Tracking Voice Change after Thyroidectomy: Application of Spectral/Cepstral Analyses

Shaheen N. Awan*, Leah B. Helou+#, Nancy Pearl Solomon* & Alexander Stojadinovic+
*Bloomsburg University of Pennsylvania; +Walter Reed Army Medical Center; #University of Pittsburgh

Introduction

Voice complaints after surgery to remove all or part of the thyroid gland have been reported in a minority of patients, and can have devastating implications for these patients. The most common operative injury involves the recurrent laryngeal nerve, resulting in transient or permanent vocal fold paresis or paralysis. Other complications can involve damage to the external branch of the superior laryngeal nerve, biomechanical arytenoid or laryngeal mucosal injury related to endotracheal intubation, cricothyroid muscle injury, and mucosal congestion/edema. The published literature generally indicates that durable voice problems occur in 1-7% of patients after thyroidectomy, but transient voice problems soon after surgery are far more common (~35%).

Voice evaluations usually include laryngeal imaging and patient report, as well as acoustic analysis. Standard acoustic analysis of voice involves the steady-state portion of a sustained vowel (most commonly /a/), though this limited voice sample may be unrepresentative of the speaker’s typical voice. Several authors have attempted to use spectral-based acoustic methods to analyze normal and disordered voice quality in running speech. In particular, cepstral analysis has shown considerable promise. A number of studies have demonstrated the effectiveness of measures derived from cepstral analysis to quantify dysphonic voice characteristics in sustained vowel and continuous speech (in particular, Hillenbrand and Houde, 1996). Recent work by Awan, Roy & Dromey (2009) has extended the work of Hillenbrand and Houde to include measures of low vs. high frequency spectral energy and the average variability of spectral and cepstral measures in addition to measures of the cepstral peak prominence (CPP).

The purpose of this study was to evaluate possible changes in the perceptual and acoustic characteristics of continuous speech samples pre- and post-thyroidectomy. Because traditional time-based acoustic measures such as jitter and shimmer are not valid with running speech samples, the acoustic analysis procedures used in this study were focused on spectral/cepstral measures.

Method

Participants: Patients scheduled for partial or total thyroidectomy were recruited for this 6-month prospective longitudinal trial. Perceptual and acoustic analyses were conducted for 70 subjects pre-thyroidectomy; 1-4 weeks; 3 months; and 6 months after surgery. Participants included 36 women and 34 men (Mean Age: 51.3 yrs.; Range: 23-78). Subjects were also classified as Normal (n=50) vs. Negative Voice Outcome (NVO; n=20) based upon blinded Consensus Auditory Perceptual Evaluation of Voice (CAPE-V; Kempster et al., 2009), Voice handicap Index (VHI), and Dysphonia Severity Index (DSI) scores at 1-week post-surgery.

Stimuli: Speech samples included the six sentences provided with the CAPE-V. Sentences were read aloud in a typical manner, and were digitized (16 bits, 25 kHz sampling rate) using an AKG C 420 head-set microphone (positioned 4-5 cm from lips) and the KayPentax CSL 4500.

Perceptual Analysis: Three certified SLPs who specialize in voice rated the CAPE-V recordings. Samples were presented quasi-randomly; session data were blocked for each subject. Judges listened under headphones in a
sound-treated booth, and used a custom-automated CAPE-V program. Median ratings for overall severity were obtained at each time point.

Acoustic Analysis: All samples were analyzed using a Windows-based computer program developed by the first author and reported in Awan et al. (2009). Measures of the cepstral peak prominence (CPP) and the ratio of low vs. high frequency spectral energy (L/H Ratio), as well as the standard deviations for the aforementioned measures were obtained for all continuous speech samples.

**Results**

A series of Repeated measures mixed-model ANOVA's (Within subjects factor (Time); Between subjects factors (Gender, and 1-wk post-op voice outcome) were computed to assess possible differences in the median CAPE-V Overall Severity ratings and acoustic measures. Results indicated that the ANOVA for Overall Severity across time was nonsignificant ($F(3,198) = 0.58; p = .63$). In contrast, significant interactions of Time x Voice Outcome were observed for CPP, CPP sd, and the L/H Ratio. Post-hoc analyses using Bonferroni corrected t-tests indicated (a) a significant reduction in mean CPP pre-op to 1-4 wk post-op within the NVO group ($p = .013$) and significant differences between NVO and normal groups at 1-4 wk post-op ($p < .001$), 3-mo post-op ($p = .03$), and 6-mo post-op ($p = .002$); (b) a strong trend for the CPP sd to be reduced 1-4 wk post-op as compared to pre-op or 3-mo post-op for the NVO group, and significant differences between NVO and normal groups at 1-4 wk post-op ($p < .001$), and 6-mo post-op ($p = .015$); (c) a significant increase in the L/H ratio between pre-op and 6-mo post-op for the normal group ($p = .045$) and a significant increase between 1-4 wk post-op and 6-mo post-op for the NVO group ($p = .016$).

**Discussion**

Changes in the overall severity of the speaking voice following thyroidectomy that appeared to be too subtle to detect via auditory-perceptual ratings were observed using spectral/cepstral analysis of continuous speech. Decrements in CPP and CPP SD soon after thyroidectomy are consistent with previous literature indicating that transient voice problems are relatively common after thyroidectomy. These findings were clearly evident for the NVO group, which included patients who were identified with voice problems at the first post-operative visit. Therefore, it appears that CPP and CPP SD are sensitive to changes that perceptual analysis alone may miss. Spectral/cepstral methods appear to provide valuable analyses of running speech samples that cannot be analyzed validly with traditional time-based acoustic measures such as jitter and shimmer during the peri-operative period in patients undergoing thyroidectomy.

**References**


(Disclaimer: The views expressed are those of the authors and do not reflect official policy of the United States Army, Department of Defense, or US Government.)