Published in final edited form as:

J Voice. 2016 July ; 30(4): 398–406. doi:10.1016/j.jvoice.2015.06.002.

AN ORAL PRESSURE CONVERSION RATIO AS A PREDICTOR OF VOCAL EFFICIENCY

Ingo R. Titze^{1,2}, Lynn Maxfield¹, and Anil Palaparthi¹

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

Voice production is an inefficient process in terms of energy expended versus acoustic energy produced. A traditional efficiency measure, glottal efficiency, relates acoustic power radiated from the mouth to aerodynamic power produced in the trachea. This efficiency ranges between 0.0001 % and 1.0 %. It involves lung pressure, and hence would appear to be a useful effort measure for a given acoustic output. Difficulty in the combined measurement of lung pressure and tracheal airflow, however, has impeded clinical application of glottal efficiency. This paper utilizes the large data base from Schutte (1980) and a few new measurements to validate a pressure conversion ratio (PCR) as a substitute for glottal efficiency. PCR has the potential for wide application due to low cost and ease of use in clinics and vocal studios.

Keywords

AC/DC ratio; vocal efficiency; oral pressure; vocal effort

INTRODUCTION

Phonation 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 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 speech science. Efficiency is usually defined as a ratio of useful energy output to required energy input. A daily energy intake from food is 8.7 million Joules (J) for a human adult (www.mydailyintake.net), or about 2000 kcal. The rate of consumption, or the power input, is about 100 watts (8.7×10^6 J per 86,000 s in a day). On the output side, acoustic power in speech ranges roughly between 0.01 – 1.0 mW. This power range is derived from a sound

Corresponding Author: Ingo R. Titze, 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.

intensity level (SIL) range of 70 – 90 dB @ 30 cm from the mouth.¹ An output/input ratio yields a global efficiency of the human body for sound production on the order of 0.0001%. Such a global efficiency has little practical use because it involves too many unrelated physical processes in the body.

Efficiency calculations are more useful if localized to an organ, if not to a subcomponent of an organ. The traditional glottal efficiency measure^{2,3} is calculated as the ratio of oral radiated acoustic power to aerodynamic power in the trachea. It has great theoretical appeal because it relates an acoustic output to an effort input (lung pressure, or more precisely, alveolar pressure). Effort in speaking and singing is a clinical and pedagogical issue. Unfortunately, glottal efficiency has seen limited clinical application. Two reasons are: (1) aerodynamic power is difficult to measure directly, and (2) the voice quality labeled “pressed voice” appears to have a high glottal efficiency because of its low airflow (and hence low aerodynamic power), yet clinicians warn against its use because of potential tissue damage.⁴ The difficulty with aerodynamic power measurement has led researchers toward indirect estimation of alveolar pressure and tracheal flow from oral pressure and flow.^{5,6,7,8,9,10} The pressed voice issue has led researchers to search for a vocal economy measure^{11,12} that maximizes acoustic output but minimizes vocal fold collision and energy dissipation in the tissues.

The economy measure proposed by Berry et al.¹¹ is an output-to-cost ratio (in dB), namely the acoustic pressure at the mouth divided by vocal fold contact stress. To make the ratio dimensionless, two reference values were selected, the usual 20 μ Pa for SPL at the mouth and 1.0 kPa for typical contact stress.¹³ This gave output/cost ratios in the range of 55 – 70 dB. It was shown computationally that the ratio varied with the adductory glottal width, 0.5 mm being an optimal value. While the ratio of acoustic output to contact stress is conceptually very appealing, its measurement is not yet clinically practical because contact stress between the vocal folds is difficult to measure.^{14,15} Results in the Berry et al.¹¹ study were based on computer simulation only, in which contact stress is an easy calculation.

More recently, Titze¹⁶ and Titze and Laukkanen¹⁷ proposed an MFDR/MADR ratio for vocal economy, where MFDR is the maximum flow declination rate at the glottis and MADR is the maximum area declination rate of the glottis. The rationale for this ratio is that MFDR is closely related to vocal intensity,^{18,19,20} whereas MADR is closely related to tissue velocity (and hence to momentum change and impact stress) prior to collision. This economy ratio is increased by raising MFDR or lowering MADR, or both. Recent advancements in high-speed kymographic imaging of vocal fold vibration have brought about the potential for obtaining this vocal economy measure on a live individual. The combined measurement techniques were demonstrated by Granqvist et al.²¹ A combination of inverse-filtering the oral flow to obtain glottal flow and its derivative, and high-speed kymographic analysis to obtain glottal area, may lead to a feasible clinical procedure. However, the cost of high-speed imaging equipment will likely remain a barrier to widespread clinical adoption of an MFDR/MADR ratio as a measure of vocal economy.

The purpose of the current study is to explore a third option, an aerodynamic to acoustic pressure conversion ratio, similar to a flow conversion ratio originally proposed by Isshiki.²²

Isshiki called his acoustic flow to steady flow ratio at the mouth an efficiency index, but the use of the word “efficiency” could be challenged because the ratio is not energy or power based. Obtaining a flow or pressure ratio at the lips is an easy measurement that, if validated, could be applied cost-effectively in a clinic with a modest advancement in technology. Given that oral pressure is easier to measure than oral flow, we adopt an aerodynamic-to-acoustic *pressure conversion ratio* (PCR). The pressure measurement is made behind the lips during phonation with a controlled small lip opening. If PCR proves to be of clinical use, questions of first interest are: (1) how does the PCR measure relate to the classical vocal efficiency ratio, (2) is there a strong correlation between high PCR and “pressed voice” as described above, and (3) does PCR vary enough across individuals to have discriminatory potential?

THEORETICAL UNDERPINNINGS FOR GLOTTAL EFFICIENCY AND PRESSURE CONVERSION RATIO

As stated, glottal efficiency has traditionally been defined as the radiated acoustic power from the mouth divided by the pulmonary aerodynamic power delivered by the lungs.^{2,3} In terms of sound pressure level (*SPL*) measured at a distance r from the mouth, glottal efficiency can be written as¹

$$E = \frac{4\pi r^2 I_0 10^{SPL/10}}{P_L U_g}, \quad (1)$$

where I_0 is the standard reference intensity (10^{-12} W/m²), P_L is the lung pressure (the term used here to represent alveolar pressure) and U_g is the mean (slow moving) airflow in the trachea. Since power is the rate of energy produced or absorbed, and since energy conservation and dissipation principles apply throughout the airway, it is guaranteed that the vocal efficiency ratio always ranges between 0.0 and 1.0. This is very satisfying from a physical standpoint, giving an exact accounting of “useful” versus “wasted” energy in the vocal system.

As mentioned briefly in the Introduction, there are several reasons why glottal efficiency has not reached wide-spread application as a vocal effort measure. First, the aerodynamic energy asymptotes to zero when the open quotient (duty ratio) in the glottis approaches zero ($U_g \rightarrow 0$). This could prevent the vocal efficiency from having a maximum value in an intermediate range of adduction, generally thought to be healthy and efficient. Second, the radiated power from the mouth is highly dependent on mouth opening, suggesting that vocal efficiency may change with vowel. A standardized mouth configuration has not yet been adopted. Third, measurement of an airborne acoustic signal requires exact specification of the mouth-microphone distance and some guarantee of room acoustic fidelity to avoid contamination from environmental noise and sound reflections. Fourth, a direct measure of lung pressure is invasive and difficult to obtain. A shuttering technique with the syllable repetition /pa-pa-pa.../ or /pæ- pæ- pæ.../ is generally used for indirect P_L measurement,^{5,6} with the assumption that oral pressure equals lung pressure in the /p/ occlusion. Airflow is measured with a

pneumotachograph flowhead or a circumferentially-vented mask placed over the mouth and nose.²³

Given the measurement challenges facing widespread adoption of the traditional measure of efficiency, we have proposed an alternative measure²⁴ using a dual cannula oral manometer. The procedure maintains a constant small lip opening to develop both a steady pressure and a time-varying (acoustic) pressure behind the lips (Fig. 1). Any measure with expected consistency benefits from standardization of at least one critical vocal tract dimension. A standard lip opening is the easiest to control. The open cannula is chosen to be short (extending the vocal tract by approximately 1–2 cm) with the tongue tip behind the lips likely being in an /o/ or an /u/ position. Subjects are asked to keep the rest of the vocal tract in a neutral position, as in /ə/. The pressure conversion ratio (PCR) is defined simply as

$$PCR = \frac{P_{ac}}{P_{dc}}, \quad (2)$$

where P_{ac} is the oral acoustic pressure (behind the lips) and P_{dc} is the oral steady pressure (also behind the lips). These pressures change with mouth and tube diameter, as discussed later, but the ratio is less geometry-dependent. With this measure, three practical shortcomings are removed. First, no airborne signals are measured, removing the need to specify mouth-microphone distance and reducing environmental interference. Second, the measurement is not dependent on the subject's selection of lip opening because a known orifice is created at the lips. This orifice, on the order of 0.6 cm diameter, is a reasonable compromise between voiced consonants and lip-rounded vowels. Third, the controlled lip opening produces a steady (DC) pressure on the same order of magnitude as the acoustic (AC) pressure, minimizing measurement error for a ratio calculation.

The major disadvantage of the PCR ratio is that it is not energy-based. Hence, it is not bounded by a 0 – 100% range. Pressure ratios anywhere along the vocal tract can increase or decrease by cross-sectional area changes (wide or narrow diameter) without corresponding power changes. Power ratios involve a pressure-flow product. Usually acoustic flow increases when acoustic pressure decreases, and vice versa. Since flow is not measured here, validation of the pressure ratio against a true efficiency calculation is essential.

The validation begins with a simple acoustic pressure transformation in a short cannula at the lips (Figure 2). If the lip tube is acoustically short (much less than a quarter wavelength at a given frequency), then there is no significant pressure drop inside the tube, and acoustic flow continuity at the mouth-tube junction allows P_{ac} to be written in terms of the radiated pressure P_o at the end of the tube,

$$P_{ac} = (A_t/A_m)P_o, \quad (3)$$

where A_t/A_m is the tube-mouth area ratio. Uncertainty of the area A_m is the weakest link in this measure. Some simulation will be needed in the future to determine the sensitivity of the

PCR ratio to this area and the entire vocal tract configuration. If the sensitivity is high, calibration may be needed on a per-subject basis. As an alternative, the uncertainty can be overcome by using a mouthpiece that houses the dual cannulas and maintains a fixed mouth area A_m for any subject. A preliminary design of the mouthpiece is based on a snorkeling mouthpiece, shown in Figure 3. It's adaptability to the *PCR* measure will be assessed in future studies. Here, the intent was to show feasibility and the relation of *PCR* to efficiency.

Returning to Eqn. (3), P_o can be calculated from the measured sound power and the radiation resistance. Given that the radiated output power is P_o^2/R_r , where R_r is the radiation resistance at the tube end, the efficiency is the ratio of the radiated power and the pulmonary power according to Eqn. (1),

$$E = \left(\frac{A_m}{A_t} \right)^2 \frac{P_{ac}^2/R_r}{P_L U_{dc}}. \quad (4)$$

Solving for P_{ac} and substituting into Equation 2, *PCR* can be related directly to efficiency,

$$PCR = \frac{A_t}{A_m} \sqrt{\frac{R_r P_L}{R_{dc} P_{dc}}} \sqrt{E}, \quad (5)$$

where R_{dc} is the steady-flow resistance of the short tube,

$$R_{dc} = P_{dc}/U_{dc}. \quad (6)$$

Thus, *PCR* is proportional to the square root of E , with a proportionality factor that depends on tube and mouth dimensions. One validation issue is whether this factor is a relatively simple scale factor or whether it is subject-specific and phonation-specific. If it is subject-specific but not phonation-specific, a calibration procedure could resolve the difference. However, if the factor varies with every phonation condition (pitch, loudness, adduction, tongue position, etc.), then E and *PCR* become rather different measures, not necessarily highly correlated. This issue needs to be resolved in a study beyond the current one.

PCR can also be expressed in terms of sound pressure level (*SPL*). The radiated power (numerator in Equation (4)) is equated to the radiated power from Equation (1)

$$(A_m/A_t)^2 P_{ac}^2/R_r = 4\pi r^2 I_o 10^{SPL/10}. \quad (7)$$

Solving for P_{ac} and dividing by P_{dc} we get

$$PCR = (4/3\pi)(rd_t/d_m^2)(\sqrt{128\rho c I_o/P_{dc}})10^{SPL/20}, \quad (8)$$

where d_t is the tube diameter, d_m is the mouth diameter ($2\sqrt{A_m/\pi}$), I_o is the ISO standard reference intensity (10^{-12} W/m² or 10^{-19} erg/s –cm²), and an explicit equation for radiation resistance has been used (Flanagan 1965, p. 33)

$$R_r = (128\rho c)/(9\pi^2 A_t). \quad (9)$$

With these relations, *PCR* will now be estimated from Schutte's 1980 data.

POST-HOC ANALYSIS OF SCHUTTE'S DATA

A landmark study on glottal efficiency was conducted by Schutte³ on 45 subjects with no voice pathology and 64 subjects with various voice 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 flowhead (≈ 3 cm diameter and 15 cm long tube) held between the lips during phonation. Figure 4 summarizes the results for both normal and disordered subjects' data. Glottal efficiency is plotted against three variables: *SPL* in the upper left, mean flow in the upper right, and lung pressure in the lower left. The individual small data points and plus signs (+) correspond to the middle of the intensity or pressure range for each subject as given by Schutte in a table at the end of his monograph. The lines with data points give the complete range as reported by regression lines for several runs on each subject. We averaged the regression lines to compute points for efficiency versus *SPL* and efficiency versus lung pressure. The new points are shown with star signs (*) for disordered subjects and plus signs (+) for normal subjects. Curve fits to the new points are given in the legends of Fig. 4.

Schutte's data did not produce much of a trend for efficiency versus mean flow rate (upper right plot). Hence, only the raw data points corresponding to mean intensity for each subject are shown. It is noteworthy that the data clustered around a mean flow rate of about 0.15 – 0.4 l/s, which is confirmed by a histogram in the lower right. Although the highest values of efficiency were for a few outliers (flow rate less than 0.1 l/s), the distribution is near Gaussian, centered on a moderate flow of 0.25 l/s. This result, not specifically emphasized by Schutte, is particularly important because it diminishes the notion that subjects lean toward pressed voice in their productions. There is only a weak correlation between increase in efficiency and decrease in mean glottal flow, which alleviates one concern with the traditional measure of efficiency. Perhaps efficiency is not the physiologic criterion for maximizing output. Transmission-line theory predicts that, for maximum power transfer from a source to a radiation load, the source impedance and the load impedance should be matched, resulting in an efficiency 50% rather than 100%.²⁵ In phonation, the efficiency is so low and raw power is so abundant that it is unclear whether speakers intend to maximize efficiency. As seen in Fig. 4, efficiency clusters around 0.01%, with values as high as 0.3% reached in Schutte's data. Bouhuys et al.² reported a values as high as 2.0% for a singer. The generally low efficiency in voice production, and the wide range (two orders of magnitude) across speakers with different conditions, make the topic of optimization of power output and efficiency worthy of further exploration in terms of internal energy losses.

The post-hoc analysis of Schutte's patient data to determine the discriminatory potential of efficiency between normal and disordered subjects showed three trends. First, on average, the 64 patients had slightly lower efficiency than the normal subjects. At 80 dB *SPL*, efficiency was 0.01 % for normals and 0.004 % for patients. At 90 dB, efficiency was 0.04 % for normal and 0.02 % for patients. Second, none of the 64 patients ever reached 100 dB *SPL*, whereas almost all of the normals did. Third, at low intensity, efficiency was much more variable in patients. At 70 dB, for example, midpoint values from the regression lines for patients (dots) varied between 0.0001 % efficiency and 0.002 % efficiency (more than an order of magnitude), while normal subjects showed little variation. This low intensity difference in variability is particularly interesting in light of phonation threshold pressure, a measure often used to assess "ease of phonation." Schutte did not measure phonation threshold pressure, but data in his monograph indicate that, at 60 dB, normals had a range of 0.2 – 0.6 kPa pressure variation (not shown here), with a mean of 0.4 kPa. Patients, on the other hand, had a range of 0.25 – 4.0 kPa, with a mean of 0.6 kPa. It appears that, for clinical assessment, low *SPL* values should be exploited to separate the groups. At mid-*SPL* values, patients were able to produce sound with only a factor of 2 less efficiency than normals, using on the order of 0.8 – 1.0 kPa of lung pressure.

Statistical analysis was also used to determine if there was a significant difference between normal and disordered subjects' data. One-way analysis of variance (ANOVA) followed by post-hoc Tukey-Kramer test with $\alpha = 0.05$ was conducted on efficiency, lung pressure, glottal flow and *SPL*. The data were normalized to their respective maximum values so that they all range between 0 and 1 before running these tests. The results suggest that the efficiency was significantly different between the two sets of data, $F = 27.3$, $P < 0.0001$. Mean glottal flow was also significantly different between normals and patients, $F = 148.6$, $P < 0.001$. There was no significant difference in lung pressures between normals and patients, $F = 0.21$, $P = 0.64$, and no significant difference between the sound pressure levels produced, $F = 0.72$, $P = 0.39$. These results suggest that for the same lung pressure, efficiency is lower in disordered subjects and more glottal flow is needed to maintain similar sound pressure levels.

The next step in the post-hoc analysis of Schutte's data was to convert his efficiency data to the new *PCR* measure. The following constants were gleaned from his publication: $d_t = 3$ cm, $A_t/A_m = 1$, $R_t/R_{dc} = 25$, $P_L/P_{dc} = 100$, giving a proportionality factor 50 between *PCR* and $\sqrt{E}$ in Equation (5). The P_L/P_{dc} value was estimated by mimicking his experiment with a tube similar to his in diameter. For simplicity, area of the mouth was assumed to be equal to tube area. As mentioned above, in future studies, the mouth/tube area ratio will be controlled more precisely with a mouthpiece. The solid line in Figure 5(a) represents curve-fitted data obtained from mean *E* - P_L regression lines (reported per normal and per disordered subject in Schutte's data). The regression data were substituted directly into Equation (5). There is no distinction between normal and disordered subjects because the same P_L/P_{dc} ratio was used for both (no differences were reported by Schutte). The dashed lines represent curve-fitted data obtained from mean *E* - *SPL* regression lines per subject in Schutte's data. The regression data were substituted directly into Equation (8), which does not contain the P_L/P_{dc} ratio. Hence, normal and patient lines are distinct. Note that *PCR* is on the order of

1.0 when glottal efficiency is on the order of 0.05 %. This relation can only be claimed for Schutte's data at this point.

Figure 5(b) shows plots of PCR versus SPL . Solid lines are from predicted SPL and dashed lines are from measured E - SPL regressions for normal and disordered subjects. The predicted SPL comes from measured E - P_L regressions by considering constant flow across the vocal tract and first calculating aerodynamic power p_a

$$p_a = P_L U_{dc}. \quad (10)$$

The radiated power is

$$p_r = E p_a, \quad (11)$$

and the radiated intensity at a distance r is

$$I = \frac{p_a}{4\pi r^2} \quad (12)$$

from which the sound pressure level is

$$SPL = 10 \log_{10}(I/I_0). \quad (13)$$

This way E was eliminated from the equations. Note that 90 dB SPL yielded a PCR value of about 1.0, with normal subjects having values about 20 % higher than patients with disorders.

With analytic calculation of PCR from Schutte's data, we have shown that there is a predictable relation between PCR and E . This relation was further established with direct PCR measurements on a few new human subjects.

METHODS

Six males and seven females participated in this study. All subjects reported normal, healthy speaking voices at the time of their participation, though one female subject had a history of surgical removal of a laryngeal polyp 1 year prior to participation in the study.

Subjects were first fitted with an electroglottograph (EGG) collar (Kay 6103) and a head-mounted microphone (Countryman Isomax) positioned 6 cm away from and lateral to the mouth. The microphone was calibrated to predict SPL by comparing its output to a sound level meter (C scale) reading at a 30 cm distance. SPL was used to compute PCR from Equation (8). EGG was used to calculate the open quotient Q_o , which is not a variable in the equations but serves as a comparison to Schutte's mean airflow data. Additionally, a thin

tube connected to a pressure transducer (Glottal Enterprises PT-25/MS-110) was maintained between the lips in the oral cavity by the subject. Signals from all devices were captured using an AD Instruments analog-to-digital converter and the company's proprietary software, Labchart7.

Once comfortable with the EGG, microphone, and oral pressure tube, subjects were given a short straw to hold between the lips, 6 mm in diameter and 3 cm in length. While the assumptions made to obtain Equation (5) were based on a 1 cm long tube, it was difficult for subjects to keep a 1 cm straw in place. Consequently, a longer tube was required at this stage. (The mouthpiece to be constructed in the future for a dual-tube system will allow a shorter tube). Subjects phonated through the straw at 3 loudness levels on 3 fundamental frequencies (f_0) within normal speaking ranges. Subjects were provided with stimulus frequencies from a piano and asked to match the pitch with their phonations. Stimulus frequencies ranged from 195 Hz to 247 Hz for females and from 125 Hz to 200 Hz for males. During each 5 – 15 s phonation length, subjects used their index finger to momentarily stop airflow through the straw, mimicking the shuttering technique used to estimate lung pressure.¹⁰ The straw was shuttered regularly throughout each phonation, at a rate of approximately 1.5 interruptions per second. We analyzed 5 – 6 s per subject, or 4 successive occlusions.

Lung pressure (P_L), sound pressure level (SPL), open quotient (Q_o) and pressure conversion ratio (PCR) were derived from the recordings for further analysis. A sliding window of 50 ms with a step size of 25 ms was used to measure these quantities. SPL , Q_o and PCR measured in each window were averaged across each straw-open segment in the shuttering regions. The pressure peak value in the straw-occluded segment was considered an estimate of the lung pressure. Figure 6 shows the DC oral pressure in the straw-open region and the straw-occluded region.

The signal from the pressure transducer provided both steady and acoustic intraoral pressure. Open quotient (Q_o) was derived from each token of the EGG signal for voiced segments. Cepstral peak prominence (CPP)²⁶ was first applied to the EGG signal to identify voiced regions. The open quotient was then measured as a ratio of open phase duration to the glottal cycle duration, using a mean value as the threshold for open-closed decision. This mean value as a threshold overestimates Q_o somewhat, but given that the EGG waveforms did not have much asymmetry across the horizontal axis, the threshold seemed reasonable for an error free calculation. Lower thresholds tend to yield more error due to EGG irregularity. While the PCR measure did not include collecting mean airflow for comparison to Schutte's³ data, it has been shown that open quotient is generally correlated with mean flow.¹⁸ Thus, by using an estimate of Q_o , this study provides a meaningful comparison to Schutte's results on efficiency versus mean glottal airflow.

RESULTS

Figure 7(a) illustrates a square root fit between PCR and P_L for one female subject. A square root relation is expected from Equation (5) if all other variables are constant or offset each other, which is unlikely. It is clear, however, that PCR must go to zero when P_L goes to zero.

With that constraint, the $\frac{1}{2}$ power fit is more defensible than a linear regression. Figures 7(b) and (c) show similar fits for all seven female subjects and six male subjects, respectively. The curves have a goodness of fit (normalized root mean square error) of 0.52 across all subjects. *PCR* varied by at least 3:1 across both male and female subjects, suggesting that *PCR* may have discriminatory potential between subjects in terms of effort expended. In Schutte's normal subject data, a range of 0.003 – 0.01 % efficiency was found for 1.0 kPa of lung pressure, which is the same 3:1 ratio.

Figure 8 shows range of *SPL* versus range of P_L , comparing our data (narrow tube) to Schutte's data (wide tube). Range of *SPL* achievable by our subjects was diminished due to the narrow (0.6 cm diameter) tube compared to Schutte's wider tube (≈ 3 cm diameter), a result that is predictable from sound radiation theory (Equation 9). The narrow tube produces a range of about 10 dB *SPL* with a 1 kPa range of pressure, while the wide tube produces a range of nearly 20 dB *SPL* with a 0.4 kPa range of pressure.

PCR versus E as calculated from Equation (5) for our tube-mouth configuration is shown in Fig. 9. The proportionality factor between *PCR* and $\sqrt{E}$ was 45 for $d_t = 0.6$ cm, $A_t/A_m = 1$, $R_r/R_{dc} = 200$, $P_L/P_{dc} = 10$. This factor is very close to that obtained from Schutte's wide tube data (recall a value of 50 in the curves of Figure 5).

PCR versus measured *SPL* is shown in Fig. 10 for all thirteen subjects, but only the minimum, maximum and mean points for individual subjects are shown on the curves for clarity. The curve obtained from analytical calculations for our tube-mouth configuration using equation (8) is also shown (thicker line and no data points). It underestimates the data for most of the subjects, suggesting a need for refinement of the constants in Eqn. (8).

Finally, Fig. 11 shows the histogram of open quotients computed during voicing regions for all thirteen subjects. It can be seen that an open quotient between 0.5 and 0.7 is prevalent, indicating that subjects were not tending towards pressed voice.

DISCUSSION AND CONCLUSIONS

The subject of vocal efficiency, almost forgotten in the literature, has been re-addressed with a simpler measure that is clinically feasible, namely a pressure conversion ratio (*PCR*) obtained at the mouth. The first question posed in the Introduction, how does the *PCR* measure relate to vocal efficiency, has been answered both theoretically and empirically. The relation is $PCR = k\sqrt{E}$, where k was 50 for Schutte's data and 45 for our subject data. The value of k may be able to be maintained reasonably constant across different phonations by using a mouthpiece with fixed mouth area and otherwise neutral vocal tract. The ratios P_L/P_{dc} and R_r/R_{dc} in Equation (5) can in the future be quantified using a bench setup and measuring P_{dc} , U_{dc} for a given P_L and the selected mouthpiece.

The second question, do efficiency and *PCR* have a bias toward pressed voice, has also been answered. When the objective is a specific loudness at a specific f_o , open quotients near 0.5 are produced more frequently than extreme Q_o values (near 0.0 or 1.0). It is speculated here that *PCR* may be governed more by maximum power transfer than by maximum efficiency. Generally, the efficiency is so low that maximization of efficiency has less pay-off than

maximization of power transfer from the glottis to free space. In electrical circuit theory, it is known that maximum power transfer occurs at a 50 % efficiency, but even this value is never reached. Thus, all we can say here is that there is no bias toward pressed voice that could hinder the usefulness of the *PCR* measure for describing ease of phonation (or, conversely, effort) in a clinical setting.

The third question, is there sufficient variation in *PCR* across subjects for a discriminatory “effort” measure, has only a preliminary answer. Normal subjects showed ranges of *PCR* from 0.2 to 3.0, which would suggest a respectable range for variations with loudness and f_o in the speech range. However, while vocally impaired patients in Schutte’s² study produced efficiency values significantly outside of the normal range, differences were not large in the mid to high range of intensity. We expect, therefore, that *PCR* can best discriminate between normal subjects and patients at low intensity (near phonation threshold).

Future studies will address how *PCR* varies in higher f_o ranges. Preliminary data suggest that, at high f_o , there is not an increase in *PCR* with increased loudness. In other words, conversion from DC to AC pressures may reach a limit with increased lung pressure when the vocal folds are stiffer. Such a result would be in agreement with classic voice range profile measurements, which usually show a mid-value of f_o has the greatest *SPL* range, rather than the extreme values of f_o .

Acknowledgments

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

References

1. Titze, IR. Principles of Voice Production. Salt Lake City, UT: National Center for Voice and Speech; 2000. p. 247Chapter 9
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. Gronigen: State University Hospital; 1980.
4. Grillo EU, Verdolini K. Evidence for distinguishing pressed, normal, resonant, and breathy voice qualities by laryngeal resistance and vocal efficiency in vocally trained subjects. *J Voice.* 2008; 22(5):546–52. [PubMed: 17400426]
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. Jiang JJ, O’Mara T, Conley D, Hanson D. Phonation threshold pressure measurements during phonation by airflow interruption. *Laryngoscope.* 1999; 109:425–432. [PubMed: 10089970]
11. Berry DA, Verdolini K, Montequin DW, Hess MM, Chan RW, Titze IR. A quantitative output-cost ratio in voice production. *J Speech Lang Hear Res.* 2001; 44:29–37. [PubMed: 11218106]

12. Titze, IR. The Myo-elastic Aerodynamic Theory of Phonation. Salt Lake City, UT: National Center for Voice and Speech; 2006a. p. 303-311.
13. Jiang JJ, Titze IR. Measurement of vocal fold intraglottal pressure and impact stress. *J Voice*. 1994; 8:132-144. [PubMed: 8061769]
14. Hess MM, Verdolini K, Bierhals W, Mansmann U, Gross M. Endolaryngeal contact pressures. *J Voice*. 1998; 12:50-67. [PubMed: 9619979]
15. Verdolini K, Hess MM, Titze IR, Bierhals W, Gross M. Investigation of vocal fold impact stress in human subjects. *J Voice*. 1999; 13:184-202. [PubMed: 10442749]
16. Titze IR. Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation. *J Speech Lang Hear Res*. 2006b; 49:439-447. [PubMed: 16671855]
17. Titze IR, Laukkanen A-M. Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. *Logoped Phoniatr Vocol*. 2007; 32:147-156. [PubMed: 17917981]
18. 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]
19. Stathopoulos E, Sapienza C. Respiratory and laryngeal function of women and men during vocal intensity variation. *J Speech Hear Res*. 1993; 36:64-75. [PubMed: 8450666]
20. Gauffin J, Sundberg J. Spectral correlates of glottal voice source waveform characteristics. *J Speech Hear Res*. 1989; 32:556-565. [PubMed: 2779199]
21. Granqvist S, Hertegard S, Larsson H, Sundberg J. Simultaneous analysis of vocal fold vibration and transglottal airflow: exploring a new experimental setup. *J Voice*. 2003; 17:319-330. [PubMed: 14513955]
22. Isshiki, N. Clinical significance of a vocal efficiency index. In: Titze, IR., Scherer, RS., editors. *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*. Denver, CO: The Denver Center for the Performing Arts; 1983. p. 230-238.
23. Rothenberg M. A New Inverse-Filtering Technique for Deriving the Glottal Airflow Waveform during Voicing. *J Acoust Soc Am*. 1973; 53:1632-1645. [PubMed: 4719255]
24. Titze IR. Phonation threshold pressure measurement with a semi-occluded vocal tract. *J Speech Language Hear Res*. 2009; 52:1062-1072.
25. Odum HT, Pinkerton RC. Time's speed regulator: The optimum efficiency for maximum output in physical and biological systems. *Am Sci*. 1955; 43:331-343.
26. Heman-Ackah YD, Heuer RJ, Michael DD, Ostrowski R, Horman M, Baroody MM, Hillenbrand J, Sataloff RT. Cepstral peak prominence: A more reliable measure of dysphonia. *Ann Otol Rhinol Laryngol*. 2003; 112:324-333. [PubMed: 12731627]

Fig. 1.
Dual cannula *PCR* instrumentation.

Fig. 2.
Pressures in short tube for relating efficiency E to PCR .

Fig. 3.
Future mouthpiece design to maintain constant mouth area and house a dual cannula.

Fig. 4. Summary of glottal efficiency data from Schutte.³ '+' represent normal subjects' data and '.' represent disordered subjects' data.

Fig. 5. Predicted PCR from Schutte's³ data. (a) PCR versus E , (b) PCR versus SPL . Solid lines are from measured P_L values and dashed lines from measured SPL values.

Fig. 6. A sample DC pressure plot showing shuttering between straw-open and straw-occluded regions from human subjects.

Fig. 7.

(a) PCR vs. P_L data from one female subject with square root fit. (b) PCR vs. P_L curves for seven female subjects. (c) PCR vs. P_L curves for six male subjects.

Fig. 8.

Relationship between P_L and SPL in narrow-tube data and wide-tube (Schutte) data.

Fig. 9.
PCR versus predicted efficiency E from analytical calculations.

Fig. 10.
PCR versus measured *SPL* for thirteen subjects.

Fig. 11.
Histogram of open quotient across all the 13 subjects.