An impedance-based inverse filtering scheme with glottal coupling

Matías Zañartu¹, Julio C. Ho², Daryush D. Mehta³,⁴, Robert E. Hillman³,⁴,⁵, and George R. Wodicka¹,²

¹School of Electrical and Computer Engineering, Purdue University; ²Weldon School of Biomedical Engineering, Purdue University; ³Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital; ⁴Speech and Hearing Biosci. and Tech., Harvard-MIT Health Sci. and Tech.; ⁵Harvard Medical School

Introduction
Continuous monitoring of vocal fold behavior after voice surgery and during voice therapy is critical. The acceleration of the skin overlying the suprasternal notch has been used to estimate a reduced number of glottal parameters. However, the linear source-filter assumption that is typically employed has not allowed for accurate estimates of many clinical parameters of interest.

Study objectives
To develop a model-based inverse filtering scheme that considers the coupling effects of a finite glottal impedance and the mechanical-acoustical properties of the skin over the suprasternal notch.
To evaluate the proposed scheme for the assessment of normal vocal function for simple and controlled conditions. Future evaluations will consider pathological cases.
To investigate the relationship between glottal aerodynamics and acoustic loading by means of the proposed inverse filtering scheme.

Experimental platform
A comprehensive set of measurements was undertaken to evaluate the accuracy of the proposed inverse filtering scheme.
Simultaneous and synchronous measurements include: neck skin acceleration (ACC), calibrated laryngeal high-speed videosecondopy (hsv), oral volume velocity (OYY), intramuscular pressure (iOP), electromyography (EGG), and radiated acoustic pressure (MIC).
Vocal gestures included repetitions of the syllables /pa/ and /pi/, different sustained vowels, and pitch glides, each at different loudness and pitch.

Estimating the system impedances
Subglottal impedance (Zsh): Start with a reference of the vocal tract area function and modify for specific human subject via tuning algorithm (Story, 2006). Target formants are obtained during the closed phase portion of the cycle (MIC and OVV).
Subglottal impedance: Start with a male-based subglottal model (Harper et al., 2001) and tune the lengths of trachea to match first resonance in the accelerometer signal (ACC). Target resonance is obtained during the closed phase portion.
Glottal impedance: Use glottal area from calibrated videosecondopy and OOV in a glottal model (Stevens, 1998).

Transfer functions and estimated glottal behavior
Once all impedances are estimated, the circuit model is used to compute the transfer functions needed to predict glottal volume velocity (U0) and its derivative (dU0/dt) from the oral volume velocity (Um) and/or accelerometer (Usk) signals.

Source-filter interactions: voice breaks vs. tract resonances
• To understand better the coupling between the system impedances, a series of vocal gestures that could generate strong source-filter interactions were studied in a normal male speaker.
• Voice breaks were observed as tissue instabilities during pitch glides, and exhibited a strong dependency on the articulated vowel.
• There was a strong correlation between the frequency of prior to the break and the changes in impedance in the supraglottal and subglottal impedance, in agreement with Titze (2008).
• The breaks occurred where 1) Fo crossed F1 and/or 2) when higher harmonics crossed the first supraglottal and subglottal resonances simultaneously.

Conclusions
• The glottal coupling and integration of the subglottal system in the inverse filtering scheme allows for the use of the neck-accelerometer signal to monitor glottal activity.
• Current results suggest that the time-invariant glottal impedance appears to be sufficiently accurate to predict the derivative of the glottal volume velocity from the acceleration signal.
• The effect of the tract impedance needs further attention due to its influence on glottal phonatory stability. These effects could be important in pathological cases.

Acknowledgements
This research was partially funded by the NSF grant number 0828903. Special thanks to Dimitar Deliyski for facilitating use of his laboratory’s high-speed video camera, which was possible through the NIH grant R01 DC007640-01A2.

Cited references