**Figure 1.** *A: computed tomography image of the head showing the airway and the larynx. B: top view of the larynx. The vocal folds are far apart at rest. C: vocal folds are brought together to close the glottis during phonation.*

during the glottal closing phase is the main mechanism for harmonic sound production, by which voices of different quality are produced and differentiated. An abrupt cessation of the glottal flow produces a voice with strong harmonic excitation at high frequencies and a bright voice quality that often carries well in a room or open space. On the other hand, a sinusoidal-like shape of the glottal flow with a gradual flow decline, often in the presence of an incomplete glottal closure, produces a voice with a limited number of higher order harmonics in the voice spectrum and a weak voice quality.

The glottal closure pattern during voice production is controlled by adductory laryngeal muscles that bring the two folds together (vocal fold approximation) to reduce the glottal gap. Indeed, phonation is impossible if the glottal gap is too large. Vocal folds that are insufficiently approximated tend to vibrate without complete glottal closure. This produces a breathy voice quality with weak excitation of harmonics and strong noise in the voice spectrum. Increasing approximation of the vocal folds leads to increased vocal fold contact and glottal closure, reducing air leakage through the glottis and increasing harmonic sound generation.

Activation of the adductory laryngeal muscles also modifies vocal fold shape and, particularly, the vertical thickness of the vocal fold medial surface. The medial surface vertical thickness plays an important role in regulating the duration of glottal closure and the produced voice quality. Increasing the vertical thickness allows the vocal folds to better maintain their position against the subglottal pressure. This is

essential to achieve complete glottal closure at high lung pressure while producing a loud voice where vocal fold approximation alone is insufficient to ensure glottal closure during phonation (Zhang, 2016).

In general, thicker vocal folds tend to close the glottis for a longer duration during phonation than thinner vocal folds. Thus, changes in vertical thickness are essential to producing voice qualities ranging from breathy (see **Multimedia 2** at [acousticstoday.org/zzhangmedia](http://acousticstoday.org/zzhangmedia)) to normal (see **Multimedia 3** at [acousticstoday.org/zzhangmedia](http://acousticstoday.org/zzhangmedia)) to pressed (see **Multimedia 4** at [acousticstoday.org/zzhangmedia](http://acousticstoday.org/zzhangmedia)). In the extreme case of very large vocal fold thickness due to strong vocal fold adduction, the folds often exhibit subharmonic or irregular vibration, producing a rough voice quality (Zhang, 2018), known as creak in the linguistic literature and more colloquially as vocal fry (see **Multimedia 5** at [acousticstoday.org/zzhangmedia](http://acousticstoday.org/zzhangmedia)).

Pitch is controlled by elongating and shortening the vocal folds, which regulates the tension and stiffness of the vocal folds. This is possible because the cover layer of each vocal fold consists of collagen and elastin fibers aligned along the anterior-posterior (front-back) direction. These fibers are in a wavy, crimped state at rest but are gradually straightened with elongation and thus become load bearing. As more fibers are gradually straightened with vocal fold elongation, the vocal folds become increasingly stiff, thus increasing pitch.

Because the laryngeal muscles that control the vocal fold length also regulate the vocal fold vertical thickness,

changes in pitch are often accompanied by changes in voice quality. For example, a pitch glide is often accompanied by changes in vocal registers. Vocal fry, produced often with increased vertical thickness and a long period of glottal closure, occurs at the lower end of the pitch range, whereas the voice at the high end of the pitch range is often in a falsetto register, produced with a reduced vertical thickness and a brief duration of glottal closure. The modal voice, which is used in conversational speech, is produced with an intermediate thickness of the vocal fold at the intermediate pitch range.

### Vocal Fold Contact Pressure and Risk of Vocal Fold Injury

During voice production, the vocal folds experience repeated mechanical stress. In particular, the contact pressure sustained by the vocal folds during repeated collision poses the greatest risk of tissue damage because this pressure acts perpendicular to the load-bearing collagen and elastin fibers within the vocal folds (Titze, 1994). For a loud voice such as screaming, the contact pressure can be as high as 20 kPa locally for extreme voicing conditions as reported in recent numerical simulations (Zhang, 2020).

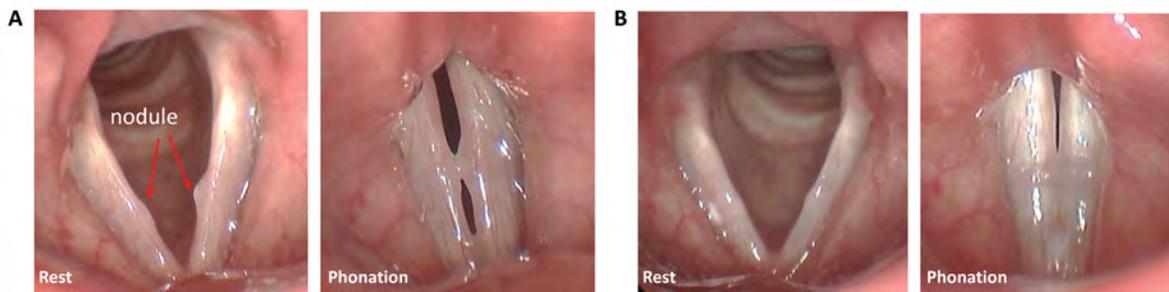
Although the vocal folds evolved to withstand the repeated contact pressure during phonation, when the contact pressure exceeds a certain level (e.g., due to talking loudly or screaming) or is sustained over an extended period (e.g., due to excessive talking or singing), it will cause injury to the vocal folds, triggering an initial inflammation response with fluid accumulation.

This often results in degraded voice quality and difficulty in producing or modulating the voice. The threshold contact pressure triggering the inflammation response appears to vary individually depending on the daily vocal load, overall health condition of the speaker, and, possibly, the microstructural composition of the vocal fold tissues. If this hyperfunction behavior (loud voice for a prolonged period) persists, there may be permanent vocal fold lesions such as vocal fold nodules (Figure 2).

The magnitude of the peak contact pressure depends primarily on the subglottal pressure used to produce the voice and, to a lesser degree, the cover layer stiffness of the vocal folds (Zhang, 2020). Soft vocal folds subject to high subglottal pressure will vibrate with a large vibration amplitude and vocal fold speed at contact, and thus a high contact pressure is required to stop the vocal folds during collision. In general, thinner vocal folds (as, e.g., in a falsetto register) tend to produce lower vocal fold contact pressure (Zhang, 2020). Although the effect of the glottal gap on the contact pressure is generally small, the contact pressure becomes excessively high when the vocal folds are tightly compressed against each other (hyperadduction).

Because the subglottal pressure has a dominant effect on both vocal fold contact pressure and vocal intensity, the risk of vocal fold injury can be significantly reduced by lowering the vocal intensity or completely eliminated by vocal rest. However, vocal rest or reduced loudness is often not socially practical due to communication needs in everyday life. A more practical strategy is to adopt laryngeal and vocal tract adjustments to minimize the subglottal pressure required

**Figure 2.** A: vocal hyperfunction can lead to vocal fold nodules on the medial edge of the vocal folds (left), which prevents complete glottal closure during phonation (right). B: vocal fold nodules almost disappear post-voice therapy (left), which significantly improves glottal closure during phonation (right).



to produce voice of desired loudness, thus minimizing vocal fold contact pressure. At the laryngeal level, this can be achieved by adopting a barely abducted (with the vocal folds just touching each other), thin vocal fold configuration (Berry et al., 2001; Zhang, 2020). This barely abducted configuration is often targeted in voice therapy (e.g., the resonant voice therapy; Verdolini-Marston et al., 1995). In voice training, register balancing between thick and thin vocal folds in singing is often promoted to minimize subglottal pressure and purportedly laryngeal pathologies over time (e.g., the Bel Canto technique).

Vocal fold contact pressure can also be lowered by vocal tract adjustments. For example, when targeting a desired loudness, vocal fold contact pressure can be lowered by constricting the epilarynx (the part of the upper airway immediately above the vocal folds) or increasing the mouth opening whenever possible. Epilaryngeal narrowing often leads to clustering of vocal tract resonances in the 2- to 3-kHz range, which is known as the singer's formant, and amplifies voice harmonics in this frequency range. Increasing the mouth opening increases the efficiency of sound radiation from the mouth. Both adjustments reduce the subglottal pressure required to produce a desired loudness, thus reducing vocal fold contact pressure (Zhang, 2021).

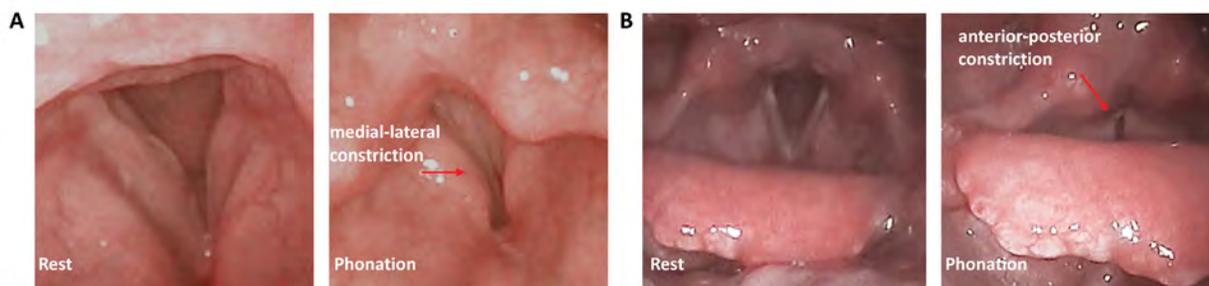
Unfortunately, untrained speakers often increase vocal fold adduction when attempting to increase vocal intensity (Isshiki, 1964), especially in an emotional situation. This is particularly the case of speakers who habitually squeeze the larynx during talking. Hyperadduction of the vocal folds may also develop as an adaptive behavior in response to transient vocal fold tissue changes. Hyperadducted vocal folds are not vocally efficient, meaning that a higher subglottal pressure is required to produce a desired loudness than that needed for barely abducted vocal folds. Because hyperadduction is often accompanied by reduced stiffness and increased thickness in the cover layer, the risk of vocal fold injury is excessively high due to the combination of the high subglottal pressure required, tightly compressed vocal folds, and low cover layer stiffness. Tightly compressed vocal folds also have the tendency to exhibit irregular vocal fold vibration with large cycle-to-cycle variations, resulting in a rough voice quality. Whenever possible, this vocal fold configuration should be avoided in loud voice production by making the appropriate adjustments at the larynx and within the vocal tract.

## Glottal Insufficiency and Adaptive Compensations

Although voice production with tightly compressed vocal folds is unhealthy, voice production with the vocal folds too far apart is also undesired. Whereas the latter vocal configuration requires the least laryngeal effort and poses the lowest risk to vocal fold injury at a low subglottal pressure, voice production is extremely inefficient due to the lack of glottal closure. Thus, attempting to talk loudly in this configuration would require excessively high subglottal pressures, resulting in a high respiratory effort and, potentially, a high vocal fold contact pressure. The produced voice is breathy in nature due to the large airflow escaping through the glottis. With the high lung volume expenditure, one may also feel short of breath and need to take another breath in the middle of an utterance, particularly when attempting to increase loudness. As a result, such a configuration is not ideal for conversational communication or loud voice production.

However, the ability to sufficiently adduct the vocal folds may be lost or weakened due to changes in vocal fold physiology, a condition known medically as glottal insufficiency. Such insufficiency may occur as a result of vocal fold paralysis or paresis due to trauma to the laryngeal nerves, vocal fold atrophy with aging, or changes in the membranous cover layer (e.g., vocal fold swelling or scarring). Under such conditions, one may develop adaptive vocal behaviors in an attempt to increase vocal efficiency and conserve air expenditure. This can be achieved by increasing activation of the adductory muscles to improve glottal closure if the neuromuscular mechanism is still intact. One may also adduct supraglottal structures such as the false folds and epiglottis (**Figure 3**), as often observed in muscle tension dysphonia. Although supraglottal adduction does not improve glottal closure, it may enhance source-tract interaction and thus increase vocal efficiency in addition to air conservation. Such adaptive behaviors often lead to increased laryngeal effort, vocal fatigue over time, and a strained voice quality. If such adaptation persists, it may lead to long-term voice disorders.

For example, vocal fold swelling often occurs after extensive shouting or screaming in a sports event or giving a lecture for a longer than the normal period. Extremely high subglottal pressures and, even more so, vocal fold hyperadduction in these situations readily lead to vocal



**Figure 3.** Adduction of the supraglottal structures may lead to medial-lateral (A: left to right) or anterior-posterior (B: front to back) constriction of the airway immediately above the vocal folds, as often observed in muscle tension dysphonia.

fold swelling. This swelling may also occur following an upper respiratory infection (such as the cold or flu), chemical exposure of the vocal folds due to laryngopharyngeal reflux (stomach acid reflux into the throat), or smoking. Vocal fold swelling makes it difficult to completely close the glottis along the length of the vocal folds, allowing air to escape through gaps around the swollen portion of the vocal folds. When vocal fold inflammation leads to an irregular medial edge of the vocal folds, irregular glottal closure may ensue, resulting in hoarse voice quality.

Vocal fold swelling is often transient and will resolve over time with vocal rest or when the underlying medical conditions have cleared. However, if one were to talk through these voice changes, one often has to increase lung pressure, tighten adduction of the vocal folds, and possibly adduct the false folds and epiglottis. This adaptation may lead to increased contact pressure between the vocal folds, further exacerbating the underlying vocal fold inflammation. If this adaptive behavior persists after the triggering conditions are resolved, the vocal fold inflammation may further develop into vocal fold lesions such as vocal fold nodules, polyps, and contact ulcers, with a more permanent change in voice quality (Hillman et al., 1989). For voice professionals, particularly singers, it is often recommended that they reduce voice use in the presence of vocal fold inflammation and avoid adaptive changes in vocal behavior.

### Muscular Tension Around the Larynx

Voice disorders may also occur from increased tension in the perilaryngeal muscles that support the larynx

(muscles connecting the larynx to other structures around the neck). This is often due to adaptive behaviors to compensate for glottal insufficiency but may also result from psychological stress (Dietrich and Verdolini Abbott, 2012).

Tension in the perilaryngeal muscles often raises the vertical position of the larynx. This results in increased adduction of the vocal folds and the squeezing of supraglottal structures such as the false vocal folds and epiglottis (Figure 3) (Vilkman et al., 1996), allowing a speaker to compensate for glottal insufficiency. However, in the absence of glottal insufficiency, such increased vocal fold adduction often leads to excessively high contact forces between the vocal folds and poses a high risk of vocal fold injury. Due to the high tension in the perilaryngeal muscles, the speaker often experiences vocal fatigue after an extended period of talking and may even feel pain around the neck.

Although voice production is primarily controlled by activities of the intrinsic laryngeal muscles (muscles with origin and insertion within the larynx), these muscles act on the laryngeal framework that is supported and stabilized by the perilaryngeal muscles. Excessive tension in the perilaryngeal muscles acting on the laryngeal cartilages makes it more difficult to adjust the relative position among the thyroid, cricoid, and arytenoid cartilages to which the vocal folds are attached. This may interfere with the delicate control of vocal fold geometry and mechanical properties by the intrinsic muscles and limit the range of vocal fold posturing. Tension in the perilaryngeal muscles may also lead to undesired

relative positions between laryngeal cartilages, which often require compensation by increased activity of the intrinsic laryngeal muscles to maintain pitch or adductory positions. This may change the relative balance between the intrinsic laryngeal muscles, resulting in increased laryngeal effort.

### **Involvement of the Respiratory System**

Adaptive behavior to tighten the larynx may also result from laryngeal-respiratory compensation. The respiratory system is responsible for providing and maintaining the subglottal pressure desired for speech production. In breathing at rest, the respiratory muscles are actively engaged during inspiration, whereas expiration often relies on a passive elastic recoil of the lungs and thorax, known as the relaxation pressure. The amount of relaxation pressure increases with the lung volume and is positive (i.e., pushes air out of the lungs) at a high lung volume and becomes negative (draws air into the lungs) at a very low lung volume. Speech production occurs during the expiration phase of breathing and takes advantage of the relaxation pressure in supplying and maintaining the desired subglottal pressure. By taking a breath to start speech at the appropriate lung volume, the desired subglottal pressure can be mostly supplied and maintained by the relaxation pressure for the entire breath group duration, without much extra respiratory muscle effort. In this sense, speech is often considered “effortless.”

However, when starting speech at either too high or too low lung volumes, extra expiratory muscle effort would be required to either overcome or supplement the relaxation pressure. This additional muscle activation increases rapidly as the lung volume approaches the lower or upper end of the lung capacity. In the extreme case of starting speech at a very low lung volume, in addition to this extra expiratory muscle activation required to maintain the desired subglottal pressure, the level of vocal fold adduction must also be increased to conserve airflow and prevent running out of air before completing an utterance. Thus, speakers who habitually start their speech at a low lung volume often produce a voice with hyperadducted vocal folds and possibly adduction of supraglottal structures (Desjardins et al., 2021), leading to vocal fatigue and undesired voice changes.

A tight laryngeal configuration at a low lung volume may also result from a reduced tracheal pull effect. Tracheal

pull is a downward force exerted by the trachea and the respiratory system on the larynx. This force applies to the cricoid cartilage and tends to reduce the degree of vocal fold adduction. Tracheal pull increases as the diaphragm descends. That is, the tracheal pull is strong when speaking at a high lung volume and decreases as the lung volume decreases (Sundberg, 1993). Thus, when speaking at a very low lung volume, vocal fold adduction may increase naturally due to reduced tracheal pull.

### **Hydration and Environmental Acoustic Support**

Hydration is another important factor in maintaining vocal health. The vocal fold surface is lined by a mucous layer that functions as lubrication to reduce the contact pressure during vocal fold collision. When the speaker is dehydrated, the mucus becomes thick and sticky instead of thin and watery, a condition that deteriorates the lubrication effect in reducing vocal fold contact pressure (Colton et al., 2011). Dehydration may also increase vocal fold stiffness and viscosity, thus increasing the lung pressure required to produce voice. Thus, maintaining good systemic hydration is essential to voice professionals who use their voice extensively in their daily life.

Voice production is mediated through auditory feedback and thus is subject to changes in the speaker’s acoustic environment. For example, with increasing background noise, we often increase vocal intensity to maintain the sufficient speech-to-noise ratio desired for communication. The increase in vocal intensity is often accompanied by a boost of high-frequency harmonic energy with respect to lower-frequency harmonic energy, indicating increased vocal fold adduction.

Similar voice changes are also observed when speaking in rooms with different reverberation characteristics. Speakers produce voice with a higher vocal intensity in rooms with a shorter reverberation time compared with rooms with a longer reverberation time in which acoustic reflections of their own voice provide strong auditory feedback and acoustic support (Brunskog et al., 2009). Thus, speaking for an extended period in a noisy environment or an acoustically “dead” environment with a very short reverberation time is likely to require an increased vocal effort and the speaker is prone to vocal fatigue and risk of vocal fold injury.

## Clinical Voice Care

Clinical voice care attempts to restore the voice through medical, behavioral, and/or surgical interventions. When the voice disorder is triggered by an underlying medical condition, such as vocal fold swelling due to an upper respiratory infection, reflux, or smoking, medical treatment is necessary to clear the medical condition. Due to the delicate structure of the vocal folds, particularly within the membranous cover layer, the initial treatment is often behavioral or voice therapy, particularly for nonorganic voice disorders but also for some organic voice disorders such as vocal fold nodules (**Figure 2**). The goal of voice therapy is to restore the best voice possible, something that is often achieved through vocal health education and modification of vocal behavior using different vocal techniques and exercises. Even for patients who eventually require surgery, pre- and postoperative voice therapy is essential to achieve an optimal voice outcome and prevent recurrence of the voice disorder. For organic voice disorders or conditions of glottic insufficiency, surgical intervention is often more effective.

One of the most common voice disorders in the clinic is muscle tension dysphonia. It involves too extensive an effort in producing the voice, with excessive muscle force and a tight larynx configuration. Some patients may also present with vocal fold lesions such as nodules, due to the chronic exposure to excessively high vocal fold contact pressure. Voice therapy is often effective in improving voice in these patients. For example, external circum-laryngeal massage is often used to relax the larynx in patients with notable tension in the musculature around the neck. Some techniques take advantage of tasks such as yawning or sighing that are naturally produced with a reduced laryngeal muscle tension and a less adducted glottal configuration, often with a lowered vertical position of the larynx. By starting with such tasks and gradually transitioning into speech, the speaker can be trained to produce voice with the same relaxed laryngeal configuration, thus reducing vocal fold contact pressure and the risk of vocal fold injury.

Various vocal exercises are also used to train speakers to produce voice with a focus on vibratory sensations around the lips and cheek and along the alveolar ridge of the palate (e.g., resonant voice therapy), thus avoiding a tight sensation at the larynx. In some exercises, the speaker is instructed to perform pitch or loudness glides

with a semi-occluded vocal tract configuration, producing either nasal sounds, trills, or phonating into a narrow tube such as a drinking straw. It is generally believed that by focusing on vibratory sensations in certain parts of the vocal tract, the speaker may adopt a vocal configuration that improves vocal efficiency and minimizes vocal fold contact pressure.

An important component of voice therapy is to reestablish the balance between respiration, phonation, and articulation. For example, for voice disorders resulting from weakened respiratory function or improper respiratory behavior, voice therapy often focuses on respiration strength training to improve respiratory function or training the speaker to begin speaking at an appropriate lung volume to ensure sufficient air supply required for speech (Desjardins et al., 2021).

For vocal fold mass lesions that are large in size, such as vocal fold polyps, cysts, and sometimes even nodules, voice therapy may have little effect and surgical removal is necessary. Because the membranous cover layer of the vocal folds is the vibrating component, it is critical that surgery remove as little tissue as possible and avoid significantly altering the delicate structure and mechanical properties of the vocal fold cover layer. Vocal fold scarring after surgery, particularly on the vocal fold medial surface where vocal fold vibration modulates airflow most effectively, often negatively impacts the patient's voice and vocal capabilities.

For patients who are unable to sufficiently adduct the vocal folds due to vocal fold paralysis, paresis, atrophy, or aging, vocal fold adduction can be improved through an office-based injection augmentation procedure in which fat or another material is injected into the vocal folds to displace the medial edge of the vocal folds toward the glottal midline. A more permanent solution is medialization laryngoplasty, in which an implant is inserted laterally to the vocal folds to permanently displace and reposition the vocal folds toward the glottal midline (Isshiki, 1989). These procedures are often able to significantly improve glottal closure and voice quality and reduce vocal effort.

In addition to adjusting the vocal fold position, vocal fold surgery also allows manipulation of vocal pitch. One way to achieve this is to adjust vocal fold tension

by surgically modifying the relative positions between laryngeal cartilages. However, this often reduces the vocal range and the amount of pitch change is relatively small. In feminization voice surgery in which a large pitch increase is desired, surgery is often performed to not only adjust vocal fold length but also to reduce the vibrating length of the vocal folds by surgically merging the anterior portions of the two vocal folds or reducing vocal fold mass. Because pitch is only one of many aspects of gender perception, voice therapy is necessary in these patients to adjust other aspects of voice use such as vowel quality, stress, inflection, choice of words and conversational style.

Surgical intervention is also effective in treating some neurological voice disorders. For example, spasmodic dysphonia is a neurological voice disorder that results from involuntary spasms in laryngeal muscle activity, which interferes with normal vocal fold vibration and leads to intermittent voice breaks and strained or breathy voice quality. Current treatment aims to weaken the affected laryngeal muscles through botulinum toxin injection or surgically denervating the affected laryngeal nerves, both of which can significantly alleviate the symptoms.

### **Bridging the Gap Between Science and Clinical Practice**

Current clinical voice care is often quite effective in at least partially improving voice production and quality. However, the voice outcome is often variable and relies heavily on the clinician's experience. Sometimes the voice still remains unsatisfactory after intervention, and the underlying reasons are often unclear. In this sense, clinical voice care is more art than science. The translation of findings from basic science voice research can play an important role in further improving clinical management of voice disorders and reducing variability in voice outcomes. For example, although vocal fold medial surface shape in the vertical dimension has been shown to be important to voice production (Zhang, 2016), it is often not monitored or targeted in current clinical voice examination and intervention, which focus on vocal fold position and glottal closure from a superior, endoscopic view. Targeting the medial surface shape in addition to other intervention goals may improve voice outcomes in patients whose voice remains unsatisfactory after intervention.

Many voice therapy techniques currently used in the clinic were modified from vocal training methods. Although many of them are effective, the underlying scientific principles often remain unclear. For example, semi-occluded vocal tract exercises are widely used in the clinic. Although some theoretical hypotheses have been put forward, they are not always consistent with the observed changes in the laryngeal and vocal tract configuration during such exercises (Vampola et al., 2011). Voice therapy and vocal training often emphasize vibratory sensations in certain parts of the airway. However, it remains unclear what laryngeal and vocal tract adjustments are elicited in patients by voice therapy and which of them are responsible for improvement in voice outcomes. A better understanding of the scientific rationale would allow clinicians to better monitor the progress of voice therapy or even adapt voice therapy toward patient-specific vocal behavior to further improve voice therapy outcomes.

Each individual voice is unique. Although some individuals are prone to vocal fold injury, others can talk loudly for an extended duration without experiencing vocal fatigue or noticeable voice changes. Little is known about the physiological and behavioral factors responsible for individual differences in vocal capabilities and vocal health. A mathematical model of voice production allowing manipulation of the voice in a physiologically realistic way would provide insights into why and how each individual voice is different (Wu and Zhang, 2019), which may lead to interesting applications both inside and outside the clinic.

### **Acknowledgments**

I thank Maude Desjardins, Bruce Gerratt, Katherine Verdolini Abbott, Lisa Bolden, and Arthur Popper for their constructive comments on an earlier draft of this paper. I also acknowledge support from the National Institute on Deafness and Other Communication Disorders (NIDCD), National Institutes of Health (NIH), Bethesda, MD.

### **References**

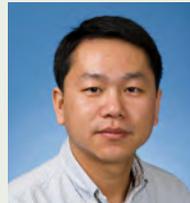
- Berry, D., Verdolini, K., Montequin, D. W., Hess, M. M., Chan, R. W., and Titze, I. R. (2001). A quantitative output-cost ratio in voice production. *Journal of Speech, Language, and Hearing Research* 44, 29-37.
- Boone, D. R., McFarlane, S. C., Von Berg, S. L., and Zraick, R. I. (2010). *The Voice and Voice Therapy*, 8th ed. Allyn & Bacon, Boston, MA.

## PHYSICS OF VOCAL HEALTH

- Brunskog, J., Gade, A. C., Bellester, G. P., and Calbo, L. R. (2009). Increase in voice level and speaker comfort in lecture rooms. *The Journal of the Acoustical Society of America* 125, 2072-2082.
- Colton, R. H., Casper, J. K., and Leonard, R. (2011). *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*. Lippincott Williams & Wilkins, Baltimore, MD.
- Desjardins, M., Verdolini Abbott, K., and Zhang, Z. (2021). Computational simulations of respiratory-laryngeal interactions and their effects on lung volume termination during phonation: Considerations for hyperfunctional voice disorders. *The Journal of the Acoustical Society of America* 149, 3988-3999.
- Dietrich, M., and Verdolini Abbott, K. (2012). Vocal function in introverts and extraverts during a psychological stress reactivity protocol. *Journal of Speech, Language, and Hearing Research* 55, 973-987.
- Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., and Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech, Language, and Hearing Research* 32, 373-392.
- Isshiki, N. (1964). Regulatory mechanism of voice intensity variation. *Journal of Speech, Language, and Hearing Research* 7, 17-29.
- Isshiki, N. (1989). *Phonosurgery: Theory and Practice*. Springer, Tokyo, Japan.
- Sundberg, J. (1993). Breathing behavior during singing. *The NATS Journal* 49, 4-51.
- Titze, I. R. (1994). Mechanical stress in phonation. *Journal of Voice* 8, 99-105.
- Vampola, T., Laukkanen, A., 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. *The Journal of the Acoustical Society of America* 129, 310-315.
- Verdolini-Marston, K., Burke, M. K., Lessac, A., Glaze, L., and Caldwell, E. (1995). Preliminary study on two methods of treatment for laryngeal nodules. *Journal of Voice* 9, 74-85.
- Vilkman, E., Sonninen, A., Hurme, P., and Korkko, P. (1996). External laryngeal frame function in voice production revisited: A review. *Journal of Voice* 10, 78-92.
- Wu, L., and Zhang, Z. (2019). Voice production in a MRI-based subject-specific vocal fold model with parametrically controlled medial surface shape. *The Journal of the Acoustical Society of America* 146, 4190-4198.
- Zhang, Z. (2016). Mechanics of human voice production and control. *The Journal of the Acoustical Society of America* 140, 2614-2635.
- Zhang, Z. (2018). Vocal instabilities in a three-dimensional body-cover phonation model. *The Journal of the Acoustical Society of America* 144, 1216-1230.

- Zhang, Z. (2020). Laryngeal strategies to minimize vocal fold contact pressure and their effect on voice production. *The Journal of the Acoustical Society of America* 148, 1039-1050.
- Zhang, Z. (2021). Interaction between epilaryngeal and laryngeal adjustments in regulating vocal fold contact pressure. *JASA Express Letters* 1, 025201.

### About the Author



#### Zhaoyan Zhang

zyzhang@ucla.edu

Department of Head and Neck Surgery  
University of California, Los Angeles  
Los Angeles, California 90095, USA

**Zhaoyan Zhang** is a professor in the Department of Head and Neck Surgery, University of California, Los Angeles (UCLA). He earned his PhD in mechanical engineering from Purdue University, West Lafayette, Indiana, with his dissertation research on the aeroacoustics of human voice production. Before joining UCLA, he did postdoctoral research on vocal tract acoustics at the University of Maryland, College Park. His current research work focuses on how changes in vocal fold physiology affect voice production and how to infer vocal fold physiology from the produced voice, leading toward clinical and speech technology applications.

### ASA School 2022



## Living in the Acoustic Environment

21-22 May 2022  
Denver, CO area

- **Two-day program:** Lectures, demonstrations, and discussions by distinguished acousticians covering interdisciplinary topics in eight technical areas
- **Participants:** Graduate students and early career acousticians in all areas of acoustics
- **Location:** A conference center/retreat/lodge near Denver
- **Dates:** 21-22 May 2022, immediately preceding the ASA spring meeting in Denver
- **Cost:** \$50 registration fee, which includes hotel, meals, and transportation from the School to the ASA meeting.
- **For information:** Application form, preliminary program, and more details will be available in November, 2021 at [www.AcousticalSociety.org](http://www.AcousticalSociety.org)



## BE SURE TO VISIT AT COLLECTIONS!

.....

[bit.ly/AT-Collections](https://bit.ly/AT-Collections)

See editorial on page 8 to learn how to contribute to Collections.