Published in final edited form as:

Speech Commun. 2018 July ; 101: 26–33. doi:10.1016/j.specom.2018.05.003.

WHERE HAS ALL THE POWER GONE? ENERGY PRODUCTION AND LOSS IN VOCALIZATION

Ingo R. Titze^{1,2}

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

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

Abstract

Human voice production for speech is an inefficient process in terms of energy expended to produce acoustic output. A traditional measure of vocal efficiency relates acoustic power radiated from the mouth to aerodynamic power produced in the trachea. This efficiency ranges between 0.001 % and 1.0 % in speech-like vocalization. Simplified Navier-Stokes equations for non-steady compressible airflow from trachea to lips were used to calculate steady aerodynamic power, acoustic power, and combined total power at seven strategic locations along the airway. A portion of the airway was allowed to collapse to produce self-sustained oscillation for sound production. A *conversion efficiency*, defined as acoustic power generated in the glottis to aerodynamic power dissipated, was found to be on the order of 10%, but wall vibration, air viscosity, and kinetic pressure losses consumed almost all of that power. This sound, reflected back and forth in the airway, was dissipated at a level on the order of 99.9 %.

Keywords

Vocal power; vocal efficiency; oral pressure; vocal effort

Vocalization involves the conversion of several forms of energy into acoustic energy. Metabolic energy is used to initiate and maintain muscle contractions, aerodynamic energy is produced in the pulmonary system to drive an airstream through the vocal tract, elastic energy is stored and retrieved in stretched tissues, and kinetic energy is developed in tissue and air movement during oscillation of the vocal folds. The efficiency of conversion of these energies into acoustic energy in the form of sound waves is a process not often addressed in voice and speech science. The likely reason is that, for humans, acoustic energy levels are so low compared to energy levels used for body movement or other physical activities, that they appear insignificant when viewed in competition with other energy needs. For smaller species that vocalize at similar loudness levels, like birds and small mammals, efficiency is a more important issue¹.

Corresponding Author: Ingo R. Titze, 136 South Main Street, Suite 320, Salt Lake City UT 84101-3306, ingo.titze@utah.edu.

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 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.

A traditional vocal efficiency measure^{2,3} is calculated as the ratio of oral radiated acoustic power to aerodynamic power in the trachea. This measure has appeal because it relates a useful acoustical output to an effort input (lung pressure, or more precisely, alveolar pressure). Effort in vocalization is of practical use for speaking and singing in humans, and for all forms of vocalization in other species.

A seminal study on glottal efficiency was conducted by Schutte³ on 45 subjects with no voice pathology and 64 subjects with various disorders. Lung pressure, mean tracheal flow, sound pressure level, and efficiency were measured for multiple repetitions of similar vocalizations. Lung pressure was derived from esophageal pressure, and flow was measured using a flow-head (≈ 3 cm diameter and 15 cm long tube) held between the lips during phonation. Vocal efficiency was measured as a function of three variables, sound pressure level (SPL), mean airflow rate, and lung pressure. Schutte's data did not produce much of a trend for efficiency versus mean flow rate, but there was a strong dependence of efficiency on lung pressure and SPL. In other words, voices became more efficient as they became louder, but only up to the point where excessive losses in the airway overcame the increase in acoustic power produced. Efficiency clustered around 0.01%, with values as high as 0.3% reached in Schutte's data. (Earlier, Bouhuys et al.² reported a values as high as 2.0% for a singer). Schutte's results also suggested that for the same lung pressure, efficiency is lower in voice-disordered subjects because more glottal flow is needed to maintain similar sound pressure levels. Even though both sets of subjects produced similar SPL values, mostly in the range of 72 to 85 dB range, there were regions of no overlap where normal subjects produce SPL above 85 dB and disordered subjects produced sound levels below 72 dB in similar lung pressure ranges.

In a very recent study⁴, a pressure conversion ratio was defined that relates to vocal efficiency and has the potential of being more clinically feasible. Many previous practical approaches have been hampered by the difficulty of simultaneous measurement of tracheal pressure and flow^{5,6,7,8,9}. By using a small fixed opening at the mouth (the lips closed around a flow-release cannula and a pressure sensor cannula) and measuring both the acoustic and the steady pressure behind the lips, the pressure ratio can be used to estimate the glottal efficiency. This study gives new hope for practical applications of efficiency measurements in voice production.

The purpose of this paper was to conduct a numerical calculation of aerodynamic and acoustic power created and absorbed along the entire airway (trachea to lips). The solution is based on the Navier-Stokes equation for non-steady, compressible airflow interacting with a soft-walled structure that can collapse and self-sustain oscillation for sound production and wave propagation. The generally low efficiency and its wide range (at least two orders of magnitude) across speakers with different airway conditions set up the questions for this study: (1) where in the airway is the power created and absorbed, and (2) what control does a vocalist have over vocal output power and efficiency at a given fundamental frequency? A new quantity, conversion efficiency, is defined. It relates to energy conversion at the location where vibration takes place. Owing to the large number of parameters involved, this study was limited to one natural frequency of oscillation, a typical male speech frequency. It is already known that vocal efficiency increases with fundamental frequency^{2,3,4}. This increase

is primarily attributed to frequency-dependent radiation efficiency, which is not the focus of this experiment.

METHODS

The National Center for Voice and Speech has developed a software package *VoxInSilico* for simulation of sound production in airways (general mathematical detail described in Chapter 6 of a textbook entitled *The Myoelastic Aerodynamic Theory of Phonation*²⁶, and later references^{10,15,16}). It includes algorithms to solve the Navier-Stokes equation, trachea to lips, in a hybrid analytical and computational paradigm. The analytical approach is used to approximate pressure and flow variations in the radial and azimuthal directions of all vocal tract sections, whereas a numerical solution is used for the axial direction (the primary flow direction). This hybrid approach is supported by empirical data for vortex shedding and turbulence¹¹, which requires an inordinate amount of computation if empirical short-cuts are not used. The effect of vorticity and turbulence in rapidly expanding sections is included in boundary condition between adjacent sections¹⁰, including all glottal sections. The approach has been validated in terms of resonance-frequency and bandwidth calculations in a benchmark paper¹⁰.

A mesh of 107 elements (sections) along the airway axis was used, 36 sections for a trachea, 2 sections for a transition into a vocal fold glottis, 5 sections for the glottis itself, and 44 sections for the supraglottal airway (Fig.1).

The model includes a 28 section nasal tract, but the velar port was kept closed for this study, preventing any nasalization. For the current study, non-steady compressible flow in soft-walled tubes was assumed so that fluid-structure interaction took place along the entire airway¹⁰. Cross-sections vary from circular in the trachea to highly elliptical in the glottis so that airway collapse could occur (Fig. 2). After glottal exit, the cross sections were circular again. Air inertance, air compliance, wall vibration losses, and boundary-layer viscosity were calculated within each section¹⁰.

The viscoelastic wall properties of the vocal tract (excluding the glottis, which is decussed below) were chosen after Ishizaka et al.¹²: Young's modulus within each section = 9.62 kPa, shear modulus between sections = 1.67 kPa, mass per unit area = 1.5 g/cm², and damping ratio = 1.26. The damping ratio produced over-damping of the walls that were not intended to be self-oscillating. This ratio is defined as $\omega_o \eta / E$, where ω_o is the natural frequency of the wall (chosen as 130 Hz), η is the viscosity of the wall tissue, and E is the Young's modulus. These values were chosen as nominal (baseline) in all sections of the airway except in the glottal sections.

The vocal tract shape above the larynx was that of a neutral /A/ vowel¹³, as shown in Figure 1. The first six sections of the supraglottal airway (the first 2.4 cm, from position 4 to 5) were subdivided into the ventricle (1 section at position 4), the false-fold glottis (1 section), and the laryngeal vestibule (4 sections leading up to position 5). Each section was nominally 4.0 mm in length and the cross-sectional areas were 0.8 cm² for the ventricle, 0.4 cm² for the false fold glottis, and 0.5 cm² for the vestibule. The tracheal airway geometry below the

larynx was modeled after MRI measurements by Story et al.¹³, with a linear 0.8 cm long transition (2 sections) from a circular geometry in the last section of the trachea to an elliptical geometry for the first section of the glottis

Given that vocal fold morphologies are highly variable across species, gender, and age, the 5-section sound source of Fig. 2 was sufficiently generic for surface-wave oscillation and some curvature on the vocal fold surfaces for variable adduction. The five short sections (1.6 mm in length, barely visible at location 3 in Fig 1) were adjusted for self-sustained oscillation with elastic, viscous, and inertial wall properties. No Bernoulli flow assumptions (i.e., incompressibility, laminar flow, and energy conservation) were made in these sections. The Navier–Stokes equations were solved the same way as in the vocal tract sections. The geometry of the elliptical laryngeal sections was derived from horizontal cross-sections reported by Hirano and Sato¹⁴. Each of the five narrow ellipses had a 10 mm major diameter and the minor diameters became the pre-phonatory glottal widths: 1.0, 0.8, 0.6, 0.6, 1.0 mm from inferior to superior, as shown in Fig. 2. The vocal fold thickness thus became $5 \times 1.6 \text{ mm} = 8.0 \text{ mm}$. A more elaborate finite-element vocal fold model^{15,16} could have been used for this study, but the generality and simplicity of having a continuous airway in which any section could be a potential sound source would have been lost, keeping in mind a variety of species, genders, and ages with different airways and vocal fold morphologies.

The viscoelastic parameters for the five self-oscillating glottal wall sections were chosen to approximate the lumped-element parameters of the two-mass model of Ishizaka and Flanagan¹⁶: Young's modulus = 4.0 kPa (represented by a spring perpendicular to the wall in Fig. 2), shear modulus = 1.0 kPa (represented by a spring coupling adjacent sections), and mass per unit area = 0.3 g/cm² (represented by the wall blocks). The nominal damping ratio was chosen to be 0.3 for all glottal sections (an average of the Ishizaka and Flanagan values; they used 0.1 for a lower mass and 0.6 for an upper mass). With these tissue parameters and glottal configuration, the glottal wall sections had a natural frequency of 130 Hz. The nominal value for lung pressure was chosen to be 1.0 kPa. Vocal fold collision always occurred in at least one section. Changes in adduction were implemented by adding and subtracting a fraction of a mm from the five minor diameters of the ellipses.

Fig. 3 shows an example of the simulated waveforms used in the pressure and power calculations. The nominal configuration was used in these simulations. On the left panel, below the vocal tract shape, top to bottom, we see contact area *ca*, glottal area *ga*, glottal flow *ug*, and the glottal flow derivative *dug*. The units are indicated on the vertical axes. On the right panel, top to bottom we see oral radiated pressure *Po*, oral flow *Uo*, pressure into the epilarynx tube *Pe*, intraglottal pressure *Pg*, and subglottal pressure *Ps*. All pressures are in kPa and flows are in L/s.

Variations in the simulation consisted of systematically changing losses in the vocal tract, altering the lung pressure, and altering the adduction. Power calculations were conducted in 7 selected locations along the airway as shown in Fig. 1: the entry to the trachea (1), the beginning of the transition into the glottis (2), within the glottis (3), the entry into the supraglottal airway (4), the entry into the pharynx (5), behind the lips (6), and at the radiation point from the lips (7).

The power calculations were as follow:

$$P = \frac{1}{N} \sum_{n=1}^N P_n U_n \quad \text{total power} \quad (1)$$

$$P_{DC} = P_{DC} U_{DC} \quad \text{steady flow (DC) power} \quad (2)$$

$$P_{AC} = P - P_{DC} \quad \text{acoustic (AC) power} \quad (3)$$

where n is the time sample index, N is the number of samples simulated (22050 in a 0.5 s window), P_n is the instantaneous pressure in a given section, U_n is the instantaneous flow, P_{DC} is the steady (DC) pressure, and U_{DC} is the steady (DC) flow, computed as time-averages over the 0.5 s window.

The instantaneous power $P_n U_n$ in Equation (1) varied dramatically in the five glottal sections where self-oscillation took place. Therefore, to get a representation of the mean intraglottal pressure, flow, and power, the calculations in Equation (1) were averaged over the five adjacent glottal sections.

RESULTS

All pressures are reported in kPa and all powers in mW. Figure 4(a) shows the steady (DC) pressure profile from the entry of the trachea to the lips. Note that almost all the DC pressure is dropped in the glottis, from point 2 to point 4; furthermore, the DC pressure has dropped to zero at the lips, as expected with an open mouth. The DC power in Fig. 4(b) follows the shape of the DC pressure because the DC flow is constant. [After some initial wall expansion of the trachea in the first 100 ms, which made the tracheal flow slightly larger than the downstream flow, the steady flow was continuous in all sections at a value of 204 cm³/s].

The total power in Fig. 4(c) has a profile nearly identical to the DC power, but there is a slightly greater decline in total power from entry (point 2) to the center of the glottis (point 3). This is a region where acoustic power is generated, which subtracts from the total power. *By conventional power definitions, positive power is power dissipated, whereas negative power is power generated by a source.*

Thus, the AC power in Fig. 4(d) is in fact negative at point 3, in the amount of −24 mW. Relating this power generated to the total 210 mW power input at the trachea introduces the measure of *power conversion efficiency*, which is on the order of 11 %. The remaining 89 % of the total power is not converted to acoustic power, but is dissipated aerodynamically. More importantly, little of the 11 % converted power is radiated from the lips. The radiated

power is so low (0.034 mW) that it cannot be distinguished from zero in Fig. 4. However, Table 1 shows all power values at the 7 locations for the nominal configuration chosen. Consider the AC power in the bottom row of the table. Beginning with location 4 (entry into the supraglottal airway), no more AC power is generated, and the power available for distribution to the downstream airway is on the order of 6 mW. In the mouth behind the lips (location 6), this power is reduced to 0.7 mW, and only 4.7 % of that oral power is radiated (0.034 mW, location 7). Interestingly, the AC power distributed backwards to the entry of the trachea (0.066 mW, location 1) is about twice the AC power radiated from the mouth.

A. Sensitivity to Increase in Vocal Tract Wall Viscoelasticity

As a first variation of the nominal configuration, the combined stiffness and damping of the vocal tract walls (excluding the glottal sections) was increased to simulate a more rigid wall. It is typical for the elastic and viscous moduli to co-vary in biological tissues¹⁸. Hence, the Young's modulus, shear modulus, and damping ratio of the subglottal and supraglottal airway walls were all increased by a factor of 5. Fig. 5 shows the results. Solid red lines are for the stiffer walls and dashed blue lines are for the original (nominal) case. The effect of the stiffer and more viscous wall was to limit the DC airflow into the trachea from 204 cm³/s to 173 cm³/s (less tracheal wall expansion due to positive tracheal pressure), which then limited the power available at the glottis (location 2). The AC power created in the glottis into the supraglottal tract decreased from 24 mW to 14 mW and the radiated power decreased from 0.034 mW to 0.025 mW (not visible).

B. Sensitivity to Air Viscosity in the vocal tract

Reduction of air viscosity from 0.000186 poise (warm air) to 0.0 poise (no air friction) in the vocal tract had little effect on the overall pressure and power contours, as shown in Fig 6, but there were small acoustic power changes. Although the power into the supraglottal vocal tract was reduced (3.9 mW instead of 6.0 mW) as shown in location 4 of Fig. 6(d), less dissipation in the supraglottal tract caused the radiated power to increase from 0.034 mW to 0.060 mW (too small to be seen in the figure). Thus, air viscosity does increase power consumption a small amount, but it is not a major factor in vocal efficiency and vocalists have no control over it.

C. Sensitivity to Vocal Fold Tissue Viscosity

A major difference in power generation was found when the damping ratio in the vibrating vocal folds, which is proportional to tissue viscosity, was increased a small amount, from 0.3 to 0.4. This is shown in Fig 7. Although the overall profiles for pressure and power distribution did not change much, the AC power produced in the glottis was reduced from 24 mW to 15 mW, with the conversion efficiency changing from 11% to 7%. The power radiated decreased by a factor of 2, from 0.034 mW to 0.017 mW. Increasing the damping ratio further to 0.5 created very small oscillation. The AC power produced at the glottis was a mere 7.0 mW and the power radiated was a mere 0.00015 mW.

Reducing the damping ratio to 0.2 (not shown) increased all the acoustic powers proportionately, but the power calculations were error-prone because the oscillations became aperiodic with less damping. In all, acoustic power generation in the airway was very

sensitive to the viscous properties of the vibrating tissue. Viscosity of biological tissues can vary over several orders of magnitude, but here a very small range of damping ratios (0.2 – 0.5) appeared to bracket the range for self-sustained oscillation.

D. Sensitivity to Lung Pressure

Lung pressure is the primary variable available for control of aerodynamic and acoustic power by the individual if fundamental frequency is not changed. Unlike wall stiffness, air viscosity, and vocal fold tissue viscosity, which are generally not under direct control, lung pressure can be changed voluntarily over a wide range. The value 1.0 kPa is typical for moderately loud speech¹⁹, but shouting and singing often requires lung pressures well above 2.0 kPa²⁰. For these larger lung pressures, nonlinear tissue elasticity and vocal fold collision forces generally limit the amplitude of vocal fold vibration. While these forces have been dealt with in detail in sophisticated finite element modeling of vocal fold structure¹⁶, they are not included in the current wall vibration model. For this reason, we limited lung pressure here to the 0.5 – 2.0 kPa range. Fig 8 shows results for this range.

In Fig. 8(a), the lung pressures are the values shown at location 1. Fig 8(b) shows that aerodynamic (DC) power increases by a factor of 4 for every doubling of lung pressure. The reason is that DC flow also doubles with a doubling of pressure, and since power is the product of pressure and flow, the power quadruples. The total power in Fig 8(c) and the AC power in Fig 8(d) followed a similar pattern, although the AC power increased by only a factor of 3 when lung pressure increased from 1.0 kPa to 2.0 kPa. Aside from this small deviation, all powers in the vocal tract basically quadrupled with each doubling of lung pressure.

E. Sensitivity to Vocal Fold Adduction

The five vocal fold sections were adducted in greater and lesser amounts by increasing the minor diameters of the ellipses. The glottal surface contour remained the same. Thus, for greater adduction the diameter set decreased from [1.0, 0.8, 0.6, 0.6, 1.0] mm to [0.6, 0.4, 0.2, 0.2, 0.6] mm. For lesser adduction, the diameter set increased to [1.6, 1.4, 1.2, 1.2, 1.6] mm. Fig. 9 shows the results. Lesser adduction (a wider glottis) reduced the glottal flow resistance, which meant that steady flow increased and the aerodynamic power increased. The AC power produced at the glottis was 24 mW for nominal adduction, 40 mW for lesser adduction, and 4 mW for greater adduction. The radiated power was 0.03 mW for nominal adduction, 0.06 mW for lesser adduction, and 0.006 mW for greater adduction. The conversion efficiency was 12 % for nominal adduction, 3.7 % for greater adduction, and 13 % for lesser adduction. Thus, it appears that power production and conversion efficiency can be regulated by adduction, but more adduction does not necessarily produce greater conversion efficiency.

SOUND LEVEL IN THE VOCAL TRACT

The AC power along the vocal tract can be expressed in terms of sound level (SL). If sound level is computed with sound intensity I , then

$$SL = 10 \log_{10}(I/I_0) = 10 \log_{10}[P/(AI_0)] \quad (4)$$

where I_0 is the reference intensity (10^{-12} W/m²), P is the power, and A is the cross sectional area of a section of the vocal tract. Table 2 shows this calculation for the seven locations along the airway for the three lung pressures. For 0.5 kPa lung pressure, the SL is 105 dB at the input to the trachea, 110 dB at the end of the trachea, 147 dB at the center of the glottis, 130 dB at the entrance to the vocal tract, 126 dB at the entrance to the pharynx, 117 dB behind the lips, and 103 dB at the point of radiation from the lips. For the 2.0 kPa lung pressure, the SL rises about 10–15 dB at every location, with a maximum of 158 dB in the glottis. It has been established that the threshold of pain for hearing is on the order of 130–140 dB^{21,22}. This level is reached in the upper larynx and lower pharynx of the airway with relatively moderate lung pressures. It is also known that the eardrum can burst with environmental noise on the order of 160 dB^{21,22}. This sound level is reached inside the glottis with higher lung pressure. Table 2 also shows the conversion efficiency (column 4) and the vocal efficiency (last column). Note that the highest lung pressure did not produce the greatest efficiency. The intermediate value (1.0 kPa) produced a conversion efficiency of 11.4 % and a vocal efficiency of 0.016 %.

COMPARISON WITH REPORTED MEASUREMENTS

Results of the empirical study on vocal efficiency conducted by Schutte (1980) on 45 normal subjects are shown in Figure 10. Vocal efficiency is plotted against lung pressure (left) and mean flow rate in the airway (right). The individual points correspond to the middle of the intensity and pressure ranges for each subject (no phonations for shouting, singing, or close to whispering). The solid lines are not regression lines, but results from the current simulation. The close agreement was unexpected, given that the vocal fold model was rather unsophisticated, a 5-section collapsible wall with elliptical cross-sections. The results indicate a high degree of confidence in the numerical simulation.

DISCUSSION AND CONCLUSIONS

The subject of vocal power and efficiency in vocalization, sparsely mentioned in the literature, has been re-addressed. Momentum and mass conservation were used to quantify airflow and tissue movement by solving the Navier-Stokes equation and the continuity equation along the airway. Power calculations revealed how aerodynamic energy was converted to acoustic energy and how this energy was distributed and absorbed in the vocal tract and finally radiated from the mouth.

For a moderate range of lung pressure (0.5 – 2.0 kPa), it was found that only about 10 % of the aerodynamic power entering the trachea was converted to acoustic power in the glottis. The remaining 90 % of the aerodynamic power was dissipated kinetically by loss of pressure recovery in constrictions and expansions of the airway (specifically the glottis). Furthermore, it was found that a very small percentage of the acoustic power generated (0.1 %) was radiated from the lips. Most of the acoustic power remained in the airway in the

form of standing waves was ultimately dissipated. The classical vocal efficiency (radiated power divided by aerodynamic input power) in simulated speech-like productions is on the order of 0.01 %, which is in remarkable agreement with empirical findings by Schutte (1980). Related to question (2) in the Introduction, control over conversion efficiency by the vocalist is mainly in terms of lung pressure and adduction at constant fundamental frequency.

Are there benefits derived from low conversion efficiency? Two likely benefits can be mentioned. First, the loss of pressure and power across the glottis is essential for self-sustained oscillation of the vocal folds. A push-pull mechanism is needed to transfer energy from the airstream to the moving vocal fold tissue surface. The push is needed for lateral movement (glottal opening) and a lesser push, or preferably a pull, is needed for medial movement (glottal closing). It has been shown repeatedly^{23,24} that an alternating convergent-divergent glottal shape can create this push-pull, but only if the pressure against the glottal surface does not recover during glottal closing. Hence, strong kinetic energy losses in a divergent glottis facilitate vocal fold vibration. Second, the fact that most of the acoustic energy remains in the vocal tract (is not radiated to the listener) allows for much variation in sound quality and phonetic variation. Vowels and consonants are distinguished by their sound frequency spectrum, which in turn is governed by the standing waves in the vocal tract. If sound were not reflected almost 100 % from the lips back into the vocal tract, there would be poor distinction between standing wave pressure patterns produced with variations in tongue, jaw, and lip gestures. Hence, speech intelligibility is facilitated by poor vocal efficiency.

A comment is offered for future investigations. For non-humans species that do not speak, efficiency of vocalization may be a bigger issue. Animals that communicate vocally over long distances may prefer to limit the inventory of contrasting vocal sounds in favor of louder sounds. Even in humans, the efficiency of shouting and unamplified singing shows a sharp contrast with speech. Klingholz et al.²⁵ reported sound pressure level (SPL) values as high as 110 dB at 30 cm from the mouth. That corresponds to a radiated acoustic power of 113 mW at the lips if the assumption is made that sound radiates spherically. This is a factor of 1000 greater than the radiated pressures computed here for speech-like conditions. Bouhys et al.² measured the vocal efficiency of a singer in a manner similar to Schutte's technique³ and found it to be 2 %, which is on the order of 200 times the efficiency calculated here (and corroborated by measurement) for speech-like vocalizations. The power calculations conducted here will hopefully be useful as a benchmark for future computational modeling of voice production, particularly at higher fundamental frequencies.

Acknowledgments

Support for this research comes from grant number 5R01 DC012045-04 by the National Institute on Deafness and Other Communication Disorders.

References

1. McCracken KG, Sheldon FH. Avian vocalizations and phylogenetic signal. *Proc Natl Acad Sci.* 1997 Apr 15; 94(8):3833–3836.

2. Bouhuys A, Mead J, Proctor DF, Stevens KN. Pressure-flow events during singing. *Ann NY Acad Sci.* 1968; 155:165–176.
3. Schutte, H. The efficiency of voice production. Groningen: State University Hospital; 1980.
4. Titze IR, Maxfield L, Palaparthi A. An oral pressure conversion ratio as a predictor of vocal efficiency. *J. Voice.* 2016; 30(4):398–406. [PubMed: 26164123]
5. Smitheran JR, Hixon TJ. A clinical method for estimating laryngeal airway resistance during vowel production. *J Speech Hear Disord.* 1981; 46:138–146. [PubMed: 7253590]
6. Rothenberg M. Interpolating subglottal pressure from oral pressure. *J Speech Hear Disord.* 1982; 47:219–223. [PubMed: 7176602]
7. Hertegard S, Gauffin J, Lindestad P-A. A comparison of subglottal and intraoral pressure measurements during phonation. *J Voice.* 1995; 9:149–155. [PubMed: 7620537]
8. Kitajima K, Fujita F. Estimation of subglottal pressure with intraoral pressure. *Acta Otolaryngol.* 1990; 109:473–478. [PubMed: 2360453]
9. Lofqvist A, Carlborg B, Kitzing P. Initial validation of an indirect measure of subglottal pressure during vowels. *J Acoust Soc Am.* 1982; 72:633–635. [PubMed: 7119286]
10. Titze IR, Palaparthi AKR, Smith SL. Benchmarks for time-domain simulation of sound propagation in soft-walled airways: steady configurations. *J Acoust Soc Am.* 2014; 136(6):3249–3261. [PubMed: 25480071]
11. Scherer RC, Torkaman S, Kucinski BR, Afjeh AA. Intraglottal pressures in a three-dimensional model with a non-rectangular glottal shape. *J Acous. Soc Am.* 2010; 128(2):828–838.
12. Ishizaka K, French IC, Flanagan JL. Direct determination of vocal tract wall impedance. *IEEE Trans Acoust Speech Sign Process.* 1975; 23:370–373.
13. Story BH. A parametric model of the vocal tract area function for vowel and consonant simulation. *J Acoust Soc Am.* 2005; 117(5):3231–3254. [PubMed: 15957790]
14. Hirano, M, Sato, K. *Histological Color Atlas of the Human Larynx.* Singular Publishing; San Diego: 1993.
15. Alipour F, Berry DA, Titze IR. A finite-element model of vocal-fold vibration. *J Acoust Soc Amer.* 2000; 108(6):3003–3012. [PubMed: 11144592]
16. Titze IR, Palaparthi A, Alipour F, Blake D. 2017; Comparison of a Fiber-gel finite element model of vocal fold vibration to a transversely isotropic stiffness model. *J. Acoust. Soc. Amer.* 142(3): 1376–1383. [PubMed: 28964045]
17. Ishizaka K, Flanagan JL. 1972; Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell System Technical Journal.* 51(6):1233–1268.
18. Fung, YC. *Bionics: Mechanical Properties of Living Tissues.* Second. Springer Verlag; New York: 1993.
19. Holmberg EB, Hillman RE, Perkell JS. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am.* 1988; 84:511–529. [PubMed: 3170944]
20. Sundberg J, Scherer R, Titze I. 1993; Phonatory control in male singing. A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. *J. Voice.* 7(1):15–29. [PubMed: 8353616]
21. Franks, JR, Stephenson, MR, Merry, CJ, editors *National Institute for Occupational Safety and Health. Jun, 1996. Preventing Occupational Hearing Loss - A Practical Guide (PDF);* 881996 June
22. Newman, EB. *American Institute of Physics handbook.* New York: McGraw-Hill; 1972. *Speech and Hearing;* 3–155. 1972
23. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Amer.* 1988; 83(4):1536–1552. [PubMed: 3372869]
24. Thomson SL, Mongeau L, Frankel SH. Aerodynamic transfer of energy to the vocal folds. *J Acoust Soc Amer.* 2005; 118(3 Pt 1):1689–700. [PubMed: 16240827]
25. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimmfeld. *Sprache-Stimme-Gehör.* 1985; (9):109–111.
26. Titze, IR. *The Myoelastic Aerodynamic Theory of Phonation.* The National Center for Voice and Speech; 136 So. Main, Salt Lake City, Utah 84101: 2006.

Highlights

- Vocal Efficiency is divided into conversion efficiency, transmission efficiency, and radiation efficiency
- Conversion efficiency is on the order of 10 %
- Efficiency control is by lung pressure and adduction

Speech Commun. Author manuscript; available in PMC 2019 July 01.

Fig. 2.
Detailed configuration of the 5 glottal sections, showing nominal adduction and glottal wall structure. Cross sections are highly elliptical (not to scale).

Fig. 3.
Simulated waveforms for power calculations.

Fig. 4.

Calculations at 7 locations along the airway. (a) steady (DC) pressure, (b) steady (DC) power, (c) total power, and (d) acoustic (AC) power. The tracheal driving pressure was 1.0 kPa and the mean (steady) flow through the airway was 204 cm³/s.

Fig. 5.

Variation with vocal tract wall stiffness. Dashed blue is original and solid red is for a 5-fold increase in wall elasticity and viscosity. (a) steady (DC) pressure, (b) steady (DC) power, (c) total power, and (d) acoustic (AC) power. The tracheal driving pressure was 1.0 kPa and the flow through the airway was $173 \text{ cm}^3/\text{s}$.

Fig. 6.

Variation with air viscosity in the vocal tract. Dashed blue is for 0.000186 poise and solid red is for zero viscosity (a) steady (DC) pressure, (b) steady (DC) power, (c) total power, and (d) acoustic (AC) power. The tracheal driving pressure was 1.0 kPa and the flow through the airway was 205 cm³/s.

Fig. 7.

Variation with greater vocal fold damping ratio. Blue is for the original value 0.3 and red is for 0.4 (a) steady (DC) pressure, (b) steady (DC) power, (c) total power, and (d) acoustic (AC) power. The tracheal driving pressure was 1.0 kPa and the flow through the airway was $180 \text{ cm}^3/\text{s}$.

Fig. 8.

Variation with lung pressure. Dashed blue lines are for the original value 1.0 kPa and solid red lines are for 0.5 kPa and 2.0 kPa. (a) steady (DC) pressure, (b) steady (DC) power, (c) total power, and (d) acoustic (AC) power. The flows through the airway were 800 cm³/s, 204 cm³/s, and 94.6 cm³/s, respectively.

Fig. 9. Variation with vocal fold adduction. Dashed blue lines are for the original diameter values 1.0, 0.8, 0.6, 0.6, 1.0 mm, as in Fig. 2, top solid red lines are for 1.6, 1.4, 1.2, 1.2, 1.6 mm, and bottom solid red lines are for 0.6, 0.4, 0.2, 0.2, 0.6 mm. (a) steady (DC) pressure, (b) steady (DC) power, (c) total power, and (d) acoustic (AC) power. The flows through the airway were 296 cm³/s, 196 cm³/s, and 105 cm³/s, respectively.

Fig. 10. Calculated vocal efficiency from simulation (solid lines) versus measured vocal efficiency by Schutte (1980) on 45 normal individuals as a function of lung pressure and mean flow rate.

Table 1

Power in mW at seven locations along the vocal tract for nominal configuration.

	Loc 1	Loc 2	Loc 3	Loc 4	Loc 5	Loc 6	Loc 7
Total pow	210.3	191.9	37.22	14.86	6.437	1.71	0.042
DC pow	210.2	192.2	61.30	8.78	3.72	0.997	0.008
AC pow	0.066	-0.251	-24.08	6.08	2.71	0.715	0.034

Sound Level and Efficiency calculated from AC power and cross-sectional area for three lung pressures (5, 10, 20 kPa in rows of three) at seven locations along the nominal vocal tract configuration.

Table 2

	Loc 1	Loc 2	Loc 3	Loc 4	Loc 5	Loc 6	Loc 7
Area, cm ²	8.00	2.55	0.1	0.8	1.17	1.54	1.0
AC pow (Mw)	0.026	-0.024	-5.59	0.885	0.440	0.081	0.0022
	0.066	-0.251	-24.08	6.083	2.713	0.715	0.034
	0.277	-0.062	-62.45	29.338	11.307	4.059	0.0842
SL (dB)	105	110	147	130	126	117	103
	109	120	154	139	134	127	115
	115	114	158	146	140	134	119
Efficiency (%)			9.82				0.0039
			11.40				0.016
			7.47				0.010