Proceedings of the 10th International Conference on
Advances in Quantitative Laryngology, Voice and Speech Research
June 3-4, 2013, Cincinnati, Ohio, USA

HIGH-SPEED VIDEOMICROSCOPY AND ACOUSTIC ANALYSIS OF *EX VIVO* VOCAL FOLD VIBRATORY ASYMMETRY

Daryush D. Mehta[1,2], Georg Luegmair[3], James B. Kobler[1,2], Robert E. Hillman[1,2,4], Alyson A. Young[1,5], Melissa L. Cooke[1,4], Michael Döllinger[3]

[1] Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA, USA
[2] Department of Surgery, Harvard Medical School, Boston, MA, USA
[3] Department for Phoniatrics and Pedaudiology, University Hospital Erlangen Medical School, Erlangen, Germany
[4] Institute of Health Professions, Massachusetts General Hospital, Boston, MA, USA
[5] Department of Speech, Language & Hearing Sciences, Boston University, Boston, MA, USA

***Keywords*: High-speed video; Vocal-fold vibration; Vocal-fold asymmetry, Acoustic voice and speech analysis**

INTRODUCTION

Voice specialists make critical diagnostic, medical, therapeutic, and surgical decisions based on coupling visual observations of vocal fold tissue motion with auditory-perceptual judgments of voice quality; i.e., in the typical setting, clinicians relate what they see in endoscopic recordings to what they hear [1]. Although clinical experience indicates that this approach provides information that is generally useful in the diagnostic process, it is inherently limited to case-by-case observations, and visual judgments of vocal fold kinematics—such as vibratory asymmetry—might not adequately reflect changes in objective measures of the acoustic signal [2].

In a high-speed videoendoscopy (HSV) study of 14 voice patients, the average degree of vocal fold vibratory asymmetry did not correlate statistically with levels of acoustic jitter, shimmer, or harmonics-to-noise ratio during sustained vowel phonation [3]. Instead, a significant amount of variation in acoustic jitter was accounted for by the standard deviation of the degree of asymmetry, pointing to cycle-to-cycle variation as being more important than absolute levels of vibratory asymmetry. In a study of 20 subjects who underwent phonomicrosurgery for vocal fold lesions, the acoustic cepstral peak magnitude correlated significantly with an HSV-based measure combining fundamental frequency deviation and average speed quotient ($r = 0.70$) [4].

The purpose of this study is to determine the correlation between measures of vocal fold vibratory asymmetry derived from high-speed videomicroscopy (HSVM) recordings and acoustic voice quality measures in an *ex vivo* larynx setup that enables precise control over the torque asymmetry of the arytenoids. Based on our previous work in human subjects, it is hypothesized that cycle-to-cycle variations in period duration and the degree of glottal closure are expected to correlate with acoustic measures that reflect degradations in voice quality. These acoustic degradations are expected to occur regardless of the absolute degree of vocal fold asymmetry.

METHODS

Fig. 1 displays the typical experimental setup of an excised human larynx whose vocal folds were driven into vibration using humidified airflow. A key component of the setup consisted of digitally-controlled manipulators of arytenoid rotation that allowed for individual control over the degree of tension on the left and right vocal folds. The manipulators held the arytenoids from above using a three-pronged rod. A thyroid cartilage load was sometimes necessary to stabilize the larynx during phonation. The whole-mount excised larynges were prepared by dissecting away supraglottal structures, suturing the ventricular folds, and leveling the thyroid cartilage to obtain full exposure superiorly of the true vocal folds. Inferiorly, the trachea was cut to a length of approximately 5 cm and mounted on a cylindrical pipe connected to the airflow supply.

Properties of sustained vocal fold vibration were captured for three excised human larynges using synchronized recordings of HSVM, arytenoid torque, subglottal pressure, and acoustics. A pneumatic pressure gauge regulated the airflow, and a low-bandwidth transducer (Glottal Enterprises) tracked the subglottal

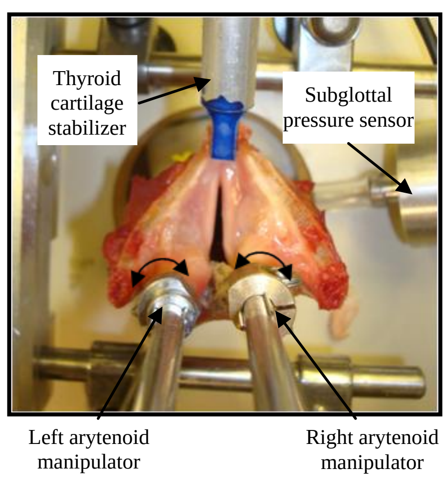

Fig. 1: Ex vivo human larynx setup.

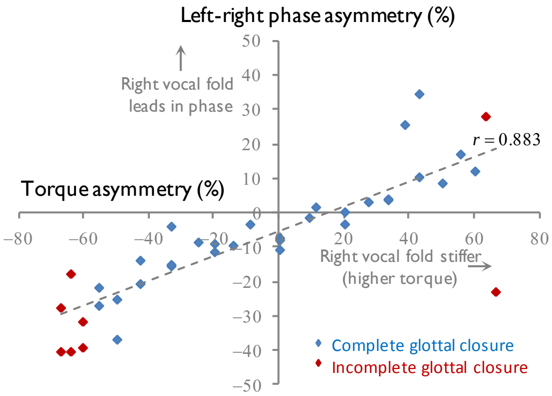

Fig. 2: HSVM-based vibratory phase asymmetry versus arytenoid torque asymmetry.

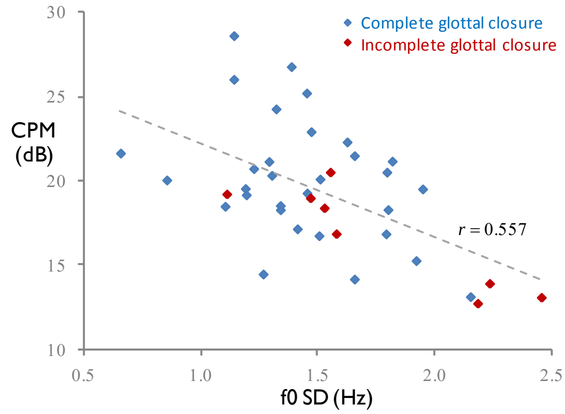

Fig. 3: Acoustic cepstral peak magnitude (CPM) versus HSVM-based f0 standard deviation (f0 SD).

pressure approximately 6 cm below the vocal folds. HSVM data were recorded through a Leica F40 surgical microscope using a monochromatic high-speed video camera (Phantom v7.1; Vision Research, Inc.) capturing 4000 images per second. To induce a particular vocal fold vibratory asymmetry, different control signals were sent to each arytenoid manipulator to apply various values of torque (in mN m) that served to stiffen a particular vocal fold. Torque asymmetry was defined as right minus left arytenoid torque, divided by their sum.

Correlation analysis (Pearson's r, $p < 0.05$) determined strengths of relationships between HSVM-based measures of vocal fold vibratory asymmetry [5] and time- and cepstral-based acoustic measures commonly linked to voice quality [4]. Data from one excised human larynx (39 trials) are reported here for which a spectrum of asymmetries was observed.

Results and Discussion

Initial results yield insight into correlations among measures of left-right vocal fold vibratory asymmetry and the acoustic cepstral peak magnitude. Fig. 2 displays the relationship between imposed torque asymmetry and vocal fold vibratory phase asymmetry ($r = 0.883$), providing validation metric of the experimental setup.

As expected, the acoustic cepstral peak magnitude (CPM) did not correlate linearly with vibratory phase asymmetry. A range of CPM values reflected stable acoustics for several levels of vibratory phase asymmetry. High levels of asymmetry—in particular, levels that created incomplete glottal closure—induced aperiodic vibratory patterns that degraded the acoustic signal in terms of a lower cepstral peak magnitude, higher aperiodicity, and increased spectral anharmonicity.

Fig. 3 displays a moderate correlation ($r = 0.557$) expected between the glottal area–based fundamental frequency deviation and the acoustic CPM, which is in agreement with measurements of postsurgical change found previously in voice patients [4].

Conclusion

Cycle-to-cycle regularity and complete glottal closure patterns continue to act as more important predictors of acoustic outcome than absolute levels of vocal fold asymmetry; although, excessively high asymmetries often are concomitant with incomplete glottal closure and a lack of vocal fold vibratory entrainment. Future work seeks to add three-dimensional measures of vocal fold kinematics to potentially complement HSVM-based measures. Long term, salient results from this study can be transferred to other laryngeal models and human studies protocols.

Acknowledgments

Research supported by the American Speech-Language-Hearing Foundation, Boston University (undergraduate grant), NIH (R01 DC007640), and DFG FOR 894/1-2.

References

[1] Zeitels SM, Blitzer A, Hillman RE, Anderson RR. Foresight in laryngology and laryngeal surgery: A 2020 vision. *Ann Otol Rhinol Laryngol* 2007;116(9) Suppl 198:1–16.

[2] Haben CM, Kost K, Papagiannis G. Lateral phase mucosal wave asymmetries in the clinical voice laboratory. *J Voice* 2003;17:3–11.

[3] Mehta DD, Deliyski DD, Zeitels SM, Quatieri TF, Hillman RE. Voice production mechanisms following phonosurgical treatment of early glottic cancer. *Ann Otol Rhinol Laryngol* 2010;119:1–9.

[4] Mehta DD, Zeitels SM, Burns JA, Friedman AD, Deliyski DD, Hillman RE. High-speed videoendoscopic analysis of relationships between cepstral-based acoustic measures and voice production mechanisms in patients undergoing phonomicrosurgery. *Ann Otol Rhinol Laryngol* 2012;121:341–347.

[5] Mehta DD, Deliyski DD, Quatieri TF, Hillman RE. Automated measurement of vocal fold vibratory asymmetry from high-speed videoendoscopy recordings. *J Speech Lang Hear Res* 2011;54:47–54.