Relationships Between the Cepstral/Spectral Index of Dysphonia and Vocal Fold Vibratory Function During Phonation

Melissa L. Cooke [1, 2]; Daryush D. Mehta [1, 2, 3]; Robert E. Hillman [1, 2, 3]

[1]Institute of Health Professions, Massachusetts General Hospital; [2]Ctr. for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital; [3]Dept. of Surgery, Harvard Medical School

Motivation

There continues to be only limited understanding of fundamental **relationships between vocal fold vibratory function and acoustic characteristics of laryngeal sound production.** Based on recent reports of high correlations between the new acoustic measure of **Cepstral/Spectral Index of Dysphonia (CSID)** and perceptual judgments of dysphonia, we hypothesized that CSID may also correlate better **with HSV-based properties of vocal fold vibratory function,** explaining more of the variance than has been previously reported[2] for other acoustic measures of voice quality.

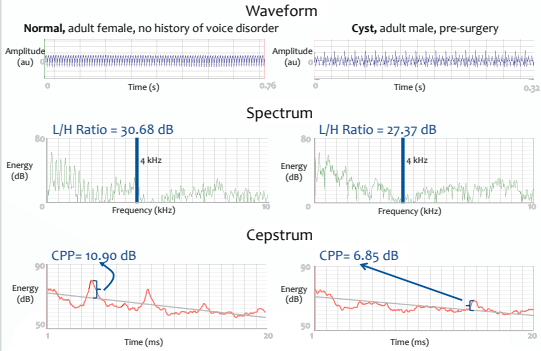
Methods

Acoustic Analysis

The **CSID** and its component variables—**cepstral peak prominence (CPP)** and **low-to-high spectral ratio (L/H Ratio)**—were calculated for **20 subjects** (12 female) who underwent phonomicrosurgery for vocal fold lesions.

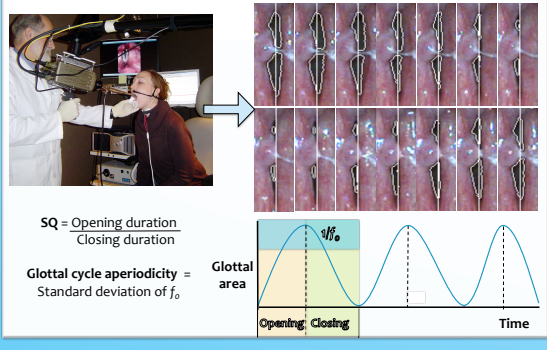
$$\text{CSID} = 84.20 - (4.40 \times \text{CPP}) + (10.62 \times \sigma_{\text{CPP}}) - (1.05 \times \text{L/H Ratio}) + (7.61 \times \sigma_{\text{L/H Ratio}}) - (10.68 \times G)$$ [1]

Analysis of Dysphonia in Speech and Voice (ADSV) Example

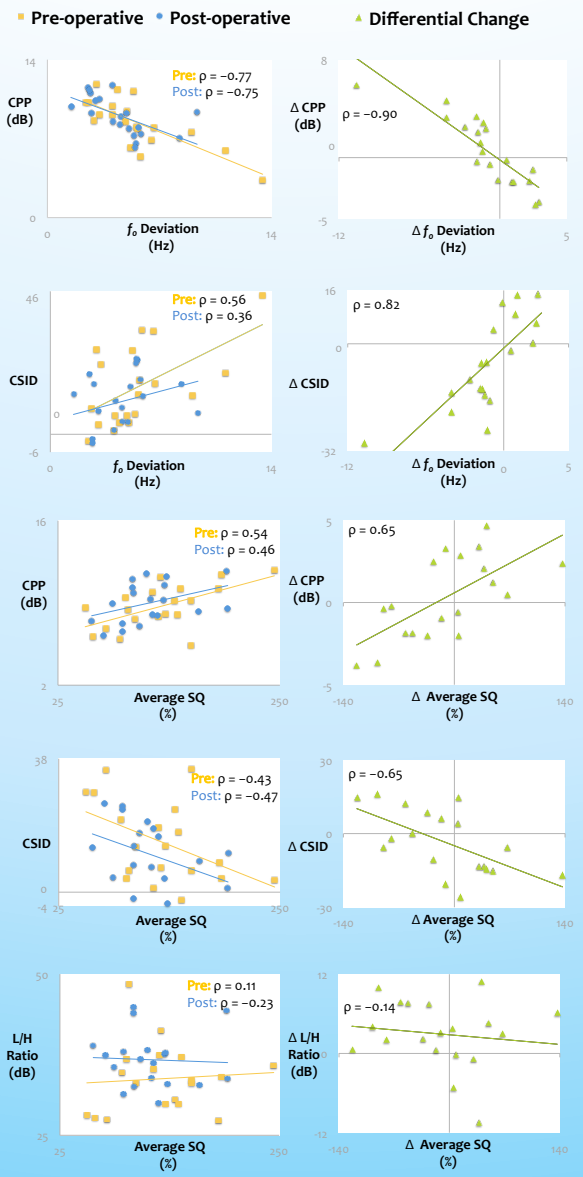


HSV–Based Analysis

Measures of vocal fold vibratory function (irregularity and asymmetry) were extracted from images acquired at 6,250 frames per second, including **glottal cycle aperiodicity (f_o Deviation)** and the mean ratio of opening phase duration to closing phase duration **(Average SQ)** for both **pre-surgical and post-surgical assessments.** Measures of asymmetry including left-right phase asymmetry and amplitude asymmetry did not significantly correlate with CSID or its individual components.



Results



Discussion

CPP correlates highly with glottal cycle aperiodicity

- **Results indicate CSID is only moderately correlated with glottal closure patterns and does not correlate as highly as CPP alone.**
- Results compare well with previous obtained relationships between acoustic CPP and HSV-based measures of voice production.[2]
- As with predicting auditory-perceptual ratings using CSID, CPP has the greatest effect on CSID when predicting vocal fold kinematics. Whereas 59% of the variance in CSID is associated with CPP, the addition of L/H Ratio, $\sigma_{\text{L/H Ratio}}$ and σ_{CPP} to the CSID equation introduces more random variance (noise) instead of explaining more of the variance of the HSV-based measures.

Average SQ and L/H Ratio correlations are not significant

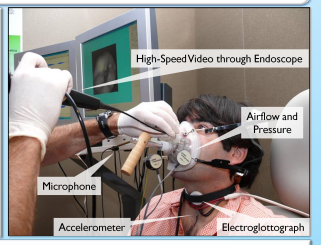
- It was hypothesized that L/H ratio would correlate with average SQ, as L/H ratio reflects harmonic spectral tilt related to the abruptness of glottal closure and the presence of aspiration noise in the higher frequencies.
- Future work is warranted to discriminate the contributions of harmonic and noise energy to L/H Ratio.

Within-subject change increases correlation

- Results corroborate previous evidence[2] that **higher correlations are obtained between acoustic and HSV-based measures when examining the within-subject change** from pre-surgical to post-surgical assessments.
- The variability in f_o and SPL influence the reliability of CPP measurement and is minimized when CPP is tracked within each subject.

Conclusions

Even though there is evidence that CSID may correlate better with perceptual judgments of dysphonia, it does not appear to correlate better with HSV-based measures of vocal fold vibratory function than previously examined acoustic measures. In fact, the CPP component of CSID actually accounts for more of the variance in vocal fold vibratory measures than the combination of measures that comprise CSID.



Future directions aim at improving our understanding of the fundamental relationships between acoustic parameters and underlying phonatory function through the simultaneous acquisition of aerodynamic estimates of glottal airflow and subglottal pressure with high-fidelity imaging and other vocal function properties.

Acknowledgments

This project is supported by grants from the NIH National Institute on Deafness and Other Communication Disorders (R01 DC007640) and by the Christopher Norman Fund. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

References

[1] Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics, 24*(9), 742–58.

[2] Mehta, D. D., Zeitels, S. M., Burns, J. A., Friedman, A. D., Deliyski, D. D., & Hillman, R. E. (2012). High-speed videoendoscopic analysis of relationships between cepstral-based acoustic measures and voice production mechanisms in patients undergoing phonomicrosurgery. *Annals of Otology, Rhinology & Laryngology, 121*(5), 341–7.