Integrative Insights into the Myoelastic-Aerodynamic Theory and Acoustics of Phonation. Scientific Tribute to Donald G. Miller

[*]Jan G. Švec, [†]Harm K. Schutte, [‡]C. Julian Chen, and [§]Ingo R. Titze, *Olomouc, Czechia, †Groningen, the Netherlands, ‡New York, New York, and §Salt Lake City, Utah

Summary: In this tribute article to D.G. Miller, we review some historical and recent contributions to understanding the myoelastic-aerodynamic (MEAD) theory of phonation and the related acoustic phenomena in subglottal and vocal tract. At the time of the formulation of MEAD by van den Berg in late 1950s, it was assumed that vocal fold oscillations are self-sustained thanks to increased subglottal pressure pushing the glottis to open and decreased subglottal pressure allowing the glottis to close. In vivo measurements of subglottal pressures during phonation invalidated these assumptions, however, and showed that at low fundamental frequencies subglottal pressure rather tends to reach a maximum value at the beginning of glottal closure and then exhibits damped oscillations. These events can be interpreted as transient acoustic resonance phenomena in the subglottal tract that are triggered by glottal closure. They are analogous to the transient acoustic phenomena seen in the vocal tract. Rather than subglottal pressure oscillations, a more efficient mechanism of transfer of aerodynamic energy to the vocal fold vibrations has been identified in the vertical phase differences (mucosal waves) making the glottal shape more convergent during glottis opening than during glottis closing. Along with other discoveries, these findings form the basis of our current understanding of MEAD.

Key Words: Voice production Vocal fold vibration−Myoelastic-aerodynamic theory of phonation−Subglottal resonances Vocal tract resonances.

INTRODUCTION

What makes the vocal folds vibrate? It is one of the oldest questions of voice science and as such, different explanations have been provided throughout history.[1−3] Nowadays, it has been generally agreed that the basic theory providing the answer is the myoelastic-aerodynamic theory of phonation (MEAD) formulated by van den Berg[1] and further elaborated by Titze and colleagues.[4,5] In this paper, we will highlight some of the basic insights into the MEAD theory that have formed our current understanding of the mechanisms of voice production. We will concentrate on the voice production in humans; the principles of MEAD, nevertheless, have been found valid across most mammals and birds.[6,7]

MEAD theory explains that the vibratory characteristics of the vocal folds depend mainly on the elasticity of the vocal fold tissues (forming the ELASTIC component of MEAD). The elastic properties of the tissues have a major influence on the fundamental frequency of the vocal fold vibrations, very much like the frequency of string oscillations is governed by the elastic properties of the string.[8,9] The elastic properties can be changed by laryngeal muscles (hence the term MYOELASTIC) making the vocal folds stiffer or slacker thus changing the fundamental frequency of the oscillations. The energy for vibration is delivered to the vocal folds by air flowing from the lungs (forming the AERODYNAMIC component of MEAD). Part of the vibrational energy of the vocal folds dissipates within each cycle due to tissue viscosity and therefore energy needs to be added to the vocal folds throughout the vibratory cycles to make the vocal fold oscillations self-sustainable.

By the end of 1950s, at the time of the formulation of MEAD by van den Berg, it was generally assumed that the pressure below the vocal folds, that is, the subglottal pressure, rises when the glottis is closed and decreases when the glottis is open. The mechanism of the vocal fold vibration was explained as a consequence of two assumed effects:

a) the glottis being pushed open by the increased subglottal pressure, and
b) the glottis getting closed when the escaped air through the glottis lowers the subglottal pressure; the closure being accomplished due to the recoil forces of the elastic vocal fold tissues with the help of a sucking force caused by the Bernoulli-effect.[4]

However, both these assumptions turned out to be incorrect and misleading when explaining the mechanism of the self-sustained vocal fold oscillations. The Bernoulli effect, that is, the negative pressure due to glottal airflow,[10] is not sensitive to the direction of vocal fold movement. Therefore, it acts on the vocal folds similarly during glottal opening and closing and does not support the vocal fold self-oscillations, unless changes due to glottis geometry or vocal tract

Accepted for publication January 26, 2021.
From the *Voice Research Laboratory, Department of Experimental Physics, Faculty of Sciences, Palacký University, Olomouc, Czechia; †Groningen Voice Research Lab, Groningen, the Netherlands; ‡Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York; and the §National Center for Voice and Speech, University of Utah, Salt Lake City, Utah.
Address correspondence and reprint requests to Palacký University Olomouc, Faculty of Science, Department of Experimental Physics, 17. listopadu 12, CZ 771 46 Olomouc, Czechia. E-mail: jan.svec@upol.cz
Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997

https://doi.org/10.1016/j.jvoice.2021.01.023

reactance are involved.[4] Furthermore, subglottal pressure was not found to be increased at the instant of glottal opening nor decreased at the instant of glottal closure. Therefore, the understanding of MEAD has changed through the years, as explained below.

MEASUREMENTS AND THEIR INTERPRETATION FOR MEAD

Subglottal pressure variations during phonation

Crucial insights into the real variations of the subglottal pressure during vocal fold oscillations were obtained by measurements that were initiated in the Netherlands by van den Berg and carried out in the 1980s by Schutte and Miller[11–15] and Cranen and Boves.[16,17] They measured the subglottal pressures in vivo using a pressure catheter that was passed via the nose, pharynx and the cartilaginous glottis, into the subglottal space just below the vocal folds. The catheter contained pressure sensors that were placed below the vocal folds registering the subglottal pressure, as well as above the vocal folds registering the supraglottal pressure. The catheter allowed the membranous part of the vocal folds to oscillate freely. Besides the subglottal and supraglottal pressures, the vocal fold vibration was also monitored using an electroglottograph, and the produced sound was registered using a microphone in front of the mouth.

The experimentally observed signals captured by H.K. Schutte on D.G.Miller are shown in Figure 1. The electroglottographic signal (Figure 1c) is used here as the reference indicating the vocal fold contacting and decontacting events from which the glottal closing and opening moments can approximately be derived [18]. The most striking finding is that the subglottal pressure reaches its highest value not at the moment of glottis opening but rather at the initial phase of glottis closure (Figure 1d, bold arrow). Rather than rising, the subglottal pressure shows damped oscillations during the glottis closed phase (interval between the bold and thin arrow in Figure 1d). The opening of the glottis occurs at an instant when subglottal pressure is at a much lower value than it was at the beginning of the closure (Figure 1d, bold versus thin arrow).

These results may seem, at first sight, contradictory to the expectations of MEAD. How can the glottis close when the subglottal pressure is high at that instant? How can the glottis open, when the subglottal pressure is low at that instant? Why is it that the subglottal pressure is lower at the opening moment than at the instant of glottal closure? The original explanation of MEAD assumed the exact opposite.

A more thorough analysis reveals, however, that these findings do not conflict with MEAD but rather provide new insights into how the MEAD and voice acoustic phenomena are related. Now, the glottal closure should rather be seen as

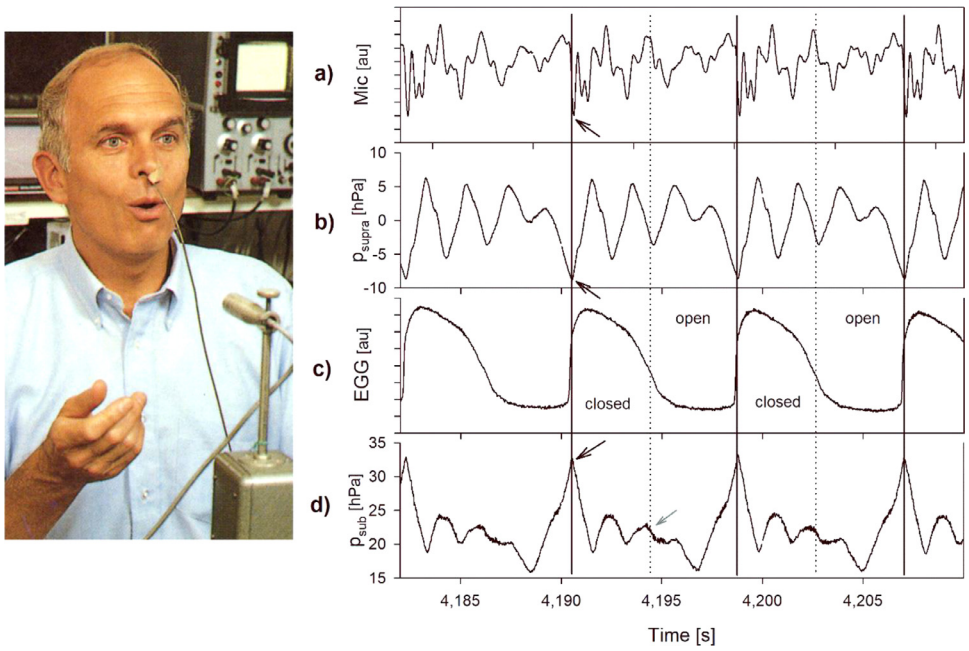

FIGURE 1. Left: Donald G. Miller in the Groningen Voice Research Lab around the year 1984, at the age of 51. He has a catheter, containing pressure sensors, inserted via his nose, pharynx, and larynx into the supraglottal and subglottal space. A microphone detects the radiated sound. For the final experiments,[11,13] also an electroglottograph (not in the image) was used to detect the vocal fold vibrations. Right: The captured waveforms in vowel [ɛ:] around the pitch B2 (120 Hz). From top to bottom, the plots show: a) the microphone signal; b) the supraglottal pressure captured c. 3 cm above the glottis; c) electroglottographic signal; d) the subglottal pressure captured c. 3 cm below the glottis. The vertical lines mark the approximate moments of glottis opening and closure. For synchronization purposes, the time axis of the microphone signal was shifted to the left by 1.4 ms to compensate for the travelling time of the sound from the glottis to the microphone. The bold arrows show the pressure values at the beginning of the closed phase; the thin arrow indicates the subglottal pressure value at the end of the closed phase. The captured waveforms are provided in the supplementary sound files S1–S4. (au = arbitrary units).

an event suddenly interrupting the glottal flow, which has important acoustic consequences by exciting the oscillations of the subglottal as well as supraglottal pressures.[13,19] The acoustic consequences of a sudden glottis closure in the vocal (supraglottal) tract were recently illustrated and educationally explained by Chen and Miller.[20] Here, we expand the concept by including also the subglottal tract.

Acoustic events in the subglottal and supraglottal tracts

Figure 2 considers the case when glottal closure suddenly stops the glottal air flow (between phases 0 and 1). First, we take a look at the events below the vocal folds, in the subglottal tract. When the airflow is suddenly stopped, the air column momentum causes a sudden rise of the subglottal pressure due to sudden accumulation of air below the vocal folds (Figure 2, phase 1). This local increase of air pressure, however, tends to propagate quickly through the whole subglottal space (Figure 2, phase 2−3). This propagation occurs with the speed of sound − the local increase of pressure below the vocal folds acts as a beginning of a sound wave, which starts to propagate away from the vocal folds through the subglottal tract. Since the subglottal tract acts as a damped acoustic resonator, the acoustic wave reflects at the boundaries and travels back and forth through the subglottal tract with a decreasing amplitude (Figure 2, phases 1−13). This is captured by the pressure sensor below the vocal folds as damped oscillations during the closed phase of the glottis (interval between the bold and thin arrow in Figure 1d). The frequency of the damped subglottal pressure oscillations corresponds to the resonance frequency of the subglottal cavities forming the subglottal formant. The lowest subglottal formant frequency has been measured in humans to be in the range of 500−650 Hz and the second subglottal formant frequency around 1200−1500 Hz.[21−23] These subglottal formant frequencies are at the boundaries of those for the vowels [ʊ, ɜː, ʌ] (as used in words "hood, book," "heard, Bert," "hud, but")[24,25] and also within the large formant ranges for the mid central vowel [əː].[26,27] The sound of the subglottal pressure signal can be listened to by playing the supplementary audio file S3.

The sound above the vocal folds is determined by the oscillating supraglottal pressure. This pressure follows a pattern similar to the subglottal pressure, but with opposite polarity (Figure 1b versus 1d). Whereas the glottal closure causes a sudden increase of pressure below the vocal folds (Figure 1d, bold arrow), above the vocal folds it causes a sudden decrease of pressure (Figure 1b, bold arrow, also Figure 2, phase 1). This is because of the inertia of the air column - during the glottal closure the air column above the vocal

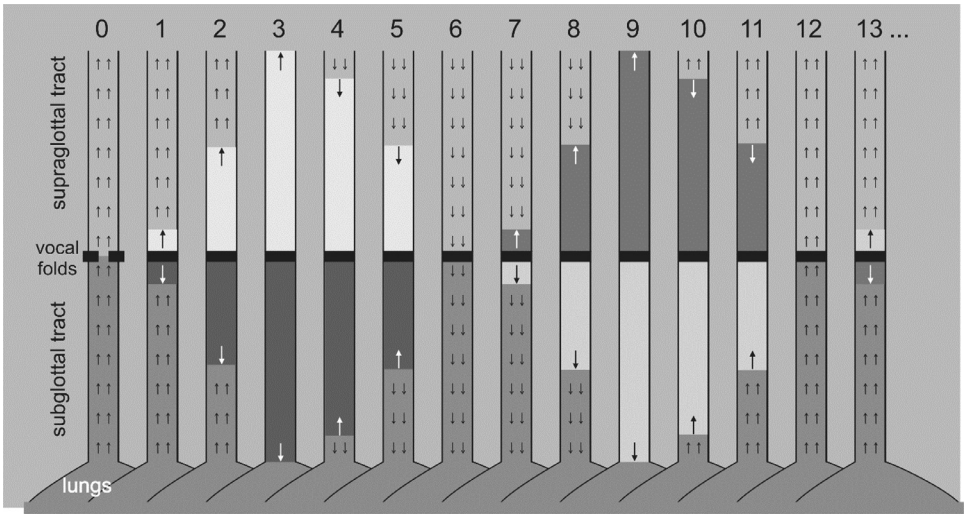

FIGURE 2. Simplified illustration of the acoustic resonance events in the subglottal and supraglottal (vocal) tracts triggered by sudden glottal closure. The glottal closure (change from phase 0 to phase 1) interrupts the airflow and causes a local compression of air below the vocal folds and local rarefaction of air above the vocal folds (phase 1). These local pressure disturbances propagate away from the vocal folds at the speed of sound (phases 1−3). When reaching the ends of the subglottal and vocal tract, the sound waves reflect and propagate back (phases 4−6). Upon reaching the vocal folds, the waves reflect again (phases 7−9). The sound waves travel back and forth (repeating phases 2−13), with a decreasing amplitude due to damping, until the glottis opens. A new glottal closure then initiates a new resonance event. For simplicity, the subglottal and supraglottal tracts are represented here by straight tubes with open ends towards the lungs and the outside space, respectively. The subglottal and supraglottal tracts are set to have the same length, so that their resonance frequencies (formants) are identical. In reality, the subglottal and supraglottal resonance events mostly occur at different frequencies as the vocal tract shape changes due to articulation. Color darkness corresponds to the air compression − the darker the color, the more compressed the air. The single arrows indicate the position and direction of the wavefront of the sound wave. The double arrows indicate the direction of the velocity of the air particles throughout the different phases. The figure is an extended version of the illustration by Chen and Miller.[19]

folds continues to move away from the vocal folds causing a local decrease of air density and thus a local decrease of pressure above the vocal folds.[20,28–30] This sudden local decrease of supraglottal pressure immediately starts to propagate away with the speed of sound in the vocal tract[20,29] (Figure 2, phase 1–3). Since the vocal tract also acts as a damped acoustic resonator, the sound wave travels back and forth in the vocal tract (Figure 2, phases 1–13) with a decreasing amplitude, similar to the subglottal tract. These damped oscillations triggered by the glottal closure were named "timbrons"[29] as they carry information about the voice timbre. The frequency of these damped supraglottal oscillations depends on the vocal tract shape and corresponds to the vowel formants.[20,29,31] Generally, when producing different vowels at an identical fundamental frequency, the supraglottal pressure waveforms and their corresponding sound differ, while the subglottal pressure waveforms and their corresponding sound remain almost identical, as demonstrated by Miller and Schutte[11,13] and by the supplementary sound files S2 and S3, respectively. In the example of Figure 1, the vowel [ɛ:] is used; its first formant has a frequency similar to that of the first subglottal formant.[24] Therefore, the periods of the damped oscillations in the subglottal waveform are similar to those in the supraglottal waveform in this particular case. The damping is slightly greater in the subglottal tract than in the vocal tract, however: notice that the within-cycle oscillations appear stronger in the supraglottal tract (Figure 1b) than in the subglottal one (Figure 1d).

The phenomena recorded in Figure 1 have been verified by repeated measurements as well as by modelling[32] and they invalidate the traditional historical explanation of the mechanism of the vocal fold oscillations, which assumes that the glottis is pushed open by the increased subglottal pressure and closes when the pressure drops. Therefore, the values of subglottal pressure at the instants of glottal closure and opening are apparently not the key elements that drive the vocal fold oscillations.

THE DRIVING MECHANISM OF THE VOCAL FOLD OSCILLATIONS

To explain what drives the vocal folds oscillations, one should take a closer look at how energy is delivered to the vocal folds. The crucial pieces of information can be found as early as in the studies of early computational models of self-sustained vocal fold oscillations by Flanagan and Landgraf,[33] Ishizaka and Flanagan,[34] and later in the theoretical work of Titze.[4,8,28]

Subglottal pressure acts on the glottis from below and is the driving force for the glottal airflow moving upwards. (In phonation based on exhalatory airflow, the mean subglottal pressure is generally higher than the supraglottal pressure causing the air to flow upwards through the open glottis). Subglottal pressure is, however, not the main element driving the vocal folds to open and close the glottis. The main driving pressure for these vocal fold movements is rather the *intraglottal pressure*, that is, the pressure in the glottis just between the vocal folds.

The intraglottal pressure is influenced by the subglottal, supraglottal, as well as by the Bernoulli pressure, depending on the geometry of the glottis.[28,35–42] In the simplest case of the one-mass model of the vocal folds (Figure 3a,b), the average intraglottal pressure p_g in an open glottis can be approximately expressed as

$$p_g = \frac{(p_{sub} + p_{supra})}{2} - p_B \quad (1)$$

where p_{sub}, p_{supra} and p_B are the subglottal, supraglottal and Bernoulli pressures, respectively.

The key for the self-sustained vocal fold vibration is that the intraglottal pressure p_g (not necessarily the subglottal pressure p_{sub}) should be larger when the glottis is opening, and smaller when the glottis is closing.[4,43] This is called the "push-pull effect," which ensures the transfer of the aerodynamic energy to the vocal fold movement.[28] This effect occurs when the following conditions are satisfied[4,8,28,43]:

A) the movement of the vocal tract air column above the glottis is delayed, due to its inertance, relative to vocal fold movement (The Vocal Tract Inertance Mechanism, Figure 3a,b); or
B) the coronal shape of glottis is convergent during opening and divergent (or, at least, less convergent) during closing (Vertical Phase Differences Mechanism, Figure 3c,d).

Below, we will discuss both mechanisms in detail.

A) The Vocal Tract Inertance Mechanism

As noted originally by Rothenberg,[30,44] the inertance of the air column of the vocal tract above the vocal folds is not negligible, which means that during glottal opening this air column can neither start its movement abruptly, nor can it slow its movement down abruptly during glottal closing. In terms of pressures, the inertance of the supraglottal air column causes a positive supraglottal acoustic pressure at glottal opening (Figure 3a) and a negative supraglottal pressure at glottal closing (Figure 3b). Mathematically, this can be expressed by the formula

$$p_{supra} = I\ dU/dt \quad (2)$$

where I is the supraglottal (vocal tract) inertance, and dU/dt is the time derivative expressing the rate of change of the air flow.[43,45] Since the rate of change dU/dt is positive for glottal opening (ie, flow entering the supraglottal tract is increasing) and negative for glottal closing (ie, flow entering the supraglottal tract is decreasing) it makes the supraglottal pressure p_{supra} to be positive for glottal opening and negative for glottal closing. The alternating supraglottal pressure is transferred to the intraglottal pressure (recall Equation 1), causing the push-pull on the vocal folds for energy transfer.

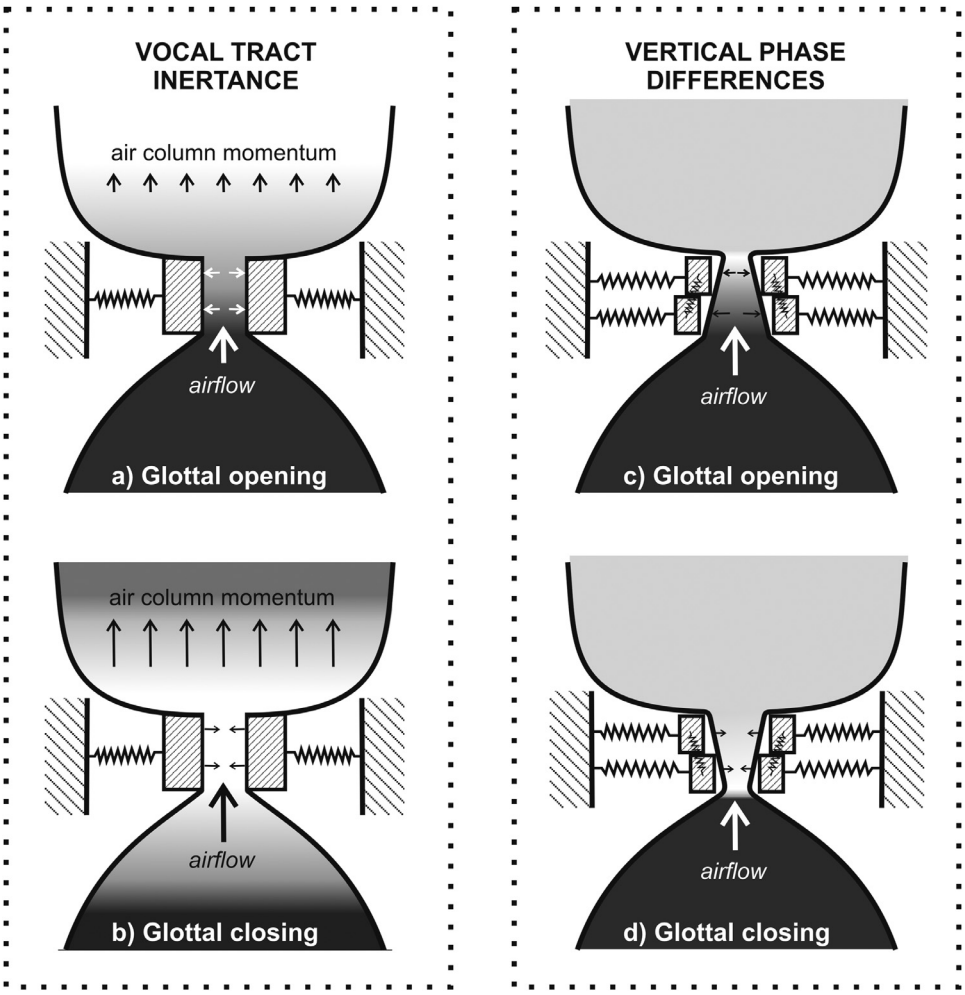

FIGURE 3. Two ways causing the intraglottal pressure to be larger during opening than during closing. A) Vocal tract inertance mechanism (on the left). A one-mass model of the vocal folds is considered here. The inertia of the supraglottal air column makes the airflow to be smaller during opening (a), than during closing (b). B) Vertical phase differences mechanism (on the right). A two-mass model of the vocal folds is considered in this case. The convergent glottis (c) makes the subglottal pressure to be dominant inside the glottis during opening, whereas the divergent glottis (d) makes the supraglottal pressure to be dominant during closing. Notice the changes of the pressure inside the glottis - the larger the pressure, the darker the air column in the figure. In (c,d), the intraglottal pressure reaches the minimum at the narrowest part of glottis where the largest flow velocities make the Bernoulli pressure strongest. Adapted from Titze.[27]

The supraglottal inertance *I* causes the peaks of the pulsating glottal airflow to occur slightly later in time than the maxima of glottal opening.[30,43–46] Consequently, the glottal airflow tends to be smaller during glottal opening than during glottal closing. This causes the Bernoulli pressure inside the glottis to be smaller during opening than during closing, again causing the push-pull on the vocal folds for energy transfer.

The inertial movement of the supraglottal air column in the direction away from the glottis delays the initial decrease of the glottal flow during closing and causes the final glottal flow shutdown to be more abrupt than the glottal area shutdown.[30,43–46] The influence of the subglottal and supraglottal tract on glottal airflow has been termed "Level 1 interaction."[45] The abrupt flow shutdown supports excitation of the acoustic subglottal and supraglottal resonance events (timbrons) illustrated in Figure 2.

The vocal tract inertance mechanism takes place naturally when the fundamental frequency of the vocal fold oscillations is below the first resonance (formant) frequency of the vocal tract, where the reactance is positive (inertive).[45,47] This is usually satisfied at the fundamental frequencies of speaking voice, but it fails for higher fundamental frequencies used, for example, in singing when the vocal tract is compliant.[45,47] That became apparent from the experiments with the first computational model of self-sustained vocal fold oscillations, that is, the one-mass model of Flanagan and Landgraf.[33] When connected to the vocal tract, such a model was able to produce phonations at the fundamental frequencies below the first resonance frequency of the vocal tract. At some regions of fundamental frequencies above the first resonance frequency, the vocal tract was compliant, and the one-mass model could not vibrate there.[34] Furthermore, without the vocal tract, the one-mass model could not vibrate at all.

B) The Vertical Phase Differences Mechanism

In order to overcome the one-mass model limitations, Ishizaka and Flanagan added a second mass to the model dividing the vocal folds into upper and lower parts.[34] These two parts can vibrate out of phase allowing the glottis to take convergent or divergent shape, thus satisfying the B) condition (Figure 3c, d). The two masses can better approximate the changes of the glottis geometry during oscillations that have been observed on excised larynges[48–51] as well as in the *in vivo* investigations.[52,53] Analyses show that, in healthy conditions when the vocal fold mucosa is pliable, the lower vocal fold margin vibrates ahead in phase compared to the upper margin.[49,54–56] This forms the basis for the so-called mucosal wave, travelling from the bottom of the vocal folds across the medial and superior vocal fold surface.[57–62] The vertical phase differences cause the coronal shape of the glottis to be convergent during the opening phase and divergent during the closing phase.[28,50,63,64] The shapes of the glottis at successive phases of the vocal fold vibration cycle and their appearance in laryngeal videoendoscopic and videokymographic images are illustrated in Figure 4.

These changes in the glottal shape greatly influence the intraglottal pressures acting on the vocal fold surface, which depend on the subglottal, supraglottal and Bernoulli pressures.[28,35–42] The convergent shape makes the subglottal pressure dominant during the opening, whereas the divergent shape causes the supraglottal pressure to be dominant during closing. This is illustrated in Figure 3c, d (for mathematical expressions of these phenomena see Titze[28]). Since the mean subglottal pressure is higher than the mean supraglottal pressure, the intraglottal pressures are generally higher in a convergent glottis than in a divergent glottis, thus providing the push-pull effect allowing the vocal folds to self-oscillate.[28,65] In principle, it is not necessary for the glottis to alternate specifically between the convergent and divergent shapes, but it is important for the glottis to take a more convergent shape during opening than during closing. When the vocal fold mucosa is stiff due to some pathology, the vertical phase differences cannot be properly established, and vocal fold oscillations are impeded.[66–68]

The vertical phase differences allow the vocal fold vibrations to be self-sustainable also without the acoustic

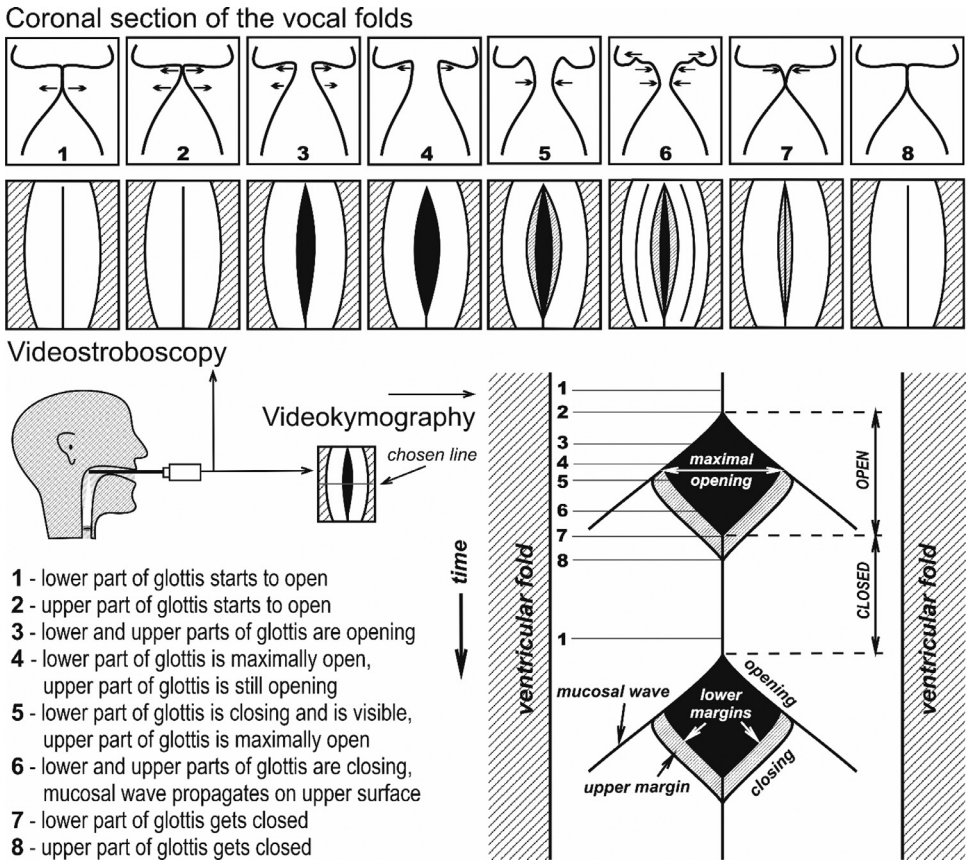

FIGURE 4. Illustration of the kinematics of the vocal fold vibration in ideal circumstances, as seen in coronal view, laryngeal videostroboscopic (or high-speed videoendoscopic) view, and videokymographic view. Mucosal waves travelling upwards on the medial vocal fold surface lead to the phase differences between the motion of the lower and upper margin of the vocal folds. These cause the coronal glottal shape to be convergent during opening and divergent during closing. In the laryngoscopic views, the lower vocal fold margin is hidden during glottis opening and becomes visible during glottal closing. When the mucosal wave reaches the upper margin, it continues to travel laterally across the upper vocal fold surface. These phenomena form the typical vibratory pattern seen in videokymography. (Adapted from Svec et al.[63,64]).

influence of the subglottal and supraglottal tracts. When present, however, the sub- and supraglottal pressure oscillations modify the magnitude of the push-pull effect and interact with the vocal fold vibratory movements. This has an effect on the amplitude and frequency of the vocal fold oscillations.[45] The influence of the subglottal and supraglottal acoustic events on vocal fold movement has been termed "Level 2 interaction."[45] The amount of the interaction between the glottal voice source and vocal tract can be strengthened or weakened particularly by narrowing or enlarging the epilaryngeal cavity of the vocal tract, respectively.[45,69,70] Problematic, compliant vocal tract regions could be minimized with a narrow epilaryngeal airway, which raises the overall vocal tract inertance.[45] The variable resonance frequencies (formants) of the vocal tract can influence the vocal fold vibrations in a variety of constructive and destructive ways. Voice instabilities have been observed when the vocal tract resonance frequencies cross the vocal fold fundamental frequency and vice versa.[71–74] In contrast to the variable vocal tract resonances, the constant subglottal resonances form predictable zones of fundamental frequencies at which the vocal fold oscillations are impeded. These zones have been hypothesized to play a role for the so called "passagio regions" for register transitions detected in the singing voice.[8,75]

CONCLUSION

The basic mechanisms of voice production summarized

Taking into account all the phenomena mentioned above, the mechanism of vocal fold self-oscillations can be, according to the myoelastic-aerodynamic theory, summarized as follows:

During phonation, when exhalatory airflow starts, the glottal constriction causes the air pressure below the vocal folds (ie, subglottal pressure) to be higher than the pressure above the vocal folds (ie, the supraglottal pressure).

The subglottal pressure pushes the lower margins of the two vocal folds apart and the glottis starts opening. The upper margin opens with a slight delay after the lower margin so that the glottis takes a convergent shape. The vertical phase differences form the basis of the mucosal wave travelling upwards across the medial vocal fold surface.

While the intraglottal pressure drives the glottis to open (the "push effect"), at some point the elasticity of the vocal fold tissues will force the vocal folds to reverse the opening movement to closing. The tenser the vocal folds, the sooner the glottal closing starts and the higher the fundamental frequency of the vocal fold oscillations.

Due to the vertical phase differences, the lower margin starts the closing movement earlier than the upper margin and the glottis takes a divergent shape (or less convergent shape than during opening).

The divergent (less convergent) shape causes the intraglottal pressure to drop, pulling the vocal folds together (the "pull effect"). This allows the glottis to close; the closure is not a necessity, however.

Due to tissue elasticity, the horizontal movement of the vocal folds reverses its direction from medial to lateral, the glottis starts opening from below again and the cycle repeats.

Notice that the vocal fold vibratory mechanism described above is, in principle, independent of the acoustic resonance phenomena in the subglottal and supraglottal tract. Indeed, excised larynx experiments have shown that the vocal folds are also capable of vibrating in anechoic laboratory conditions when vocal tract and subglottal resonances are missing.[76] Under normal *in vivo* conditions, however, the push-pull on the vocal fold surfaces due to alternating pressures from an inertive vocal tract can add to the push-pull obtained from the convergent-divergent shape change.

The glottal opening and closing modulates the airflow passing through the glottis. The modulated airflow acts as the sound source, which is well known from the source-filter acoustic theory of voice production formulated by Fant.[19,31] Nonlinear effects due to the presence of subglottal and supraglottal tracts (such as the effect of the non-negligible vocal tract inertance), cause the glottal airflow waveform to be skewed compared to the glottal area waveform, however.[30,43–46] Furthermore, the presence of the subglottal and supraglottal tracts triggers the acoustic resonance phenomena, as indicated in Figure 2. The acoustic resonance phenomena add formants to the resulting subglottal and supraglottal sound and cause the subglottal and supraglottal pressure waveforms to be rather complex, as shown in Figure 1. Such complex pressure oscillations just below and above the glottis then in turn interact with the voice source, influence the vocal fold vibrations and can trigger nonlinear-dynamic voice phenomena, such as pitch shifts, register breaks, etc.[71–74,77]

At the fundamental frequencies that are well below the subglottal and supraglottal formant frequencies, the subglottal and supraglottal acoustic resonances are strongly excited particularly when there is full glottal closure – such a case was considered in Figures 1 and 2. However, for phonation to occur the glottis does not need to close fully - the mechanisms mentioned above work similarly also in phonations with incomplete glottal closure, as can often be observed in persons with no voice complaints.[78]

DEDICATION

This article is devoted to the memory of Donald Gray Miller, Ph.D. (February 21, 1933 – April 22, 2020), singer, voice pedagogue, one of very few actively performing voice professionals who spent part of his life as a researcher following his drive to understand more scientifically the singing voice and inspiring many followers. He was our dear colleague in the Groningen Voice Research Lab for more than 30 years (Head: prof.dr. H.K. Schutte).[79] His many

contributions to singing voice pedagogy include, for example, studying and describing the resonance strategies in different singing voice registers,[80,81] development of the "VoceVista" software tool for visualization of the sound of voice,[82] development of a portable electroglottograph "EGGs for singers" to be used in singing studios,[83] and much more. These contributions have been widely recognized in the community of singing voice pedagogues and scientists. His contributions to the fundamentals of voice science, however, have remained to be known only to a narrower circle of voice scientists. This paper acknowledges some of these within a wider context of the development of basic voice science and reduces the highly complex process of voice production to its essential elements so that the information is accessible also to nonscientists.

CONFLICTS OF INTEREST

The authors declare that they have no financial conflicts of interest to disclose.

Acknowledgments

The work of JGS has been supported by the Czech Science Foundation (GA CR) project no. 19-04477S. An initial version of this paper was presented at the KTH Stockholm Popup Voice Science Coffee Break No.2 online seminar on April 23, 2020 (https://www.kth.se/tmh/research/voice-science-popups-1.976261). The authors appreciate the help of Pedro Amarante Andrade with digitizing the original multichannel tape recordings of Schutte and Miller.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2021.01.023.

REFERENCES

1. Berg van den J. Myoelastic-aerodynamic theory of voice production. *J Speech Hear Res*. 1958;3:227–244.
2. Cooper DS. Research in laryngeal physiology with excised larynges. In: Cummings CW, Fredrickson JM, Harker LA, Schuller DE, Krause CJ, eds. *Otolaryngology - Head and Neck Surgery*. Vol 3. 2 ed. St. Louis and Toronto: C. V. Mosby; 1986:1728-1737.
3. Schutte HK. Historical approaches in revealing the singing voice, Part 1. In: Welch G, Howard DM, Nix J, eds. *Oxford Handbook of Singing*. Oxford, UK: Oxford University Press; 2019:965–990.
4. Titze IR. Comments on the myoelastic - aerodynamic theory of phonation. *J Speech Hear Res*. 1980;23:495–510.
5. Titze IR. *The Myoelastic Aerodynamic Theory of Phonation*. Denver COIowa City IA: National Center for Voice and Speech; 2006.
6. Elemans CPH, Rasmussen JH, Herbst CT, et al. Universal mechanisms of sound production and control in birds and mammals. *Nat Commun*. 2015;6:1–13.
7. Herbst CT. Biophysics of vocal production in mammals. In: Suthers RA, Fitch WT, Fay RR, Popper AN, eds. *Vertebrate Sound Production and Acoustic Communication*. 1 ed Springer International Publishing Switzerland; 2016:159–189.
8. Titze IR. *Principles of Voice Production (second printing)*. Iowa City, IA: National Center for Voice and Speech; 2000.
9. Verdolini K, Titze IR. The application of laboratory formulas to clinical voice management. *Natl Center Voice Speech Status Progr Rep*. 1994;7:197–205.
10. Berg van den J, Zantema JT, Doornebal Jr P. On the air resistance and the Bernoulli effect of the human larynx. *J Acoust Soc Am*. 1957;29:626–631.
11. Miller DG, Schutte HK. Characteristic patterns of sub- and supraglottal pressure variations within the glottal cycle. In: Van Lawrence L, ed. *Transcripts of the XIIIth Symposium: Care of the Professional Voice, New York, 1984*. New York, NY: The Voice Foundation; 1985:70–75.
12. Schutte HK, Miller DG. The effect of F0/F1 coincidence in soprano high notes on pressure at the glottis. *J Phonetics*. 1986;14:385–392.
13. Schutte HK, Miller DG. Resonanzspiele der Gesangsstimme in ihren Beziehungen zu supra- und subglottalen Druckverläufen: Konsequenzen für die Stimmbildungstheorie. [Play of resonances in the singing voice in the supra- and subglottal pressure changes: consequences for the theory of voice production]. *Folia Phoniatr (Basel)*. 1988;40:65–73.
14. Miller DG, Schutte HK. Formant tuning in a proffesional baritone. *J Voice*. 1990;4:231–237.
15. Miller DG, Schutte HK. Effects of downstream occlusions on pressures near the glottis in singing. In: Gauffin J, Hammarberg B, eds. *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*. San Diego, California: Singular Publishing Group; 1991:91–98.
16. Cranen B, Boves L. A set-up for testing the validity of the two mass model of the vocal folds. In: Titze IR, Scherer RC, eds. *Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control*. Denver, CO: The Denver Center for the Performing Arts; 1983:500–513.
17. Cranen B, Boves L. Pressure measurements during speech production using semicondductor miniature pressure transducers - impact on models for speech production. *J Acoust Soc Am*. 1985;77:1543–1551.
18. Herbst CT, Lohscheller J, Svec JG, et al. Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings. *J Exp Biol*. 2014;217:955–963.
19. Fant G. Glottal source and excitation analysis. *STL-QPSR*. 1979:85–107. 1/1979.
20. Chen CJ, Miller DG. Pitch-synchronous analysis of human voice. *J Voice*. 2019;34:494–502.
21. Cranen B, Boves L. On subglottal formant analysis. *J Acoust Soc Am*. 1987;81:734–746.
22. Arsikere H, Leung GK, Lulich SM, et al. Automatic estimation of the first three subglottal resonances from adults speech signals with application to speaker height estimation. *Speech Commun*. 2013;55:51–70.
23. Sundberg J, Scherer R, Hess M, et al. Subglottal pressure oscillations accompanying phonation. *J Voice*. 2013;27:411–421.
24. Peterson GE, Barney HL. Control methods used in study of the vowels. *J Acoust Soc Am*. 1952;24:175–184.
25. Childers DG, Wu K. Gender recognition from speech. Part II: fine analysis. *J Acoust Soc Am*. 1991;90(4 Pt 1):1841–1856.
26. Koopmans-van Beinum FJ. What's in a schwa? *Phonetica*. 1994;51:68–79.
27. Flemming E. The phonetics of schwa vowels. *MIT*. Available at: http://web.mit.edu/flemming/www/paper/schwaphonetics.pdf. Accessed December 14, 2020.
28. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am*. 1988;83:1536–1552.
29. Chen CJ. *Elements of Human Voice*. New Jersey: World Scientific; 2016.
30. Rothenberg M. Acoustic interaction between the glottal source and the vocal tract. In: Stevens KN, Hirano M, eds. *Vocal Fold Physiology*. Tokyo: University of Tokyo Press; 1981:305–328.
31. Fant GM. *Acoustic Theory of Speech Production*. The Hague, NL: Mouton; 1960.
32. Ho JC, Zanartu M, Wodicka GR. An anatomically based, time-domain acoustic model of the subglottal system for speech production. *J Acoust Soc Am*. 2011;129:1531–1547.
33. Flanagan JL, Landgraf LL. Self oscillating source for vocal tract synthesizers. *IEEE Trans Audio Electroacoust*. 1968;AU-16:57–64.
34. Ishizaka K, Flanagan JL. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst Tech J*. 1972;51:1233–1268.

35. Ishizaka K. Air resistance and intraglottal pressure in a model of the larynx. In: Titze IR, Scherer RC, eds. *Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control*. Denver, CO: The Denver Center for the Performing Arts; 1983:414–424.
36. Li S, Scherer RC, Wan MX, et al. The effect of glottal angle on intraglottal pressure. *J Acoust Soc Am*. 2006;119:539–548.
37. Decker GZ, Thomson SL. Computational simulations of vocal fold vibration: Bernoulli versus Navier-Stokes. *J Voice*. 2007;21:273–284.
38. Khosla S, Oren L, Ying J, et al. Direct simultaneous measurement of intraglottal geometry and velocity fields in excised larynges. *Laryngoscope*. 2014;124:S1–S13.
39. Oren L, Khosla S, Gutmark E. Intraglottal pressure distribution computed from empirical velocity data in canine larynx. *J Biomech*. 2014;47:1287–1293.
40. Alipour F, Scherer RC. Time-dependent pressure and flow behavior of a self-oscillating laryngeal model with ventricular folds. *J Voice*. 2015;29:649–659.
41. Sidlof P, Zörner S, Hüppe A. A hybrid approach to the computational aeroacoustics of human voice production. *Biomech Model Mechanobiol*. 2015;14:473–488.
42. DeJonckere PH, Lebacq J. In vivo quantification of the intraglottal pressure: modal phonation and voice onset. *J Voice*. 2020;34:645.e619–645.e639.
43. Story BH. Mechanisms of voice production. In: Redford MA, ed. *The Handbook of Speech Production*. edition: 1 Wiley-Blackwell; 2015:34–58.
44. Rothenberg M, Zahorian S. Nonlinear inverse filtering technique for estimating glottal-area waveform. *J Acoust Soc Am*. 1977;61:1063–1071.
45. Titze IR. Nonlinear source-filter coupling in phonation: theory. *J Acoust Soc Am*. 2008;123:2733–2749.
46. Granqvist S, Hertegard S, Larsson H, et al. Simultaneous analysis of vocal fold vibration and transglottal airflow: exploring a new experimental setup. *J Voice*. 2003;17:319–330.
47. Titze IR. Source-filter interaction in speaking and singing is nonlinear. *Echoes Newslett Acoust Soc Am*. 2007;17:1–3.
48. Hiroto I. *Vibration of Vocal Cords: an Ultra High-Speed Cinematographic Study. (Film)*. Kurume, Japan: Department of Otolaryngology, Kurume University; 1968.
49. Matsushita H. The vibratory mode of the vocal folds in the excised larynx. *Folia Phoniatr (Basel)*. 1975;27:7–18.
50. Berry DA, Montequin DW, Tayama N. High-speed digital imaging of the medial surface of the vocal folds. *J Acoust Soc Am*. 2001;110(5, Pt.1):2539–2547.
51. Doellinger M, Berry DA. Visualization and quantification of the medial surface dynamics of an excised human vocal fold during phonation. *J Voice*. 2006;20:401–413.
52. Yumoto E, Kadota Y, Mori T. Vocal fold vibration viewed from the tracheal side in living human beings. *Otolaryngol Head Neck Surg*. 1996;115:329–334.
53. Berke GS, Moore DM, Hantke DR, et al. Laryngeal modeling: theoretical, in vitro, in vivo. *Laryngoscope*. 1987;97:871–881.
54. Hiroto I. The mechanism of phonation: its pathophysiological aspect. (In Japanese). *Nippon Jibiinkoka Gakkai Kaiho*. 1966;69:2097–2106.
55. Baer T. *Investigation of Phonation Using Excised Larynxes. (Doctoral dissertation)*. Cambridge, Mass: Massachusetts Institute of Technology; 1975.
56. Titze IR, Jiang JJ, Hsiao TY. Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration. *Ann Otol Rhinol Laryngol*. 1993;102:58–63.
57. Musehold A. Stroboskopische und fotografische Studien über die Stellung der Stimmlippen im Brust- und Falsett-Register. [Stroboscopic and photographic studies about the vocal fold adjustments in chest and falsetto registers]. *Arch Lar Rhinol*. 1898;7:1–21.
58. Farnsworth DW. High-speed motion pictures of the human vocal cords. *Bell Lab Record*. 1940;18:203–208.
59. Saito S, Fukuda H, Isogai Y, et al. X-ray stroboscopy. In: Stevens KN, Hirano M, eds. *Vocal Fold Physiology*. Tokyo: University of Tokyo Press; 1981:95–106.
60. Boessenecker A, Berry DA, Lohscheller J, et al. Mucosal wave properties of a human vocal fold. *Acta Acust United Acust*. 2007;93:815–823.
61. Kobler JB, Chang EW, Zeitels SM, et al. Dynamic imaging of vocal fold oscillation with four-dimensional optical coherence tomography. *Laryngoscope*. 2010;120:1354–1362.
62. Kumar SP, Svec JG. Kinematic model for simulating mucosal wave phenomena on vocal folds. *Biomed Signal Process Control*. 2019;49 (March):328–337.
63. Svec JG, Sram F, Schutte HK. Videokymografie: nová vysokofrekvenční metoda vyšetřování kmitů hlasivek. [Videokymography: a new high-speed method for the examination of vocal-fold vibrations]. *Otorinolaryngol (Prague)*. 1999;48:155–162.
64. Svec JG. *On vibration Properties of Human Vocal Folds: Voice Registers, Bifurcations, Resonance Characteristics, Development and Application of Videokymography*. Groningen, the Netherlands: University of Groningen; 2000. Ph.D. thesis.
65. Thomson SL, Mongeau L, Frankel SH. Aerodynamic transfer of energy to the vocal folds. *J Acoust Soc Am*. 2005;118(3, pt.1):1689–1700.
66. Hirano M. *Clinical Examination of Voice*. 5. Wien, Austria: Springer-Verlag; 1981.
67. Bless DM, Hirano M, Feder RJ. Videostroboscopic evaluation of the larynx. *Ear Nose Throat J*. 1987;66(July 1987):289–296.
68. Hirano M, Bless DM. *Videostroboscopic Examination of the Larynx*. San Diego, California: Singular Publishing Group; 1993.
69. Titze IR, Palaparthi A. Sensitivity of source-filter interaction to specific vocal tract shapes. *IEEE/ACM Trans Audio Speech Lang Process*. 2016;24:2507–2515.
70. Lucero JC, Lourenco KG, Hermant N, et al. Effect of source-tract acoustical coupling on the oscillation onset of the vocal folds. *J Acoust Soc Am*. 2012;132:403–411.
71. Titze I, Riede T, Popolo P. Nonlinear source-filter coupling in phonation: vocal exercises. *J Acoust Soc Am*. 2008;123:1902–1915.
72. Uezu Y, Kaburagi T. A measurement study on voice instabilities during modal-falsetto register transition. *Acoust Sci Technol*. 2016;37:267–276.
73. Wade L, Hanna N, Smith J, et al. The role of vocal tract and subglottal resonances in producing vocal instabilities. *J Acoust Soc Am*. 2017; 141:1546–1559.
74. Murtola T, Aalto A, Malinen J, et al. Modal locking between vocal fold oscillations and vocal tract acoustics. *Acta Acustica United Acustica*. 2018;104:323–327.
75. Titze IR. A framework for the study of vocal registers. *J Voice*. 1988; 2:183–194.
76. Lehoux H, Hampala V, Svec JG. Subglottal pressure oscillations in anechoic and resonant conditions and their influence on excised larynx phonations. *Sci Rep*. 2021;11(Art. no. 28):1–14.
77. Titze IR, Baken RJ, Herzel H. Evidence of chaos in vocal fold vibration. In: Titze IR, ed. *Vocal Fold Physiology: Frontiers in Basic Science*. San Diego, CA: Singular Publishing Group; 1993: 143–188.
78. Sulter AM, Schutte HK, Miller DG. Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. *J Voice*. 1996;10:175–189.
79. Schutte HK. Historical approaches in revealing the singing voice, Part 2. In: Welch G, Howard DM, Nix J, eds. *Oxford Handbook of Singing*. Oxford, UK: Oxford University Press; 2019:991–1026.
80. Miller DG. *Registers in Singing: Empirical and Systematic Studies in the Theory of the Singing Voice*. Groningen, the Netherlands: University of Groningen; 2000. Ph.D. thesis.
81. Miller DG. *Resonance in Singing: Voice Building Through Acoustic Feedback*. Princeton, NJ: Inside View Press; 2008.
82. Miller DG, Schutte HK. The use of spectrum analysis in the voice studio. In: Nair G, ed. *Voice – Tradition and Technology. A State-of-the-Art Studio*. San Diego, CA: Singular Publishing Group; 1999:189–210.
83. Miller DG, Schutte HK. The use of the electroglottograph in the voice studio. In: Nair G, ed. *Voice – Tradition and Technology. A State-of-the-Art Studio*. San Diego, CA: Singular Publishing Group; 1999:211–225.