EXAMINING THE ACOUSTIC CONTRIBUTIONS OF THE EPILARYNGEAL TUBE TO THE VOICE SOURCE AND VOCAL TRACT RESONANCE

Scott R. Moisik[1] and John H. Esling[1]
[1]Dept. of Linguistics, University of Victoria, BC, Canada, srmoisik@uvic.ca

1. INTRODUCTION

In this study, we provide observations from videofluoroscopic and high-speed laryngoscopic techniques, combined with spectral analysis, to increase our understanding of the acoustic nature of aryepiglottic trilling and the epilaryngeal tube in general. The underlying goal is to work towards a complete biomechanical model of the laryngeal framework that acknowledges the role of the supraglottal structure of the larynx in shaping the nature of the voice source and of vocal tract resonance.

The epilaryngeal tube (sometimes referred to as the laryngeal vestibule; see Fig. 1) constitutes an important component of the larynx that can modulate the voice source in several ways and partly determines the resonances of the lower vocal tract. Under static configurations, Titze (2008) demonstrates how epilaryngeal tube constriction changes the acoustic coupling between the sub- and supralaryngeal vocal tract sections, giving rise to changes in vocal fold dynamics of a non-linear nature (where the filter is influencing the source). With sufficient airflow, various parts of the epilaryngeal tube can also be set into oscillation. This means that the acoustic coupling effect of the epilarygeal tube can vary as a function of time.

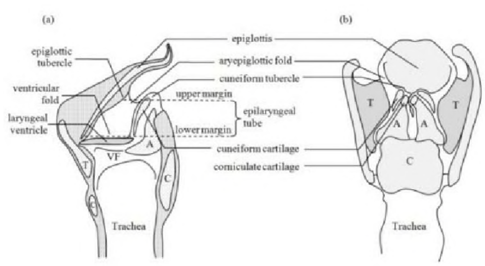

Fig. 1. Epilaryngeal tube anatomy. Mid-sagittal (a); Posterior (b).

There are three main components of the epilaryngeal tube that self-oscillate: the ventricular folds, the aryepiglottic folds, and the epiglottis. Ventricular fold oscillation (or ventricular phonation) in non-pathological cases occurs simultaneously with vocal fold oscillation (Fuks et al. 1998; Lindestad et al. 2002). It is most familiar in the ethnic singing styles of East Asia (e.g. Mongolian, Tibetan, and Tuvan throat singing). Notably, the ventricular folds undergo an irregular transient phase before reaching a highly stable steady state of oscillation where they become entrained with the vocal folds (at 1/2, 1/3, or equal to the glottal F0). Aryepiglottic (AE) oscillation (or trilling), impressionistcally labeled 'growling', is less thoroughly documented (although see Sakakibara et al. 2004). AE trilling is commonly heard in pop music singing styles (most famously by Louis Armstrong), but can be heard in throat clearing and coughing, and also occurs as a phonetic variation of pharyngeal segments in languages such as Iraqi Arabic and in a subset of phonatory registers, most notably in !Xóõ (Painter 1986; Esling & Harris 2005). When the epilaryngeal tube is heavily constricted, the aryepiglottic folds run perpendicular to the longitudinal dimension of the vocal and ventricular folds, and, depending on individual anatomical configuration, there is the possiblity for several epilaryngeal tube apertures or air-channels to exist. The region from the lateral margin of the epiglottis to the cuneiform tubercles defines the left and right apertures associated with the left and right AE folds. It is also common to see a medial aperture/channel formed by the medial surfaces of the cuneiforms. Furthermore, the arytenoid-corniculate complex can also form an aperture. Sakakibara et al. (2004) estimate that the lumen of the epilaryngeal tube contributes a pole in the vocal tract transfer function around 1.5 kHz but they note that there also acoustic losses due to the laryngeal ventricle. Finally, we have observed epiglottal oscillation (or epiglottal trilling; estimated to be at ~30-40 Hz; see Fig. 2) during the production of voiceless pharyngeals in Somali. The speaker in Fig. 2 possesses a flat epiglottis, which likely facilitates the oscillation. It remains an open question how common the linguistic use of epiglottal trilling is compared with aryepiglottic trilling. Even in this example, the aryepiglottic folds can be seen to oscillate.

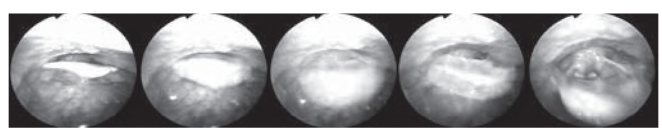

Fig. 2. An epiglottal trill (234 ms) in Somali (/ħaniin/ 'testicle').

2. METHOD

Videofluoroscopic (VIF) and high-speed laryngoscopic (HSL; 500 Hz) data of voiced aryepiglottic trills (produced at a glottal F0 of 100 Hz in an [a] context) were obtained. An estimate of epilaryngeal tube compliance is made based on the VIF data using ImageJ to obtain pixel areas of the epilaryngeal tube lumen which were then converted to metric by using an estimated scaling factor. A laryngographic (EGG) and acoustic signal were simultaneously collected with the HSL data, and these are analyzed in conjunction with AE aperture information obtained from analysis of the HSL using kymographic and simple image analysis techniques (see Moisik et al. 2010 for more details based on a similar analytical approach).

3. RESULTS & DISCUSSION

The lumen of the epilaryngeal tube expands as airflow from the glottis impinges upon the surface of the tube immediately below the tubercle of the epiglottis (Fig. 3). This compliance is suspected to play an important role in self-sustained oscillation of the AE folds by inducing a phase shift between the volume velocity and the AE aperture area, effectively breaking the symmetry in energy exchange between the AE folds and the air flow.

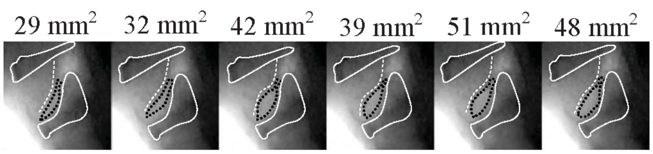

Fig. 3. Estimated epilaryngeal tube sagittal cross-sectional area using VIF. Black dotted line = measured area; white outlines = hyoid bone and cricoid-arytenoid cartilages; white dashed line = epiglottis surface contour.

Although the AE trill is highly unstable/irregular in its vibratory pattern (see Moisik et al. 2010), there are moments when entrainment between the vocal folds and the AE folds stabilizes; such a moment is presented in Fig. 4. In this figure, the right AE aperture (the left is statically patent during this sequence) is compared with the glottal aperture (indicated by the inverted EGG signal which shows moments of glottal opening as peaks). During the AE opening phase (markers 1, 3, and 5 in Fig. 4) vocal fold closure velocity is monotonic and rapid; the acoustic state becomes resonant and is characterized by strong spectral density up to the resonances associated with the F1 and F2 at which point the energy in the signal dips until around 1.5-1.7 kHz, where another resonance is encountered (F3; although this value is close to the resonance said to be associated with the epilaryngeal tube by Sakakibara et al. (2004)). During the AE closure phase (markers 2 and 4) there is an increase in the high frequency components as evident in the acoustic signal and spectral envelopes. This is likely attributeable to the airflow during the glottal open phase impacting with the inferior surface of the right AE fold. Vocal fold movement during the closure phase of the AE cycle is delayed, which may be a result of increasing back pressure on the vocal folds due to the ever narrowing epilaryngeal tube air channel. The spectral analysis reveals that at this point in the cycle a zero is introduced into the vocal tract filter in the region of 400-500 Hz. This acoustic loss may reflect the geometry of the epilaryngeal tube, which, in a way similar to lateral sounds, has a medial obstruction to airflow due to the cuneiform tubercles.

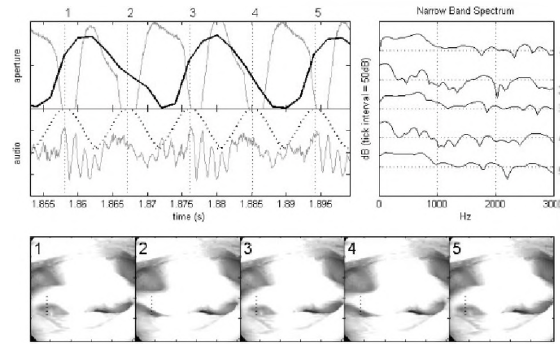

Fig. 4. Right AE aperture vs. inverted EGG (to show opening pattern). Dotted lines in audio plot show 10 ms Hamming windows used to calculate the FFT for spectral moments corresponding to different phases of the trill at glottal closure. Dotted black line in HSL = kymographic line used for estimating right AE aperture.

4. CONCLUSION

The epilaryngeal tube is a unique structure in the vocal tract in that it can simultaneously modulate and contribute to the vocal source as well as shape the acoustic coupling between the source and filter and yield its own resonances. This work represents a continual effort to expand the knowledge we have of this structure with the hopes of incorporating this knowledge into a more robust compuational model of laryngeal dynamics and acoustics.

REFERENCES

Esling, J. H. & Harris, J. G. 2005. States of the glottis: an articulatory phonetic model based on laryngoscopic observations. In *A figure of speech: a Festschrift for John Laver*. Hardcastle, W. J. and J. M. Beck (eds.). Mahwah, NJ: Erlbaum. 347–383.

Fuks, L., Hammarberg, B., & Sundberg, J. (1998). A self-sustained vocal-ventricular phonation mode: acoustical, aerodynamic and glottographic evidences, *KTH TMH-QPSR*, 3, 49-59.

Moisik, S. R., Esling, J. H., & Crevier-Buchman, L. 2010. A high-speed laryngoscopic investigation of aryepiglottic trilling. JASA 127 (3). 1548-1558.

Titze, I. R. (2008). Nonlinear source-filter coupling in phonation: Theory. JASA 125 (5), 2733-3749.

Traill, A. 1986. The laryngeal sphincter as a phonatory mechanism in !Xóõ (Bushman). In R. Singer and J. K. Lundy (Eds.), Variation, Culture and Evolution in African Populations: Papers in Honour of Dr. Hertha de Villiers. 123-131. Johannesburg: Witwatersrand University Press.

Sakakibara, K. I., Fuks, L., Imagawa, H., & Tayama, N. (2004). Growl voice in pop and ethnic styles. In Proceedings of the ISMA. Nara, Japan.