Laryngeal flow due to longitudinal sweeping motion of the vocal folds and its contribution to auto-oscillation (L)

Henri Boutin, a) John Smith, and Joe Wolfe
School of Physics, The University of New South Wales, Sydney, New South Wales 2052, Australia

(Received 10 November 2014; revised 23 May 2015; accepted 1 June 2015; published online 8 July 2015)

Analysis of published depth-kymography data [George, de Mul, Qiu, Rakhorst, and Schutte (2008). Phys. Med. Biol. 53, 2667–2675] shows that, for the subject studied, the flow due to the longitudinal sweeping motion of the vocal folds contributes several percent of a typical acoustic flow at the larynx. This sweeping flow is a maximum when the glottis is closed. This observation suggests that assumption of zero laryngeal flow during the closed phase as a criterion when determining parameters in inverse filtering should be used with caution. Further, these data suggest that the swinging motion contributes work to overcome mechanical losses and thus to assist auto-oscillation.

I. INTRODUCTION

The acoustic flow entering the vocal tract at the larynx (hereafter the laryngeal flow) plays a fundamental role in the production of speech. However, it cannot usually be measured directly. The waveform of this laryngeal flow (often called the glottal flow) is often estimated using inverse filtering of the radiated sound (e.g., Alku 1992, 2011; Alku et al., 2006a; Alku et al., 2006b; Alku et al., 2009; Airas, 2008). This approach involves separating two unknowns with complex behavior (the laryngeal flow and the transfer function of the tract) using only one measured signal (the radiated sound). This procedure cannot be unequivocal (Drugman, 2012; Chu et al., 2013), and consequently assumptions about the waveform of the laryngeal flow are often made in order to adjust the anti-resonances of a set of inverse filters to remove the filtering effects of the vocal tract and the radiation from the lips. To this end, it is sometimes assumed that the laryngeal flow is very close to zero while the glottis is closed. This assumption is not necessarily true: if the vocal folds, while closed, are moving in the vertical direction, then they displace air in that direction, as observed by Granqvist et al. (2003) in the case of “flow phonation,” using inverse filtering. We refer to this hereafter as sweeping flow: the component of laryngeal flow due to the sweeping motion of the folds. If one used the term glottal flow to describe only the mass injection associated with the flow through the glottis, i.e., through the (vibrating) aperture between the vocal folds, while closed, are moving in the vertical direction, then the effective length would be 9.1 mm/2 = 4.6 mm. The effecting flow corresponds closely with the beginning of

II. SWEEPING FLOW DUE TO THE VOCAL FOLDS’ MOTION

For an upright speaker, define the lateral direction (i.e., the horizontal left-right axis) as x, the forwards direction as y, and the (nearly vertical, longitudinal) direction parallel to the supraglottal tract, as z. The z motion of the vocal folds has been quantified by George et al. (2008), who used triangulation of laser beams in combination with nasendoscopy. (The cited study measured only one subject, and the motions of this subject’s vocal folds were not symmetrical—in the following analysis the data for both vocal folds are combined.) George et al. plot zx(t) at a constant y position roughly midway across the glottis. Its amplitude, about 2 mm, has the same order of magnitude as the estimate given by Granqvist et al. (2003), between 1 and 2 mm, and the measurements of Larsson (2009), between 0.3 and 1.5 mm.

To calculate the sweeping flow from these data, z(x, t) is integrated along the x-direction for both vocal folds, differentiated with respect to time and multiplied by an effective length L of the moving vocal folds in the y-direction. L is obviously less than the total length (9.1 mm) measured on Fig. 2 from George et al. (2008). If the shape along this length were a sine function from 0 to π, the effective length would be 9.1 mm * 2/π = 5.8 mm. If it varied as sin² from 0 to π, the effective length would be 9.1 mm/2 = 4.6 mm. The effective length is set at 5 mm for this order-of-magnitude calculation. The differentiation is performed by taking the Fourier transform, multiplying each component at angular frequency ω by jω and then taking the inverse transform. Figure 1(top) displays this sweeping flow averaged over four complete cycles before (pale curve) and after (dark curve) the harmonics above the third have been removed. Figure 1(bottom) shows the open area of the glottis (i.e., the glottal aperture) A(t) calculated from the same data set. (Two complete cycles are shown.)

Considering the filtered curve in Fig. 1(top), the peak of the sweeping flow corresponds closely with the beginning of


Copyright 2015 Acoustical Society of America

Pages: 146–149

a)Electronic mail: boutin@lam.jussieu.fr
the closed glottis phase and has a magnitude of 34 mL s$^{-1}$. Ideally, this sweeping flow would be compared with total laryngeal flow measured simultaneously. However, the laryngeal flow was not measured by George et al. (2008), whose subject was “a trained professional able to produce a sustained phonation over a long period of time.” A lower limit on the ratio of sweeping flow to glottal flow can be estimated by considering the extreme case of a phonation in which 5 liters of air is expended in 5 s—a very loud phonation. Approximating the laryngeal flow as a square wave, the peak value would be 2 liters s$^{-1}$. In this case, the sweeping flow in Fig. 1 would be only 1.7% of the total laryngeal flow. On the other hand, if the laryngeal flow corresponding to the data of Fig. 1 had a more typical amplitude of 0.3 liters s$^{-1}$ (see examples given by Södersten et al., 1995), then the sweeping flow would be about 11% of the laryngeal flow. Further, its maximum occurs near the beginning of the period of glottal closure. Using a numerical model of vocal fold motion that allows both longitudinal and lateral motion, Flanagan and Ishizaka (1978) calculate the sweeping and laryngeal flows. With the parameters used, they find that the peak-to-peak sweeping flow is five percent of the total peak-to-peak laryngeal flow. Their lower value presumably reflects differences between their model and the vocal folds measured by George et al.

A sweeping flow with a relative magnitude of several percent or more might have implications for inverse filtering if the filtering algorithms used (whether automated or manual) assume a zero laryngeal flow during the closed glottis phase, or indeed a ripple-free flow over that phase (e.g., Bäckström et al., 2005; Lehto et al., 2007; Airas, 2008).

It is also interesting that other researchers often find a small flow during the “closed” phase, which is usually attributed to glottal leakage (e.g., Holmberg et al., 1988; Cranen and Schroeter, 1995; Iwarsson et al., 1998). This could be at least partially a consequence of the presence of a sweeping flow.

III. AUTO-OSCILLATION OF THE VOCAL FOLDS

Apart from its possible implications for inverse filtering, the presence of a non-negligible longitudinal motion has another potentially important implication for speech science.

In the following discussion, several phenomena are not mentioned. These include not only the modulation of the flow between the vocal folds, but also the deformation of the folds on collision. This omission from discussion does not, of course, imply that these are negligible.

Consider the highly simplified, one-mass, two-spring model for the vocal folds shown in Fig. 2, which is similar to that of Awrejcewicz (1990), Adachi and Yu (2005), and others. Consider the case where the frequency of vibration is well below those of the resonances of the supraglottal and subglottal tracts. (This is the case of the data analysed here, and is usually the case for both men’s and women’s speech.) At such a frequency, the supraglottal and subglottal impedances are small, so the supraglottal and subglottal pressures are approximately constant. Flow separation at such a constriction is complicated (see, e.g., Hirschberg, 1992). For the very simple model presented here, flow separation is assumed to occur at the narrowest aperture with the consequence that the pressure between the folds and everywhere downstream from them is approximately uniform at $P_2$, and that the pressure upstream is uniform at $P_1 > P_2$. What is the work done on the folds by $(P_1 - P_2)$ over one complete cycle?

Assume that (as shown in Fig. 2) the phase of the motion in the $z$ direction is ahead of that of the separation of the vocal folds in the $x$ direction, in agreement with the observations of vocal fold vibration of Baer (1981) and the model of Adachi and Yu (2005). In consequence, the length of the folds in the $x$ direction is greater when the folds are moving upwards [phase (a)–(b) in Fig. 2] than when they are moving downwards [phase (c)–(d) in Fig. 2]. Consequently, the positive component of longitudinal laryngeal flow due to the upwards swinging motion of the longer folds is greater than the negative component due to downwards swinging motion of the shorter folds, see Fig. 2. Subject to the conditions listed above, positive work is then performed by the pressure difference $(P_1 - P_2)$ over one complete cycle (Boutin et al., 2015) solely due to the swinging motion of the vocal folds. This work done can overcome losses, and might provide a significant contribution towards auto-oscillation, particularly in cases where the phases of the load impedances may not be favorable.
An order-of-magnitude estimate of the rate of positive work done on the vocal folds by \((P_1 - P_2)\) can be calculated using the data of George et al. (2008). Assuming a transglottal pressure difference of 1 kPa, the average power produced by this swinging motion of the vocal folds is 0.5 mW. The radiated acoustic power of the voice is in the range 10^{-3} to 10 mW (Titze, 1992), but much of this power comes from the modulation by the vibrating vocal folds of the flow between them. The energy stored in the vibratory motion of the folds themselves and the fraction of this that is lost each cycle are not known. If, as seems likely, the rate of loss of vibratory energy is on the order of a milliwatt, the power calculated above may be a significant contribution to auto-oscillation.

IV. CONCLUSION

The measurements of George et al. (2008) were made on just one subject. Consequently much could be learned from similar measurements made on a larger sample, especially if simultaneous measurements could be made of flow at the lips, of the acoustical properties of the tract and/or of glottal behavior using electroglossotography. One of the purposes of this paper is to encourage that. However, one immediate conclusion from this analysis is that caution should be exercised in assuming zero laryngeal flow in the closed glottis. The dark and light gray areas represent respectively positive and negative volumes swept by the vocal folds during the intervals (a)-(b) and (c)-(d). In the lower part, the black and gray sine waves are approximations to the oscillations in the z and x directions, and illustrate their phase relation. (This sine approximation requires deformation of the vocal folds during their collision.)

ACKNOWLEDGMENTS

We thank Harm K. Schutte and Frits F.M. de Mul for sending us the original versions of the plots in George et al. (2008) and Associate Editor A. Hirschberg for helpful comments. The support of the Australian Research Council is gratefully acknowledged.


