Published in final edited form as:

J Voice. 2017 November ; 31(6): 649–661. doi:10.1016/j.jvoice.2017.03.001.

Dynamics of the driving force during the normal vocal fold vibration cycle

P.H. DeJonckere¹, J. Lebacqz², and I.R. Titze³

¹Neurosciences, University of Leuven & Federal Agency for Occupational Risks, Brussels, Belgium

²Neurosciences Institute, University of Louvain B-1200 Brussels, Belgium

³National Center for Voice and Speech, University of Utah, Salt Lake City, Utah, USA

Abstract

Intraglottal pressure is the driving force of vocal fold vibration. Theoretically, simultaneous quantification of glottal area and transglottal airflow allows calculating the intraglottal pressure waveform during a single vibration cycle. In this study, we show that, by combining photoglottography (transglottal light transmission) and airflow (Rothenberg's mask) measurements during sustained vocal emissions *in vivo*, the intraglottal pressure wave can be approximated in a way similar to what has been done in models. The results confirm *in vivo* that the intraglottal pressure is systematically larger during the opening phase than during the closing phase, so that over one whole cycle, the driving force performs net positive work, accounting for sustained vocal fold motion. A component of this driving force asymmetry is related to vocal tract inertance, which also accounts for the skewing of the airflow waveform compared to the area waveform. Furthermore, the intraglottal pressure ratio (opening/closing) increases with voicing intensity, reaches a maximum around 76 dB, and significantly decreases at higher intensities. This rise and fall suggests that there is a range of intensity values in which, mechanically, a maximum of the driving force is imparted to the vocal fold mass. This finding could have implications for voice economy in professional speakers.

Keywords

Intraglottal pressure; transglottal flow; glottal area; photoglottography; flow glottography; glottal mechanics

Introduction

Intraglottal pressure produces the driving force of vocal fold vibration¹. Owing to their technical difficulty, direct experimental recordings of the intraglottal pressure (or even subglottal or supraglottal pressure) with respect to glottal dynamics during the vibratory

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.

cycle are very scarce. Miller & Schutte² used simultaneous electroglottography (EGG) and a transglottal pressure transducer to study pressures above and below the glottis. In sustained phonation in a trained male singer, they found a sharp maximum peak in the subglottal pressure signal at, or just after the time of closure of the vocal folds. The peak probably reflects the acoustic wave generated by the collision of the vocal folds, the pressure transducer behaving as a microphone positioned very close to the glottis. Using a similar (Millar) pressure transducer introduced subglottally through the crico-thyroid space, Sundberg & al.³ observed a peak of subglottal pressure that tended to occur just after the time of maximum flow declination rate (MFDR), i.e. nearly at the time of vocal fold impact. However, the shape of the subglottal pressure waveform was found to be clearly affected by the vowel produced. Ursino & al.⁴ suggested that the synchrony of a sharp pressure peak with the time of glottal closure could be the result rather than the cause of the movement of the vocal folds. Indeed, the source of the pressure peak can be compression of the air resulting from the vocal fold collision. Several in vitro and modeling studies have been published, but one older publication⁵ reported a subglottal pressure waveform directly measured *in vivo* (filtered by the narrow Teflon cannula) showing a phase lead of slightly less than 90° relative to the glottal area waveform. Such a phase lead can also be seen in the tracings reported by Alipour & al.⁶ in excised canine larynges.

According to Titze^{7,8}, a positive flow of energy from the airstream to the tissue can be realized only if the net aerodynamic driving force has a component in phase with the tissue velocity (i.e. the first derivative of displacement, thus with a phase lead of 90° with respect to displacement). Whenever the driving force is in the direction of the velocity, energy is imparted to the tissue mass. Conversely, when the driving force is opposite to the direction of the velocity, energy is taken from the vibrating tissue mass.

Fig. 1 (adapted from Titze⁸) shows a simulation of a single vibration cycle of the vocal folds in a typical normal phonation of a male subject (modal register). The figure constitutes our hypothesis for the measurements to be described. The two upper traces represent glottal area and airflow as a function of time. Both signals increase upwards. The glottal area corresponds to the tissue displacement. Inertance of the vocal tract air column—as well as, to some extent, air compressibility - causes the build-up and the collapse of the airflow to be slightly delayed with respect to the opening and closing of the vocal folds. After the work of Rothenberg⁹, the shape of the airflow has been discussed and modeled in detail by several authors: typically there is a slower rise and a more abrupt fall, causing a skewing of the airflow waveform to the right.

The middle trace in Fig. 1 represents the shape of the air particle velocity waveform, obtained by dividing the flow waveform by the glottal area waveform. Only the open part of the vibration cycle is shown and is relevant in the scope of this discussion. Due to the skewing of the airflow curve, the air particle velocity is asymmetric, with an average increase to the right.

The lower trace in Fig. 1 is the waveform of the intraglottal pressure computed on the basis of Bernoulli's energy law, which gives a first approximation of the driving force on the tissue on the basis of conservation of energy in incompressible fluid flow that remains

attached to the glottal wall and has negligible viscous loss. The total energy in the fluid at any point has two components, one directly related to pressure and one related to particle velocity, thus representing the work energy and the kinetic energy. The kinetic energy is proportional to the square of the particle velocity. The energy conservation is written in terms of pressure as:

$$P + 1/2 \rho v^2 = \text{constant} \quad (1)$$

where ρ is fluid density and v is particle velocity. If the energy in the fluid stream is constant, an increase in particle velocity must be accompanied by a decrease in pressure. This energy conservation holds for inviscid laminar flow that remains attached to the glottal wall. Flow separation from the wall and vortex formation in a divergent glottal duct causes a departure from this simple energy conservation law. The pressure does not recover completely in an expanding (diverging) duct, which helps to create the pressure asymmetry required for self-sustained oscillation⁷, provided that the divergence occurs during glottal closing. Also, Bernoulli's law is not valid for compressible flows (variable air density). However, for Mach numbers < 0.3 , it may be assumed that the glottal air flow is incompressible¹⁰. Actually the compression rate of air at the glottal level is limited: for speaking voices (including loud voicing), the volume changes due to air compressibility are in the range of 0.7 to 4 % (subglottal pressures of 7 to 40 hPa).

Hence, as air particle velocity gradually increases during the major part of the open phase of the vibration cycle, according to the total energy formula, intraglottal pressure must decrease over the same time period. This estimate of intraglottal pressure is valid over a fraction of the open period, namely the opening phase. If flow separation from the glottal wall occurs in the closing phase, as well as flow vorticity, the intraglottal pressure is affected by the supraglottal acoustic pressures, which modifies the overall pressure distribution in the glottis.

If flow separation and supraglottal acoustic pressure are included, the intraglottal pressure can theoretically be divided into two parts, one upstream and the other downstream of the flow separation,

$$P = P_s - k (1/2) \rho v^2 \text{ upstream of flow separation} \quad (2)$$

$$P = I dU/dt \text{ downstream of flow separation} \quad (3)$$

In the above, P_s is subglottal pressure (the lung pressure created by the contraction of expiratory muscles and/or elastic thoracic forces), k is a pressure loss coefficient for glottal entry and viscous drag, I is supraglottal acoustic inertance, and U is airflow. However, exactly defining the boundary of the flow separation is practically impossible. Downstream of the flow separation, dU/dt is mainly positive during glottal opening (when flow is

increasing) and mainly negative during glottal closing (when flow is decreasing). Thus, a push-pull alternation is felt on the upper part of the vocal folds, in phase with the tissue velocity. The inertance I can be estimated as

$$I = \rho L / S \quad (4)$$

where S is the mean cross-sectional area of the epilarynx portion and L is the effective length of the air column above the vocal folds. L and S may be considered as constant during emission of a sustained vowel, as is the case in our experiments.

As stated by Titze⁸ (Chapter 4), from a mechanical point of view, the only requirement for keeping the vocal folds in vibratory motion is that the driving force be less positive during closing (including tissue recoil) than during opening, and that the net driving force over the whole cycle be sufficient to overcome frictional forces. The important point is the asymmetry of the pressure curve in the opening portion of the cycle *vs.* the closing portion. The asymmetry can be demonstrated by calculating the rate of change of the pressure in equation (2) which becomes:

$$dP/dt = (\partial P / \partial A) (dA/dt) + (\partial P / \partial U) (dU/dt) \quad (5)$$

$$= -k \rho (U/A^2) (-U/A) dA/dt + (-k \rho U/A^2) dU/dt \quad (6)$$

$$= -k \rho (U/A^2) [dU/dt - (U/A) dA/dt] \quad (7)$$

in which A is glottal area. P_s is no longer in the equation and k is only a scaling constant in the rate of change of pressure. Everything else is dependent on the functions U and A , which are measured over time. The function in brackets can be evaluated point by point in time series from the area and flow measurements. Once dP/dt is integrated, the pressure is known only within a constant, but it can be demonstrated that the pressure is larger when dA/dt is positive (glottal opening) than when dA/dt is negative (glottal closing).

The first aim of the present work is to verify whether these assumptions are actually supported by direct simultaneous physiological measurements of glottal area and flow. Considering that, on the basis of glottal area measurements, it is impossible to precisely define when either equation (2) or equation (3) must be applied, - this consideration being supported by experimental observations - we have applied the general equation (1), as well as equation (7) derived from equation (2) for computing intraglottal pressure on the basis of our measured area and flow data. As regards equation (7), the assumption of a different intraglottal pressure when dA/dt is positive than when dA/dt is negative will be verified. As regards equation (3), I is a constant whereas it may be assumed that dU/dt is mainly positive

during glottal opening (when flow is increasing) and mainly negative during glottal closing (when flow is decreasing).

The second aim is to compute the ratio of the intraglottal pressure during the opening and the closing phases, *i.e.* the areas under the curves of intraglottal pressure *vs* time, thus reflecting the mechanical efficiency at glottal level—in different voicing conditions. Indeed, this intraglottal pressure ratio is likely to be influenced by physical parameters like sound pressure level (SPL) of vocal emission and closed quotient.

Material and Methods

Subject

The subject for this experiment was one of the authors, a trained vocalist with extensive experience in voice production experiments.

Glottal area

The glottal area was derived from a photoglottographic recording, obtained by transilluminating the trachea and detected by a photovoltaic transducer in the pharynx¹¹ (Fig. 2). The light source was a 15 W tungsten filament light bulb driven by a 12 V battery as a constant current source. The lamp was mounted in a cylindrical housing closed by a condenser lens, which was placed on the skin of the neck facing the trachea below the cricoid cartilage.

The transducer, a BP104 silicon photodiode (Vishay Precision Group, Malvern, PA, USA), was glued onto a small laryngoscopic mirror (Nr. 3), the handle of which was introduced - together with the sensor lead - through the hermetically sealed hole normally intended for the handpiece of the Rothenberg mask (see Fig. 2). The current produced by the photodiode was amplified by a home-made current-to-voltage converter with a linear response up to 2 kHz. Saturation of the output signal was observed only at light intensities 10 to 20 times larger than those recorded in the experiments.

A function generator (model 3311A, HPE, Palo Alto, CAL, USA) was used to produce calibrated sine waves added to a constant positive offset. This signal was fed to light up an LED placed in front of the detecting photodiode as used in the records. The aim of the constant base current was to keep the LED lit continuously, to avoid a possible threshold effect of the LED when applying low frequency sine waves. The light emitted by the LED was readily detected by the photodiode and the magnitude of the exciting currents was adjusted to produce photodiode signals of the same order of magnitude as those of signals from the glottis. By displaying both the LED current and the photodiode output on the oscilloscope screen, the exact time delay of the whole chain of photo signal recording was measured over a range of frequencies from 80 to 220 Hz. The result was a systematic mean delay of 0.102 ms with no significant change over the tested range of frequencies, with an effective time resolution of 10 μ s. All records were corrected for the delay.

The peak of the photoglottographic trace clearly marks the time of maximal glottal area. The other critical glottal events (beginning of glottal opening and end of glottal closing) were

defined according to the method of Gerrat & al.¹²: they occur when the rising and falling trace intersects a horizontal line drawn at 90 % down from the positive peak. This line is parallel to a line drawn between the negative peaks preceding and following the positive peak of maximal opening.

Glottal flow

The glottal flow waveform was recorded using a Rothenberg mask and the MSIF2 inverse filtering system of Glottal Enterprises, Syracuse, NY. The mask is equipped with a compressible seal and is firmly pressed against the face of the subject, to avoid any air leakage.

By injecting air from an external source in the mask, saturation of the output of the filtering system was found to occur at flow values 20 % larger than the largest values measured in the experiments.

Rothenberg specifies, in the original papers describing the use of his flowmeter, that a constant correction of 1.25 ms must be applied to all records¹³. The delay is the sum of two terms: delay in the electronic detection and preamplification circuit, and delay due to propagation of the pressure wave from the level of the glottis to the pressure detector inserted in the wall of the mask. The delay due to the electronic circuitry was measured by injecting calibrated signals at the input of the preamplifier, and a constant delay of 0.75 ms was found between the input and the output of the circuit, independent of frequency. In the subject tested in this work, a distance of 16 cm was measured from the glottis to the pressure detector. The speed of sound at sea level, 30°C and 100 % RH is 351 m.s⁻¹, and the delay due to propagation of the pressure wave is thus 0.455 ms. All records were corrected for the total delay.

The flow measuring system of Rothenberg was calibrated as follows: air was introduced in the mask from a controlled source at a range of flow values, and the output of the electronic circuit was recorded; at each value of flow, the absolute value of the flow was measured using a calibrated pneumotachograph connected in series with the mask. The slope of the regression analysis of the results was 85.7ml.s⁻¹/V.

For actual *in vivo* combined measurements, the practical difficulty is to succeed in obtaining (with adequate formant filtering) a flow signal with a closed plateau that is as horizontal as possible (without ripple), while simultaneously configuring the oropharyngeal tract in such a way that the glottal light signal is optimally detected by the photovoltaic sensor. As a matter of fact, the absolute magnitude of the light signal is critically influenced by the precise positioning of the sensor in the pharynx. Nonoptimal positioning results in a low amplitude, noisy signal, although its pattern remains unchanged. In practice, it is unrelated to SPL, contrary to the flow signal. For this reason, in all recordings, light signals have been expressed as percentages of the maximum value of the signal, thus ranging between 0 and 100.

Recording

Both signals were recorded using a PicoScope 2205A module (Pico Technology Ltd, St Neots, England U.K.) and stored in a PC computer for later analysis.

Voice material

About 120 recordings were performed on a sustained vowel (sounding like something between a denasal/a/ and a denasal/ə/, due to the experimental setup) always in modal (chest) register at various intensities ranging from very soft - although clearly voiced - to nearly the maximum possible, taking into account the constraints of the recording conditions. A complete velopharyngeal closure is difficult to permanently control, but any possible nasal air escape is collected by the Rothenberg mask. Apart from the experiment itself, average subglottal pressure values were estimated by the flow interruption method in the same subject and for similar utterances: the values roughly range from 4 to 30 hPa.

Measurements

The following parameters were measured on the photoglottographic trace:

Opening time = time of increasing glottal area: time (ms) from the beginning of the opening to the peak of maximal opening

Closing time = time of decreasing glottal area: time (ms) from the peak of the maximal opening to the end of closing

Open time = opening + closing time

Closed time = period - open time

Closed quotient (CQ) = closed time/period

Open quotient = 1 - closed quotient

Speed quotient = opening time/closing time

Mutatis mutandis, similar computations were achieved on the glottal flow waveform. An important additional parameter of the glottal flow waveform is the maximum flow declination rate (MFDR)⁸, the maximal rate of change of the, i.e. just before closure of the glottis.

The flow and photoglottographic signals were used to calculate the velocity of air particles:

$$v = \text{flow/area of the glottis (measured by the photoglottographic signal)}$$

Pressure values were computed according to equation (1).

In principle, we should use equation (2) as long as the glottal shape is convergent and equation (3) when it becomes divergent, i.e. when intraglottal flow separation occurs, and when the entrained flow can develop into flow separation vortices that possibly affect the pressure distribution within the glottis. This holds only for high subglottal pressures, as at low subglottal pressures, the glottis assumes a minimal diverging angle during closing and the flow separates from the superior surface of the vocal folds at the glottal exit^{14,15,16}. However the photoglottographic area signal does not allow defining the critical moment

between the two phases. The time derivative of the photoglottographic signal always goes smoothly through zero and no discontinuity is ever observed at the peak of the area waveform. Hence we assume that the average driving pressure (bottom to top) is close to the Bernoulli pressure.

The value of constant k in equations (2) and (7) is not accurately known. It varies with glottal configuration and lung pressure. Using a model of human larynx, van den Berg *et al.*¹⁷ found a mean value of 1.37. This point has been thoroughly investigated by Fulcher *et al.*¹⁸. In the computation of pressure curves (from the flow and area curves), k was set equal to 1.37 in all cases. This procedure introduces errors in the computation of values of pressure. Considering the estimated mean glottal diameter and the estimated mean subglottal pressure in comparable conditions¹⁹, and according to the computations of Fulcher *et al.* on a static model, we may assume that k varies—in our experiments - from about 1.39 for soft to medium voicing to about 1.0 for medium to loud voicing. However, in this work, the important point is not the absolute values of intraglottal pressure but rather the ratio of the intraglottal pressure during opening and closing phases respectively, and both of these phases are affected in the same way by the inaccuracy.

The sampling frequency was 391 kHz for each signal. According to the sound intensity level, durations of open phases ranged from about 3 ms at high intensity to about 6 ms at low intensity, and the number of samples of both signals ranged from about 575 to about 1,150: these are the numbers of pressure values used in computing the intraglottal pressure. The intraglottal pressure during the opening phase and during the closing phase were calculated separately by summing values of pressure between the vertical limits set in the area waveform, and the ratios of the intraglottal pressure were calculated. It must be noted that in this procedure, no assumption is made about the absolute pressure.

Sound pressure level

A small condenser microphone (Ø 5.6 mm) was fixed laterally inside the Rothenberg mask (see Fig. 2): it exactly fits into an opening of the mask on the side opposite the pressure transducer. Processing of the voice samples for SPL analysis was achieved using the PRAAT software (www.praat.org). The microphone sound levels were calibrated with a Wärtsilä 7178 (Wärtsilä Electronics, Helsinki, Finland) sound level meter in a position corresponding to a direct measurement at 10 cm from the lips. The exact moment chosen for the area and pressure recording could be clearly identified on the sound recording, allowing a precise fit of the area/airflow curves with the dB (A) value.

Data analysis

Statistica 13 software (Dell Inc., Tulsa, USA) was used for graphics and statistics.

Results

All vocal emissions were in modal register, within the range 58 to 86 dB. Fundamental frequencies were found to range between 108.1 and 155.2 Hz.

Fifty recordings were selected for detailed analysis. The main reason for rejection was a too small photovoltaic signal. This may be due either to nonoptimal positioning and orientation of the sensor, or to inadequate morphological configuration of the vocal tract (position of tongue basis and epiglottis) causing an obstacle between glottis and sensor. Indeed, in the computation procedure, the flow signal is divided by the light signal, and the signal-to-noise ratio of the result is mainly determined by that of the light signal. At the beginning of the opening phase and at the end of the closing phase, both signals are small and their ratio becomes extremely noisy, all the more so when the light signal is so small that the ratio becomes meaningless.

Figs. 3 to 5 show three typical recordings with the juxtaposed glottal area, airflow and the calculated intraglottal pressure curves. Pressure values were computed according to equation (1). During the closed phase, as both variables are near zero, the squared ratio becomes meaningless. For this reason, that part of the curve has not been represented in Figs. 3 to 5. The air velocity curves are not shown, as they mirror (upside down) the intraglottal pressure curves.

Fig. 3 corresponds to a soft voicing at 58.92 dB, with a very short closed plateau. The airflow signal follows a slightly skewed sinusoidal pattern. The very broad peak of the pressure curve just precedes the maximal opening of the glottis, and the intraglottal pressure during the opening phase just exceeds that during the closing phase (ratio near 1).

In Fig. 4, the SPL is 76.84 dB, with a long closed phase. The airflow trace is clearly skewed to the right. A ripple can be seen in the airflow trace just after the closing; it is due to imperfect filtering. The pressure peak occurs slightly less than a quarter of a period before the maximum glottal opening, and the pressure abruptly drops shortly after the maximum. The intraglottal pressure during the opening phase is about 7 times larger than that during the closing phase.

Fig. 5 corresponds to a louder voice utterance (83.78 dB), again with a long closed phase, although not longer than that of Fig. 4, and with ripple just after the closing. The pressure drop occurs clearly later with respect to the maximum glottal opening, causing the pressure ratio to dip compared to the case of Fig. 4.

For the voice utterance considered in Fig. 4, which is, as regards loudness, quite representative of characteristics of professionally speaking in public, a separate application of equations (2/7) is shown.

Fig. 6 shows the glottal area (top) and P (bottom) calculated by integration of eq. 7, derived from eq. 2. The intraglottal pressure abruptly drops after the maximal glottal opening is reached, and becomes clearly smaller during the closing phase.

Fig. 7 shows the pressure ratio (intraglottal pressure during the opening phase/intraglottal pressure during the closing phase according to the glottal area display) as a function of the intensity of voice production (dB). The red curve shows the best fit through the points, using a least squares method. Even though there is a large spread of the points, the ratio clearly increases with voice loudness, from the lowest intensities on up to around 76.5dB. For

intensities higher than around 76.5 dB, the opposite trend is observed. This is confirmed by a statistical analysis: when the 25 ratio-points corresponding to intensities larger than 76.5 dB are considered separately, they correlate negatively with the SPL values ($r = -.44$; $p < 0.05$).

Fig. 8 shows the closed quotient of the airflow wave as a function of intensity. The closed quotient clearly increases when the voice becomes louder, but reaches a plateau at around 76 dB, remaining constant at larger intensities.

In Fig. 9, the skewing quotient of the flow is plotted as a function of intensity: similarly to Fig. 6, there seems to be a maximum skewing around 76–78 dB. Again, when the 25 skewing quotient values corresponding to intensities larger than 76.5 dB are considered separately, they correlate negatively with the SPL values ($r = -.41$; $p < 0.05$).

Figs. 10 shows how the MFDR gets stronger negative values (steeper falling slope) as intensity of voice increases, without a clear peak. The MFDR is actually the maximum value (in mL/ms/ms) of the first derivative of the airflow curve.

As seen in Fig. 11, the peak flow clearly increases with intensity, up till about 76 dB, and does not significantly change at larger intensities.

Discussion

Two main findings result from our computations

First, as to the basic mechanism of the time relation between the driving force and the vocal fold movement, our recordings demonstrate that whatever the SPL, the intraglottal pressure is systematically larger during the opening phase than during the closing phase, so that over one whole cycle, the driving force performs net positive work (Fig. 1). Importantly - according to Titze's model^{7,8} - the general downward trend of the mean intraglottal pressure during the open phase of the cycle matches the downward trend of the tissue velocity, in other words the tissue displacement shows a phase delay of $\pi/2$ radians (in ideal conditions, without friction) with respect to the driving force. This downward trend, which creates the asymmetry in the flow waveform, results from the skewing of the airflow wave with respect to the area wave, and is a key factor in obtaining a velocity-dependent driving force. Quite frequently the intraglottal pressure curve shows a blunt peak (see e.g. Fig. 5): such a pressure peak always precedes the maximal opening by less than $\pi/2$ radians.

This confirms previous results obtained in vivo as well as in excised canine larynges^{5,6}. The nearly always sharp positive pressure peak found by some authors² at the time of closure could be, as already suggested by Ursino et al.⁴, the consequence rather than the cause of the movement of the vocal folds. Jiang & Titze¹, using a hemilarynx procedure (dog hemilarynxes with vocal fold vibrating against a plexiglas plate with included pressure sensors) observed (1) a sharp pressure peak (impact stress) of about 3 kPa when the vocal fold makes contact with the plate, and (2) a large positive intraglottal pressure wave when the pressure sensor has no contact with the vocal fold and the pressure waveform represents subglottal pressure. The maximum of this large positive pressure wave occurs slightly before the maximal glottal opening.

As mentioned above, our pressure computation using equation (1) remains an approximation. On the one hand, we do not separate the convergent and divergent phases of the glottal shape, for which eq. 3 resp. 2/7 are applicable. However, we show (Fig. 6) that application of eq. (7) and eq. (1) provides qualitatively similar results. As regards equation (3) (cf. Introduction), I is a constant whereas dU/dt is mainly positive during glottal opening and mainly negative during glottal closing. On the other hand, we assume that air is incompressible, which is requested for applying Bernoulli's formula, but in our working conditions this compression rate of air at glottal level is limited: 0.7 to 4 %. Further, the maximum intraglottal air speed is 70 m/s²⁰ which remains below Mach 0.3.

Second, it clearly appears that the ratio of the intraglottal pressure during respectively glottal opening and glottal closing strongly varies with the SPL of voice emission, at least within the frequency range of the male speaking voice. The essential finding is the biphasic shape of the pressure ratio curve, when plotted against voicing intensity (Fig. 6). There is a zone of intensity values where the ratio of pressures is maximal, i.e. where mechanically, a maximum of the driving force is imparted to the vocal fold mass. In our experiments, this zone of optimal mechanical efficiency - at glottal level - is around 76.6 dB.

This idea is not new: on the basis of open quotient, skewing quotient and MFDR values, Titze⁸, using computer models, hypothesized that there is a 'Open Quotient' value for getting an optimal voice power (around a 'Open Quotient' of about 0.5 to 0.6). According to Titze, although this could be of pedagogical value, finding the optimal adjustment may not be an easy task for every vocalist.

Tang & Stathopoulos²¹ measured the vocal efficiency (ratio of radiated acoustical power to the supplied aerodynamic power), and found that—in speech tasks - vocal efficiency increased with vocal intensity, but they limited the intensity to 5 +/- 1 dB SPL above the comfortably produced vocal intensity level, which remains inferior to the range investigated in this work. It must also be noted that they estimated the total vocal efficiency, whereas in our work, the efficiency is considered strictly at the glottis level, i.e. only part of the whole system.

The determining factor of the biphasic shape of the pressure ratio curve actually seems to be the changing skewing quotient of the flow waveform, although it has been reported that, in normal male subjects, there are no significant changes in photoglottographic speed quotient with variations of intensity or frequency²². Actually, the skewing quotient curve of the flow waveform as a function of voice intensity closely resembles that of the pressure ratio, with a significant decline beyond about 76 dB (Fig. 8). Dromey et al.²³ also observed, in normal subjects, that the skewing quotient (of the airflow) initially increased with vocal intensity but decreased near the intensity ceiling.

Vocal intensity is highly correlated with MFDR, as has been demonstrated both theoretically^{24,25} and experimentally^{26,27,23}: SPL is directly determined by the maximum steepness of the flow decrease during the closing phase, i.e. the magnitude of the negative peak of the derivative of the airflow^{28,25}. Our data fully confirm this point (Fig. 9). A larger maximum slope has the spectral effect of increasing the energy of the partials within the first

and second formants (/a/-/ə/), actually the most important spectral portion for the overall SPL^{29,30,31}. Conversely, the larger the peak-to-peak amplitude of the glottal flow, the larger is the intensity of the voice source fundamental^{27,28,31}.

There are actually three different ways by which this steepness can be enhanced²⁸:

1. by increasing the amplitude of the flow pulse: this occurs up to about 76 dB, but no more for higher SPL values, as can be seen in Fig. 10;
2. by increasing the closed quotient: Fig. 7 shows this to be the case for intensities up to about 76 dB, but not beyond;
3. by increasing the skewing of the pulses: this possibility is suitable for low and medium intensities, but at higher intensities, the flow pulse becomes significantly less skewed (Fig. 8).

From a certain intensity level on, the closed quotient ceases to increase and remains stable at higher intensity values. The pressure ratio begins to decrease, as does the skewing quotient. Nevertheless, the MFDR continues to get larger negative values (Fig. 9). Comparing Figs. 8, 9 and 10 suggests that, at higher intensities, a change could occur in the pattern of the airflow wave in the sense of a broadening ('squaring') of the shape (see also Fig. 5).

The concept of an optimal energy transfer from driving force to vocal fold motion at a set zone of voice intensity could have relevant implications for voice care in professional speakers (singing voice is of course more complex). Beyond the critical zone, external voice amplification is to be recommended.

Our findings rely upon measurements of one single trained vocalist. There are probably interindividual differences in this critical intensity zone. For example, Howard³² showed that vocally trained subjects make use of higher closed quotients in read passages than untrained subjects.

In further studies, our photoglottographic method could be replaced by high speed transnasal videoendoscopy, the fiberscope being introduced by a small hole in the Rothenberg mask³³.

Finally, it seems also quite clear that a breathy phonation, produced by leaving a persistent opening in the dorsal cartilaginous glottis (insufficient adduction), with presence of a non-zero minimum flow and reduction in MFDR³⁴, will probably recreate the flow and area patterns of Fig. 1, involving a significant reduction in the energy transfer from the fluid to the vocal folds.

Conclusion

Simultaneous measurements of glottal area and transglottal flow allow to compute the intraglottal pressure waveform during a single vibration cycle, and to plot it against the glottal area for different voicing intensities. Whatever the SPL, the intraglottal pressure is systematically larger during the opening phase of the glottis than during the closing phase, so that over one whole cycle, the driving force performs a net positive work, accounting for the vocal fold motion.

There is a zone of intensity values where the ratio of pressures is maximal, i.e. where mechanically, a maximum of the driving force is imparted to the vocal fold mass. In our experiments, this zone of optimal mechanical efficiency—at glottal level - is around 76,6 dB.

The concept of an optimal energy transfer from driving force to vocal fold motion at a set zone of voice intensity could have relevant implications for voice hygiene in professional speakers. Beyond the critical zone, external voice amplification is to be recommended.

Acknowledgments

This study was, in part, supported by grant number 5R01DC012045-05 from the National Institutes of Health.

References

1. Jiang JJ, Titze IR. Measurement of vocal fold intraglottal pressure and impact stress. *J Voice*. 1994; 8:132–144. [PubMed: 8061769]
2. Miller DG, Schutte HK. Characteristic patterns of sub- and supraglottal pressure variations within the glottal cycle. *Transcripts of the XIIIth Symposium Care of the Professional Voice (Philadelphia) Part 1*. 1985:70–75. Also in *Proceedings PAS-Conference October 3–5 Groningen*, H.K. Schutte, ed.
3. Sundberg J, Scherer R, Hess M, Müller F, Granqvist S. Subglottal pressure oscillations accompanying phonation. *J Voice*. 2013; 27:411–421. [PubMed: 23809566]
4. Ursino F, Pardini L, Panattoni F, Matteucci F, Grosjacques M. A study of EGG and simultaneous subglottal pressure signals. *Folia Phoniatr*. 1991; 43:220–225.
5. Dejonckere P, Lebacqz J. Phase relationship between dynamics of the subglottic pressure and oscillatory movement of the vocal folds. I. Sustained phonation. *Archives of Physiology and Biochemistry*. 1980; 88:333–341.
6. Alipour F, Scherer RC, Finnegan E. Pressure-flow relationships during phonation as a function of adduction. *J Voice*. 1997; 11:187–194. [PubMed: 9181542]
7. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am*. 1988; 83:1536–1552. [PubMed: 3372869]
8. Titze I *Principles of voice production*. National Center for Voice and Speech; Iowa City IA USA: 2000. 2nd Printing
9. Rothenberg M. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *J Acoust Soc Am*. 1973; 53:1632–1645. [PubMed: 4719255]
10. McGraw-Hill Concise Encyclopedia of Physics. © 2002. The McGraw-Hill Companies, Inc; New York, USA:
11. Dejonckere PH. Comparison of two methods of photoglottography in relation to electroglottography. *Folia Phoniatr (Basel)*. 1981; 33(6):338–47. [PubMed: 7343452]
12. Gerratt BR, Hanson DG, Berke GS, Precoda K. Photoglottography: A clinical synopsis. *J Voice*. 1991; 5:98–105.
13. Rothenberg, M. Source-tract acoustic interaction in breathy voice. In: Titze, IR., Scherer, RC., editors. *Vocal fold physiology: biomechanics, acoustics and phonatory control*. The Denver Center for the Performing Arts; Denver, CO: 1984. p. 465–481.
14. Oren L, Khosla S, Gutmark E. Intraglottal geometry and velocity measurements in canine larynges. *J Acoust Soc Am*. 2014; 135:380–388. [PubMed: 24437778]
15. Khosla S, Oren L, Ying J, Gutmark E. Direct simultaneous measurement of intraglottal geometry and velocity fields in excised larynges. *Laryngoscope*. 2014; 124(Suppl 2):S1–13.
16. Oren L, Gutmark E, Khosla S. Intraglottal velocity and pressure measurements in a hemilarynx model. *J Acoust Soc Am*. 2015; 137(2):935–943. [PubMed: 25698025]
17. Van den Berg, Jw, Zantema, J., Doornenbal, P, Jr. On the air resistance and the Bernoulli effect of the human larynx. *J Acoust Soc Am*. 1957; 29:626–631.

18. Fulcher L, Scherer RC, Anderson N. Entrance loss coefficients and exit coefficients for a physical model of the glottis with convergent angles. *J Acoust Soc Am*. 2014; 136/3:1312–1319.
19. Dejonckere PH, Lebacq J, Bocchi L, Orlandi S, Manfredi C. Automated tracking of quantitative parameters from single line scanning of vocal folds: A case study of the ‘messa di voce’ exercise. *Logopedics Phoniatrics Vocology*. 2015; 40:44–54.
20. Berke, GS., Gerratt, BR. Biomechanics of vocal fold motion revisited. *Proceedings of the 10th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research*; June 3–4, 2013; Cincinnati, Ohio, USA. p. 3-4.
21. Tang J, Stathopoulos ET. Vocal efficiency as a function of vocal intensity: A study of children, women, and men. *J Acoust Soc Am*. 1995; 97:1885–1892. [PubMed: 7699170]
22. Hanson D, Gerratt BR, Berke G. Frequency, intensity and target matching effects on photoglottographic measures of open quotient and speed quotient. *J Speech Hear Res*. 1990; 33:45–50. [PubMed: 2314084]
23. Dromey C, Stathopoulos ET, Sapienza CM. Glottal airflow and electroglottographic measures of vocal function at multiple intensities. *J Voice*. 1992; 6:44–54.
24. Fant, G. STL-Quarterly Progress and Status Report. Vol. 23. Dept for Speech, Music and Hearing KTH; Stockholm: 1982. Preliminaries to analysis of the human voice source; p. 1-27.
25. Titze IR. Theoretical analysis of maximum flow declination rate versus maximum area declination rate. *J Speech Language and Hearing Research*. 2006; 49:439–447.
26. 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]
27. Gauffin J, Sundberg J. Spectral correlates of glottal voice source waveform characteristics. *J Speech Hear Res*. 1989; 32:556–565. [PubMed: 2779199]
28. Sundberg, J. STL- Quarterly Progress and Status Report. Vol. 35. Dept for Speech, Music and Hearing KTH; Stockholm: 1994. Vocal fold vibration patterns and phonatory modes; p. 69-80.
29. Gobl, C., Karlsson, I. Male and female voice source dynamics. In: Gauffin, J., Hammarberg, B., editors. *Vocal Fold Physiology: Acoustic, perceptual and physiological aspects of voice mechanisms*. San Diego CA: Singular Publishing Group; 1991. p. 121-128.
30. Scherer, R. Laryngeal function during phonation. In *Vocal health and pedagogy*. In: Sataloff, RT., editor. *Science and assessment*. 2. Plural Publishing; San Diego: 2006. p. 85-102.
31. Scherer, R. Laryngeal function during phonation. In: Rubin, JS, Sataloff, RT., Korovin, GS., editors. *Diagnosis and Treatment of Voice Disorders*. Plural Publishing Inc; San Diego CA: 2014.
32. Howard DM. Toward the quantification of vocal efficiency. *J Voice*. 1990; 4:205–212.
33. Mehta, DD., Deliyiski, DD., Zeitels, SM., Zañartu, M., Hillman, RE. Integration of transnasal fiberoptic high-speed videoendoscopy with time-synchronized recordings of vocal function. In: Izdebski, K, Yan, Y, Ward, RR, Wong, B, JF., Cruz, RM., editors. “Normal & Abnormal Vocal Folds Kinematics: High Speed Digital Phonoscopy (HSDP), Optical Coherence Tomography (OCT) & Narrow Band Imaging (NBI®)”, Volume I: Technology. San Francisco, CA: Pacific Voice & Speech Foundation; 2015. p. 105-114.
34. Zañartu M, Galindo GE. Modelling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction. *J Acoust Soc Am*. 2014; 136:3262–3271. [PubMed: 25480072]

Fig. 1.

(adapted from Titze) Graphic simulation of a single vibration cycle of the vocal folds in a typical normal phonation of a male subject (modal register). The two upper traces represent the glottal area and the airflow as a function of time. Both signals increase upwards. The middle trace represents the shape of the air particle velocity waveform, obtained by dividing the (skewed) flow waveform by the displacement waveform. Only the open part of the vibration cycle is shown. The lower trace is the waveform of the intraglottal pressure, computed on the basis of Bernoulli's energy law.

Fig. 2.

Schematic configuration of the experimental setup, showing the positioning of the subglottal light source, the supraglottal photovoltaic sensor glued onto a small laryngoscopic mirror, the Rothenberg flow mask with its compressible facial seal and its differential pressure transducer, the condenser microphone fitting into an opening of the mask, and the handle of the light sensor hermetically sealed in the hole of the removed handpiece.

Fig. 3.

Typical recording with the juxtaposed glottal area, airflow and calculated intraglottal pressure curves, for a soft voicing (/a/) at 58.92 dB. Closed plateau is very short. Airflow is a slightly skewed sinus pattern. The very broad peak of the pressure curve just precedes the maximal opening of the glottis, and the intraglottal pressure during the opening phase just exceeds that during the closing phase (ratio near 1).

Fig. 4.

Typical recording with the juxtaposed glottal area, airflow and calculated intraglottal pressure curves, for voicing at 76.84 dB (/a/). Closed phase is long. Airflow trace is clearly skewed to the right. The pressure peak occurs slightly less than a quarter of a period before the maximum glottal opening, and the pressure drops abruptly shortly after that maximum. The intraglottal pressure the opening phase is about 7 times larger than that during the closing phase.

Fig. 5.

Typical recording with the juxtaposed glottal area, airflow and calculated intraglottal pressure curves, for loud voicing at 83.78 dB (/a/). Closed phase is long, but not longer than in Fig. 3. The pressure drop occurs clearly later with respect to the maximum glottal opening, causing the pressure ratio to dip compared to the case of Fig. 3.

Fig. 6.

Glottal area (top) and, after integration, P calculated following eq. 7, derived from eq. 2 (bottom) for the same data as in Fig. 4. The area trace starts and stops when the rising and falling trace intersects a horizontal line drawn at 90 % down from the positive peak. The vertical line corresponds to the moment when the area reaches its maximum. The intraglottal pressure abruptly drops after the maximal glottal opening is reached, and becomes clearly smaller during the closing phase.

Fig. 7.

Pressure ratio (intraglottal pressure during the opening phase/intraglottal pressure during the closing phase) as a function of the intensity of voicing. The red curve shows the best fit through the points, using a least squares method. There is a maximum around 76.5 dB.

Fig. 8. Closed quotient of the airflow wave as a function of voicing intensity. The closed quotient clearly increases when the voice becomes louder, but reaches a plateau at around 76 dB, remaining constant at larger intensities.

Fig. 9. Skewing quotient of the flow wave as a function of voicing intensity: similarly to Fig. 5, there seems to be a maximum skewing around 76–78 dB.

Fig. 10. Maximum flow declination rate (= magnitude in mL/ms/ms) of the first derivative of the airflow curve) as a function of voicing intensity. MFDR gets stronger negative values as intensity of voice increases. The line represents the result of a linear regression analysis.

Fig. 11. Peak flow (amplitude of flow wave) as a function of voicing intensity: It clearly increases with intensity, up till about 76 dB, and does not significantly change at larger intensities.