Published in final edited form as:

J Voice. 2017 March ; 31(2): 135–141. doi:10.1016/j.jvoice.2016.06.003.

HUMAN SPEECH: A RESTRICTED USE OF THE MAMMALIAN LARYNX

Ingo R. Titze^{1,2}

¹National Center for Voice and Speech, The University of Utah, lead institution, Salt Lake City, UT

²Department of Communication Sciences and Disorders, The University of Iowa, Iowa City, IA

Abstract

Purpose—Speech has been hailed as unique to human evolution. While the inventory of distinct sounds producible with vocal tract articulators is a great advantage in human oral communication, it is argued here that the larynx as a sound source in speech is limited in its range and capability because a low fundamental frequency is ideal for phonemic intelligibility and source-filter independence.

Method—Four existing data sets were combined to make an argument regarding exclusive use of the larynx for speech: (1) range of fundamental frequency, (2) laryngeal muscle activation, (3) vocal fold length in relation to sarcomere length of the major laryngeal muscles, and (4) vocal fold morphological development.

Results—Limited data support the notion that speech tends to produce a contracture of the larynx. The morphological design of the human vocal folds, like that of primates and other mammals, is optimized for vocal communication over distances for which higher fundamental frequency, higher intensity, and fewer unvoiced segments are utilized than in conversational speech.

Conclusion—The positive message is that raising one's voice to call, shout, or sing, or executing pitch glides to stretch the vocal folds, can counteract this trend toward a contracted state.

Keywords

speech; larynx; contracture; singing; muscles

INTRODUCTION

Mammals, birds, amphibians, and reptiles communicate with sound produced in respiratory airways, using either a larynx or a syrinx (birds) as a sound source. Sonic, infrasonic, or ultrasonic vibration is sustainable by nonlinear interaction between the airstream and a

Corresponding Author: Ingo R. Titze, National Center for Voice and Speech, 136 South Main Street, Suite 320, Salt Lake City UT 84101-3306, ingo.titze@ncvs2.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

collapsible (soft tissue) segment of the airway wall, producing a fundamental frequency f_0 and a spectrum of higher frequencies. The vibrating tissue disturbs the airstream so that acoustic waves are propagated in the airways. These waves travel along the slowly-moving airstream, with a small portion of the sound being radiated from the mouth, beak, or nostrils into free space to the listener. A much larger portion of the sound is retained in the airway in the form of multiple reflections from irregular boundaries, producing standing waves that form distinct classes of sound. Amplitude and frequency modulations of the fundamental frequency and higher partials are added to allow rhythmic and melodic patterns to be produced.

Communication strategies vary across species (Taylor & Reby, 2010). A single tone with constant fundamental frequency and amplitude, considered the *carrier* of vocal communication, reveals the presence and location of the animal, and possibly its size and some degree of identity. Modulations of this carrier are needed to create a sufficient inventory of sounds for all of the animal's communication needs. In analogy with radio communication theory, the carrier is of a much higher frequency than the modulations, which are in the form of variations in amplitude, frequency, duration, or frequency spectrum. Ideally, the modulations are not so large that the carrier (voicing) is interrupted. For short-distance communication, however, unvoiced segments have become part of the sound inventory. Prosodic variations, such as melody and accent, do not need to interrupt the carrier and are therefore good candidates for long-range vocal communication. They are used by most species that vocalize. Range of frequency and amplitude, speed of change of frequency and amplitude, and frequency spectrum of the overtones of the carrier frequency, become important criteria for a large inventory of vocal signals that a species can produce.

With the evolution of speech on the order of 100,000 years ago (Lieberman, 1991), humans discovered that a larger inventory of modulations could be produced by changing the airway structures rather than simply the sound-source characteristics. Moving the tongue, the lips, the jaw, or the velum, allowed the frequency spectrum of the carrier overtones to be modulated amply to produce vowels and consonants. Articulation became the dominant modulation in vocal communication, so much so that whispered speech, or speech with a buzzer held against the neck, is viable today for close-range communication. Source modulations are of secondary importance. The invention of electronic amplification makes source modulations with wide frequency and amplitude ranges even less essential for speech.

The purpose of this paper is to put forth an argument, with existing data never presented in combination, that adaptation of the mammalian larynx for long-range unamplified vocal communication could eventually be reversed with excessive or exclusive use of speech over short distances. Calling, with its many variations (howling, shouting, hooting, roaring, screaming, chanting) is practiced so little that a predisposition to motor control problems in the larynx may exist. Furthermore, infrequent mechanical stretching of laryngeal tissues may diminish the need for a multi-layered vocal fold morphology. Fragmentary evidence is given here in the nature of acoustic requirements, vocal fold morphology, vocal fold posturing for speech, and approaches used for voice therapy and training. The methods are in the form of corroborating data sets produced over several decades.

METHODS

Acoustic Requirements for Speech

The ideal sound carrier for speech articulation has a low fundamental frequency (f_0). A logical argument is that intelligibility of phonemes is improved if source harmonics are closely spaced, given that vocal tract resonances can then be sampled with a greater number of frequencies. With identical formant structure, male speech should by this argument be more intelligible than female speech. This has not been shown, however. In fact, the evidence is somewhat in the opposite direction (Byrd, 1994; Bradlow et al., 1997). The answer lies in the fact that formant frequencies are higher in females than males, which tends to equalize the sampling issue. At female fundamental frequencies beyond those typically used in speech (i.e., singing), it has been shown that intelligibility is indeed reduced (Joliveau, Smith, & Wolfe, 2004; Collister & Huron, 2008; Sundberg & Ternström, 2008). There is poorer sampling of the resonances of the vocal tract. For example, since most vowels are identified by the first two resonant frequencies of the supraglottal airway (the exception being rhotic vowels), and since these frequencies center around 500 Hz and 1500 Hz, respectively, a fundamental frequency f_0 that is not well below 500 Hz provides too few harmonics to sample the resonance peaks effectively. The same can be said for voiced consonants, which also have low resonance frequencies. Thus, there is a strong tendency for humans to keep f_0 low during speech. Unvoiced consonants use secondary sound sources with greater spectral density (hisses, clicks, and pops), but these sounds are not carried over distances more than a few meters. They are generally not effective for long-range vocal communication unless amplification is used.

Figure 1 shows the range of sound pressure level (SPL) plotted against range of f_0 , known as a voice range profile (VRP). The curves labeled loud and soft show the non-speech range of SPL obtainable 30 cm from the mouth over the entire f_0 range of ten males (Gramming, 1988; Gramming & Sundberg, 1988). Superimposed are four speech contours from male classroom teachers who speak daily on the order of 6 hours (Hunter & Titze, 2010). Note that 70 % of the speech f_0 falls below 200 Hz, while fundamental frequencies beyond 500 Hz are available from the male larynx with equal or greater SPL variation.

Aside from the requirement of low f_0 for vowel and voiced consonant intelligibility, strong nonlinear coupling between the source and the filter is avoided when f_0 is low. This creates a stability region for source harmonics in speech. It has been shown that low harmonics of the source spectrum (in particular f_0 or $2 f_0$) passing through airway resonances can destabilize vocal fold vibration (Titze & Riede, 2008). Avoiding these crossings keeps the source spectrum more constant (less registered in voice science terminology, or less bifurcated in nonlinear dynamics terminology) so that modulation with articulation is not confused with spectral uncertainty at the source. Hence, a linear source-filter theory of speech production that separates the sound source from the resonator has been the hallmark of speech science and technology for over half a century, beginning with Fant (1960).

A current trend among young people (notably females) is to speak in vocal fry, also known as pulse register (Wolk, & Abdelli-Beruh, 2012; Abdelli-Beruh, Wolk, & Slavin, 2014). A subharmonic series (period doubling or tripling) is often produced, which increases the

density of source frequencies in the spectrum. This would appear to increase speech intelligibility. Vowels and consonants are extremely well sampled, but overall vocal intensity is low and the voice quality is rough. With amplification or small speaker-listener distances, low intensity does not seem to be an issue. However, prosodic variations (i.e., intonation that takes the voice into and out of fry without linguistic intent) appear to be more difficult to produce. Without ample prosodic variation to communicate mood, personality, or other paralinguistic characteristics, speech can degenerate to *vocal texting*.

Prior to acquisition of the major articulatory components of speech, human infants vocalize at high f_o (around 400 – 500 Hz). Mothers encourage high f_o vocalization with “mothereses” that modulate the high f_o carrier with lots of prosodic variation (Kuhl, 2000). The fundamental frequency steadily drops in the first 3 years (to around 300 Hz) as more and more articulation develops (Kent, 1976). Calling, shouting, laughing, and crying continue towards puberty, but are often suppressed thereafter for social reasons. Even singing, which in some cultures is a daily family or school activity, is becoming less habilitated with electronic amplification that can extend the acoustic dimensions of the voice artificially. People who sing appear to retain physiologically younger voices into advanced age. Their voices are louder and their f_o is categorically higher (Brown, Morris, Hollien, & Howell, 1991).

Vocal Fold Morphology

Mammalian vocal folds (or vocal cords) are generally multi-layered in their tissue construct (Figure 2). An epithelium (skin) encapsulates a soft tissue structure known as the lamina propria, which in itself can have one, two, or three compartments: a superficial layer, an intermediate layer, and a deep layer (Hirano & Sato, 1993). They consist of a matrix of elastin and collagen fibers with interstitial fluids (proteoglycans and glycoproteins) (Gray, Titze, Chan, & Hammond, 1999). Lateral to the deep layer, and firmly attached to it, is the thyroarytenoid (TA) muscle (Hirano, 1975).

The design is functional in that the superficial layer (under the skin) needs to be pliable, like a gel, to support surface modes of vibration (Titze, 1988). These surface modes allow energy to be transferred from the airstream to the tissue. An alternating convergent and divergent glottal duct is produced between the vocal folds every cycle of vibration with these surface modes, such that a push-pull action is exerted on the vocal fold surface by the airstream (Titze, 1988). Deformation of the superficial layer is a requirement for the alternating shape and its resulting push-pull action. Thus, a single gel-like superficial layer underneath the epithelium is a minimum requirement for oscillation. Human infants begin with this morphology, i.e., a thick single-layer *mucosa* (Hirano & Sato, 1993). The deeper layers develop around ages 3 – 4 but a full adult-like lamination may not be achieved until the age of puberty (Sato, Nakashima, Nonaka, & Harabuchi, 2008).

The intermediate and deep layers of the lamina propria form a ligament, with dense collagen and elastin fibers aligned in parallel, like a string. The alignment and strength of the fibers is functionally driven. It has been shown that children whose vocal folds were never vibrated due to cerebral palsy did not develop a vocal ligament (Sato, Umeno, Nakashima, Nonaka, & Harabuchi, 2012). Likewise, a human adult who never phonated past 64 years of age due

to a cerebral hemorrhage, lost the ligament architecture. The vocal fold morphology returned to its infant-like state, with a uniform rather than a layered structure (Sato, Umeno, Ono, & Nakashima, 2011). From these studies it is reasoned that evolution has provided the cell structure (stellate cells at the endpoints of the vocal folds) for multiple layer development, but tissue stress from vocalization is needed to build the structure.

A major benefit of multiple layers is the possibility of a large f_o range. A homogeneous (non-layered) mucosa cannot be stiffened to produce a wide f_o range while at the same time maintaining the surface pliability requirement for energy transfer by surface waves, which sustains oscillation. A fiber-gel structure is ideal for the combined requirement of string-like tension and surface pliability. As the vocal folds grow in length, however, f_o drops because the natural frequencies in the fibers are lowered. This lowering of natural frequencies with laryngeal growth can be countered only by a nonlinear stress-strain relation in the tissue fibers (Titze, 2000), or by active contraction of muscle fibers. For a moderate range of f_o , a two-layer system (mucosa and muscle) suffices. TA muscle fibers can stiffen to provide a workable range of f_o . Such an architecture has been developed by some domestic canines, who bark in a low to medium f_o range.

A much wider pitch range, however, is achieved with development a vocal ligament in a tri-layer system (mucosa, ligament, muscle). Stiffness in muscle fibers can then gradually be shifted to stiffness in ligament fibers as vocal fold length is increased with cricothyroid (CT) muscle activation. The human vocal folds evolved and develop into a tri-layer system, presumably for calling over long distances, combat shouting, screaming, imitating infants or animal sounds, or chanting and singing. Pigs have a ligament for squealing, Rocky Mountain elk for extremely high-pitched “bugling,” and chimpanzees for screaming and pant-hooting. For humans who do not vocalize out of the context of speech, however, the vocal ligament may become superfluous. Surgeons have found that removal of the ligament (known as a cordectomy) may not have a profound effect on conversational speech (Hillel et al., 2013; Mau et al., 2015). Caution should be used, however, in applying this perspective to current surgical management of the vocal ligament. Successful voice outcomes from a few subligamentous cordectomies could be based on secondary (compensatory) effects. At this stage, therefore, it is risky to consider the vocal ligament superfluous in reconstructive surgery for speech. For singing and calling, on the other hand, one can say categorically that ligament removal would likely result in a dramatic impairment.

Physiology of Posturing With Laryngeal Muscles

Muscles that are used for posturing usually consist of an agonist-antagonist pair. One muscle adducts while the other abducts, or one elongates while the other shortens. For optimal strength and control, it is desirable for both muscles to be operating near their sarcomere length (the length at which the actin-myosin overlap produces the maximum contractile force). The *in situ* resting length is often near the sarcomere length (Rack and Westbury, 1969; Vaz et al., 2012). In this resting position, between two boundary attachments, the muscle is stretched slightly by endpoint tendons. From the sarcomere length, the muscle can be shortened or elongated in comparable amounts. In speech, however, there is a tendency to perpetually shorten the vocal folds to keep f_o low. The principle of least effort in speech

production (Traunmuller, H., & Eriksson, A., 1995; Cielo et al., 2010; Holmberg et al., 1994; Lindblom, 2000) may not be observed for the larynx if articulation is of higher priority than range of f_o . The process of adduction of the vocal processes of the arytenoid cartilages is a rocking action on top of the cricoid cartilage (Selbie et al., 1998) that moves the processes ventrally, caudally, and medially (forward, downward, and inward). This shortens the vocal folds 10 – 30 % from the neutral resting length for a range of low (speech-like) fundamental frequencies (Nishizawa et al., 1988). While the vocal folds can also be elongated about 30 % from the resting length to tense the tissue fibers, this lengthening rarely happens in speech.

Fig. 3 shows a muscle activation plot for f_o change produced by two intrinsic muscles of the larynx, the cricothyroid (CT) and the thyroarytenoid (TA). Electromyography (EMG) was conducted on a human male with hooked-wire electrodes (Titze et al., 1989). Six fundamental frequencies were elicited with multiple tokens of vowel phonations. The muscle activities, ranging between 0.0 and 1.0, were normalized to maximum values of contraction by calibration maneuvers (a high squeal for maximum CT and a forceful adduction [as in heavy lifting] for maximum TA). Note that typical male speech fundamental frequencies (98 Hz and 131 Hz) are all in the lower left quadrant of the muscle activation plot. The subject used mainly TA activity to increase f_o in this region, moving from the lowest contour line to the second from the bottom. This range of f_o agrees with the speech range shown in Figure 1. It also agrees with small vocal fold lengths (lower third of the range reported by Nishizawa et al., 1988) for corresponding frequencies.

The stress in the tissue fibers of the TA muscle can be estimated from the frequency of a string under tension, fixed at both ends,

$$f_o = \frac{1}{2L} \sqrt{\frac{\sigma}{\rho}} \quad (1)$$

For a vocal fold length L of 1.0 cm, a fundamental frequency f_o of 150 Hz, and a density ρ of 1.04 g/cm³, the stress σ in the muscle fibers is predicted to be 9 kPa. The exact range of stress in human muscle fibers is not known, but it has been measured in vitro in canine TA muscle to be 0 – 100 kPa (Alipour-Haghighi and Titze, 1985). Assuming similar muscle fibers for human and canine TA muscles, the required stress for speech in humans would be less than 10 % of the maximum active stress achievable at sarcomere length. Alipour-Haghighi and Titze (1985) showed that the maximum stress occurs at 20 – 30 % elongation of the muscle, rather than at shortening.

Chhetri et al. (2012) used anesthetized canines (larynx size comparable to humans) to stimulate the motor nerves of the five intrinsic laryngeal muscles (CT, TA, lateral cricoarytenoid LCA, posterior cricoarytenoid PCA, and interarytenoid IA). The vocal fold was shortened by 2 % at 200 Hz phonation and 12 % at 100 Hz phonation. This study, together with the study of Alipour-Haghighi and Titze (1985) on the same species, make it clear that the TA muscle is biased far from its sarcomere length in the 100 – 200 Hz speech fundamental frequency range. The TA muscle needs to be opposed by the CT muscle to

stretch to sarcomere length. Both canines and humans have a strong CT muscle. Howling by dogs and wolves may accomplish this lengthening if the CT muscle stretches the TA muscle passively. Some head lifting with neck stretching may help in lengthening the vocal folds for howling. Humans can rely on the CT muscle to stretch both the TA muscle fibers and the vocal ligament, but only with high-pitched phonations like singing and calling, but these vocalizations are not generally part of speech.

A secondary f_o -raising mechanism involves lung pressure. Experiments on excised canine larynges (Titze, 1989) have shown that a range of 0 – 2.5 kPa of lung pressure can change f_o by 100 – 200 Hz, depending on the vocal fold length (Fig. 4). This change is attributed to a dynamic stretch increase in the TA fibers with increasing vibrational amplitude. Thus, canines can reach their maximum barking f_o and perhaps sarcomere length, with raised lung pressure. For human speech, a two-layer system (without a ligament) would also appear to be sufficient if “loud” phonation were practiced, as in calling, but calling is practiced mainly at construction sites, at sports events, or in playgrounds. Loud speech, as in noisy restaurants or moving vehicles, is not an effective exercise because it retains the articulatory constraints mentioned earlier.

To achieve the full physiological f_o range shown in Figures 1 and 3, a vocal ligament is available to humans. Collagen fiber stress can be much higher than the active or passive stress in the TA muscle, but it requires substantial vocal fold lengthening (rather than shortening). Figure 5 shows the stress-strain curves for three layers of vocal fold tissue: mucosa, ligament, and 10 % contracted TA muscle. Min et al. (1995) measured the ligament stress on humans. For negative strain, the ligament tissue stress is similar to the mucosal stress. The activated TA muscle alone can control f_o , as discussed above. However, if the CT muscle is activated enough to lengthen the vocal folds to the point where the ligament curve crosses the TA curve, the TA can operate near the sarcomere length (strain near 0.2 – 0.3, or 20–30%) and the ligament is stretched as well. This requires f_o to be in the 300 – 400 Hz range for males (near the top of the voice range profile in Fig. 1).

A good exercise for the laryngeal musculature, and perhaps for the entire framework of the larynx, is to stretch the vocal folds often by raising f_o above speech levels. In human phonation today, calling over long distances for identity and location, as well as uttering intense vocal expression for emotion, tend to be minimized. Singing appears to be one of a few viable modalities for exercising the CT/TA muscle pair. As cited earlier, Brown et al. (1991) showed that singers maintained a higher f_o than non-singers over their entire lifespan.

A second postural issue in speech is frequent and aggressive adduction of the vocal folds. This again involves an imbalance in the use of agonist-antagonist muscles, in this case the LCA/TA combination for adduction and the PCA for abduction. It has been shown that adduction from the neutral position and return to the neutral position occurs on the order of 10,000 times per day in classroom teachers (Titze et al., 2007). This adduction, often quite forceful in speech if a syllable needs to be stressed, or if the teacher needs to increase vocal effort in the presence of classroom noise, involves mostly TA and LCA activation, leaving the antagonist PCA dormant. Rapid inhalation or sniffing, which would engage the PCA, is relatively infrequent laryngeal activity. At high f_o , however, when the vocal folds need to be

stretched, PCA activation helps to stabilize the position of the arytenoid cartilages on the cricoid cartilage, thereby contributing to the vocal fold stretch produced by CT activity. Thus, the CT/PCA combination needs to be activated often to oppose the LCA/TA combination if the entire musculature is to make use of its physiologic range. But again, if f_o is kept perpetually and chronically low, neither CT nor PCA will ever be well exercised.

DISCUSSION ON VOICE THERAPY AND VOCAL TRAINING

Traditionally, voice disorders have been broadly characterized as either organic or functional. Organic disorders have an identified lesion, while functional disorders exhibit voice limitations with a visually-intact system. Spasmodic dysphonia, paradoxical vocal fold movement, muscle tension dysphonia, and the laryngeal component of stuttering show no structural disorder in the larynx, but the muscular system produces abnormal vocal fold posturing and movement. There is as yet no direct link of this abnormal posturing or movement to a condition one might call laryngeal muscle *contracture*, but many therapy techniques for functional voice disorder embrace the idea of stretching the vocal folds with f_o glides (Stemple et al., 1994; Titze and Verdolini-Abbott, 2012), reducing forceful adduction of the vocal folds (Boone, 1996; Colton and Casper, 1996), and stretching of muscles and connective tissue by laryngeal massage (Roy et al., 1998). Muscle contracture in the limbs has been reported as a result of agonist-antagonist imbalance (e.g., paralysis of one or the other) or as a result of abnormal flexing or shortening over extended periods of time. The general impression by clinicians is that vocal fold movement in functional disorders is restricted, locked up, or even paradoxical. Laryngeal massage, chanting, and singing are generally considered to be “unlocking” activities. Patients with Parkinsons Disease have overcome vocal limitations with therapy geared towards louder, more energetic voice production (Ramig et al., 1994). Severely symptomatic spasmodic dysphonia patients can often sing with ease. It is not yet known to what extent laryngeal muscle contracture can be reversed (temporarily or permanently) by therapy.

In addition to these motor control issues, it is known that mechanical stress causes a re-modeling of soft tissue. Fibrocytes respond to stress by replenishing structural proteins. If fibers are not frequently stretched, they will weaken. Exercises known as semi-occluded vocal tract (SOVT) exercises or vocal function exercises (non-speech exercises) have been shown to accomplish simultaneous stretching and unpressing of the vocal folds (Stemple et al., 1994; Titze, 2006; Kapsner-Smith et al., 2014), thereby alleviating vocal fatigue.

Clinicians are not totally in agreement about the context in which vocal habilitation and rehabilitation should take place. General physical exercise principles are *specificity* (building specific muscles for the context in which they will be used) and *overload* (exercising muscles slightly beyond their normal limits for a short period of time). Training in the context of speech, however, is not likely to make ideal use of these principles. As has been shown, low f_o will not elicit an overload on the CT muscle. If extreme f_o glides, vocal fold adductory glides, or long sustained phonation were part of speech, training within the context of speech would be ideal. Given the restricted range of muscle use, however, non-speech exercises may accomplish the task more effectively, with the hope that the results will transfer to speech. Current use of semi-occluded vocal tract exercises or vocal function

exercises for laryngeal muscle training is becoming wide-spread as a means of “taking the voice to the gym” several times a day for a few minutes.

CONCLUSION

Using small and fragmentary data sets, arguments have been presented that current trends in the use of the human larynx for speech runs counter to its fundamental design. Low fundamental frequency, low intensity, and frequent adduction for accent in speech prevent laryngeal muscles from being exercised with the appropriate ranges of motion and optimal lengthening. It has been suggested that voice disorders such as muscle tension dysphonia, spasmodic dysphonia, or other larynx-based dysfluencies in speech, may be linked to the mismatch between mammalian design and contemporary use of the larynx. It would appear, however, that regular exercise *out of the context of speech* could have a health benefit. For some people, this may parallel the need to exercise the spine and postural muscles of the body to overcome the effects of excessive bending, sitting, or otherwise distorting natural equilibria. Vocology, the science and practice of voice habilitation, is a new field that addresses these challenges.

Acknowledgments

Support for this research comes from grant number 5 R01DC012045-04 by the National Institute on Deafness and Other Communication Disorders.

References

- Abdelli-Beruh NB, Wolk L, Slavin D. Prevalence of vocal fry in young adult male American English speakers. *Journal of Voice*. 2014; 28:185–190. [PubMed: 24315658]
- Alipour-Haghighi F, Titze IR. Viscoelastic modeling of canine vocalis muscle in relaxation. *Journal of the Acoustical Society of America*. 1985; 78(6):1939–1943. [PubMed: 4078169]
- Alipour-Haghighi F, Titze IR. Elastic models of vocal fold tissues. *Journal of the Acoustical Society of America*. 1991; 90(3):1326–1331. [PubMed: 1939897]
- Boone, DR. *The voice and voice therapy*. 5. Englewood Cliffs, NJ: Prentice-Hall, Inc; 1996.
- Bradlow AR, Toretta GM, Pisoni DB. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*. 1997; 20:255–272.
- Brown W, Morris RJ, Hollien H, Howell E. Speaking fundamental frequency characteristics as a function of age and professional singing. *Journal of Voice*. 1991; 5:310–315.
- Byrd D. Relation of sex and dialect to reduction. *Speech Communication*. 1994; 15:39–54.
- Chhetri DK, Neubauer J, Berry DA. Neuromuscular control of fundamental frequency and glottal posture at phonation onset. *Journal of the Acoustical Society of America*. 2012; 131(2):1401–1412. [PubMed: 22352513]
- Cielo CA, Elias VS, Brum DM, Ferreira FV. Thyroarytenoid muscle and vocal fry: a literature review. *Revista Da Sociedade Brasileira De Fonoaudiologia*. 2010; 16(3):362–369.
- Collister LB, Huron D. Comparison of Word Intelligibility in Spoken and Sung Phrases. *Empirical Musicology Review*. 2008; 3(3):109–125.
- Colton, RH., Casper, J. *Understanding voice problems: a physiological perspective for diagnosis and treatment*. 2. Baltimore MD: Williams & Wilkins; 1996.
- Fant, G. *The Acoustic Theory of Speech Production*. The Hague: Moulton; 1960.
- Gramming, P. *The phonetogram: an experimental and clinical study*. Malmö, Sweden: Malmö General Hospital; 1988.

- Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. *Journal of the Acoustical Society of America*. 1989; 83:2352–2360.
- Gray S, Titze IR, Chan R, Hammond T. Vocal fold proteoglycans and their influence on biomechanics. *Laryngoscope*. 1999; 109:845–854. [PubMed: 10369269]
- Hillel AT, Johns MM III, Hapner ER, Shah M. Voice outcomes from subligamentous cordectomy for early glottic cancer. *Annals of Otolaryngology, Rhinology, and Laryngology*. 2013; 122(3):190–196.
- Hirano M. Phonosurgery: basic and clinical investigations. *Otologia (Fukuoka)*. 1975; 21:239–440.
- Hirano, M., Sato, K. *Histological Color Atlas of the Human Larynx*. San Diego: Singular Publishing Group, Inc; 1993.
- Holmberg EB, Perkell JS, Hillman RE, Gress C. Individual variation in measures of voice. *Phonetica*. 1994; 51:30–37. [PubMed: 8052674]
- Hunter EJ, Titze IR. Variation of intensity, fundamental frequency, and voicing for teachers in occupational versus non-occupational settings. *Journal of Speech Language Hearing Research*. 2010; 53(4):862–875.
- Joliveau E, Smith J, Wolfe J. Tuning of vocal tract resonances by sopranos. *Nature*. 2004; 427:p116.
- Kent RD. Anatomical and neuromuscular maturation of the speech mechanism Evidence from acoustic studies. *American Journal of Anatomy*. 1976; 151:11–20.
- Kuhl PK. A new view of language acquisition. *Proceedings of the National Academy of Science*. 2000; 97(22):11850–11857.
- Lieberman, P. *Uniquely Human: the evolution of speech, thought, and selfless behavior*. Cambridge, MA: Harvard University Press; 1991.
- Lindblom, B. Emergent phonology. *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*; 2000. p. 195-209.
- Mau T, Palaparthi A, Riede T, Titze IR. Effect of resection depth of early glottic cancer on vocal outcome: An optimized finite element simulation. *Laryngoscope*. 2015; 125(8):1892–1899. [PubMed: 26010240]
- Min YB, Titze IR, Alipour-Haghighi F. Stress-strain response of the human vocal ligament. *Annals of Otolaryngology, Rhinology, and Laryngology*. 1995; 104(7):563–569.
- Nishizawa, N., Sawashima, M., Yonemoto, K. Vocal fold length in vocal pitch change. In: Fujimura, O., editor. *Vocal fold Physiology: voice production, mechanisms and functions*. New York, NY: Raven Press; 1988. p. 75-82.
- Rack PM, Westbury DR. The effects of length and stimulus rate on tension in the isometric cat soleus muscle. *Journal of Physiology*. 1969; 204(2):443–460. [PubMed: 5824646]
- Ramig LO, Bonitati CM, Lemke JH, Horii Y. Voice treatment for patients with Parkinson disease: Development of an approach and preliminary efficacy data. *Journal of Medical Speech-Language Pathology*. 1994; 2(3):191–209.
- Roy N, Bless DM. Manual circumlaryngeal techniques in the assessment and treatment of voice disorders. *Current Opinion in Otolaryngology & Head and Neck Surgery*. 1998; 6(3):151–155.
- Sato K, Nakashima T, Nonaka S, Harabuchi Y. Histopathologic investigations of the unphonated human vocal fold mucosa. *Acta Oto-laryngologica*. 2008; 128:694–701. [PubMed: 18568507]
- Sato K, Umeno H, Ono T, Nakashima T. Histopathologic study of human vocal fold mucosa unphonated over a decade. *Acta Oto-laryngologica*. 2011; 131:1319–1325. [PubMed: 21916773]
- Sato K, Umeno H, Nakashima T, Nonaka S, Harabuchi Y. Histopathologic investigations of the unphonated human child vocal fold mucosa. *Journal of Voice*. 2012; 26(1):37–43. [PubMed: 21227641]
- Selbie WS, Zhang L, Levine WS, Ludlow CL. Using joint geometry to determine the motion of the cricoarytenoid joint. *Journal of the Acoustical Society of America*. 1998; 103:1115–1127. [PubMed: 9479765]
- Stemple JC, Lee L, D’Amico B, Pickup B. Efficacy of vocal function exercises as a method of improving voice production. *J Voice*. 1994; 8(3):271–278. [PubMed: 7987430]
- Sundberg J, Ternstrom S. Commentary on “Comparison of Word Intelligibility in Spoken and Sung Phrases” by Lauren Collister and David Huron. *Empirical Musicology Review*. 2008; 3(4):215–217.

- Taylor AM, Reby D. The contribution of source-filter theory to mammal vocal communication research. *Journal of Zoology*. 2009; 280:221–236.
- Titze IR. The physics of small-amplitude oscillation of the vocal folds. *Journal of the Acoustical Society of America*. 1988; 83(4):1536–1552. [PubMed: 3372869]
- Titze IR. On the relation between subglottal pressure and fundamental frequency in phonation. *Journal of the Acoustical Society of America*. 1989; 85(2):901–906. [PubMed: 2926005]
- Titze, IR. *Principles of Voice Production*. Vol. Chapter 8. Denver, CO: National Center for Voice and Speech; 2000.
- Titze, IR. *The Myoelastic-Aerodynamic Theory of Phonation*. Salt Lake City, Utah: National Center for Voice and Speech; 2006.
- Titze IR, Luschei ES, Hirano M. Role of the thyroarytenoid muscle in regulation of fundamental frequency. *Journal of Voice*. 1989; 3(3):213–224.
- Titze IR, Hunter EJ, Švec JG. Voicing and silence periods in daily and weekly vocalizations of teachers. *Journal of the Acoustical Society of America*. 2007; 121(1):469–478. [PubMed: 17297801]
- Titze IR, Riede T, Popolo P. Nonlinear source-filter coupling in phonation: Vocal Exercises. *Journal of the Acoustical Society of America*. 2008; 123(4):1902–1915. [PubMed: 18396999]
- Titze, IR., Verdolini-Abbott, K. *Vocology: the science and practice of voice habilitation*. Salt Lake City, UT: National Center for Voice and Speech; 2012.
- Traunmuller H, Eriksson A. The frequency range of the voice fundamental in the speech of male and female adults. 1995 Unpublished Manuscript.
- Vaz MA, de la Rocha FC, Leonard T, Herzog W. The force-length relationship of the cat soleus muscle. *Muscles Ligaments Tendons J*. 2012; 2(2):79–84. [PubMed: 23738278]
- Wolk L, Abdelli-Beruh NB, Slavin D. Habitual use of vocal fry in young adult female speakers. *Journal of Voice*. 2012; 26:e111–e116. [PubMed: 21917418]

Figure 1.
Voice range profile with superimposed speech range contours of school teachers (after Hunter et al., 2010).

Fig. 2.
Layered structure of human vocal folds (after Hirano, 1975).

Fig. 3.
Muscle activation plot (MAP) for a human male subject (after Titze et al., 1989).

Figure 4. Increase in f_o achievable with increased lung pressure. Data are from canine larynges. (after Titze, 1989).

Figure 5.
Stress-strain curves for human vocal fold tissue layers. (After Alipour-Haghighi and Titze, 1985; 1991).