An Oral Pressure Conversion Ratio as a Predictor of Vocal Efficiency

*,†Ingo R. Titze, *Lynn Maxfield, and *Anil Palaparthi, *Salt Lake City, Utah, and †Iowa City, Iowa

Summary: 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 article uses 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 because of low cost and ease of use in clinics and vocal studios.

Key Words: 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 for a human adult (www.mydailyintake.net) or about 2000 kcal. The rate of consumption, or the power input, is about 100 W (8.7×10^6 J/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 intensity level (SIL) range of 70–90 dB at 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.[5–10] 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 μPa 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. Although 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, Titze[16] and Titze and Laukkanen[17] 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,[18–20] whereas MADR is closely related to tissue velocity (and hence to momentum change and impact stress) before 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

Accepted for publication June 2, 2015.

From the *National Center for Voice and Speech, The University of Utah, Salt Lake City, Utah; and the †Department of Communication Sciences and Disorders, The University of Iowa, Iowa City, Iowa.

Address correspondence and reprint requests to Ingo R. Titze, National Center for Voice and Speech, The University of Utah, 136 South Main Street, Suite 320, Salt Lake City, UT 84101-3306. E-mail: ingo.titze@ncvs2.org

Journal of Voice, Vol. 30, No. 4, pp. 398-406
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2015.06.002

to widespread clinical adoption of an MFDR/MADR ratio as a measure of vocal economy.

The purpose of the present study was to explore a third option, an aerodynamic-to-acoustic pressure conversion ratio (PCR), 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 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 previously, 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 as[1]

$$E = \frac{4\pi r^2 I_0 10^{\mathrm{SPL}/10}}{P_\mathrm{L} U_\mathrm{g}}, \tag{1}$$

where I_0 is the standard reference intensity (10^{-12} W/m^2), 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. Because power is the rate of energy produced or absorbed, and because 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 widespread application as a vocal effort measure. First, the aerodynamic energy asymptotes to zero when the open quotient (duty ratio) in the glottis approaches zero $(U_\mathrm{g} \rightarrow 0)$. 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 measure[24] 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 (Figure 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 PCR is defined simply as

$$\mathrm{PCR} = \frac{P_\mathrm{ac}}{P_\mathrm{dc}}, \tag{2}$$

where P_ac is the oral acoustic pressure (behind the lips) and P_dc is the oral steady pressure (also behind the lips). These

FIGURE 1. Dual cannula PCR instrumentation.