

# High-Speed Videoendoscopic Analysis of Relationships Between Cepstral-Based Acoustic Measures and Voice Production Mechanisms in Patients Undergoing Phonomicrosurgery

Daryush D. Mehta, PhD; Steven M. Zeitels, MD; James A. Burns, MD;  
Aaron D. Friedman, MD; Dimitar D. Deliyski, PhD; Robert E. Hillman, PhD

**Objectives:** There is increased interest in using cepstral-based acoustic measures for objective clinical voice assessment because of their apparent advantages over more time-honored methods, but there is a paucity of information about how these newer measures relate to underlying phonatory mechanisms.

**Methods:** We investigated the relationships between the acoustic cepstral peak magnitude (CPM) and high-speed videoendoscopy (HSV)-based measures of vocal fold phonatory function in 20 subjects who underwent phonomicrosurgery for vocal fold lesions. Acoustic and imaging data were acquired during sustained vowel phonation before and after surgery.

**Results:** The changes in the measures between presurgical and postsurgical assessments showed that the CPM correlated significantly with an HSV-based measure combining fundamental frequency deviation and average speed quotient ( $r = 0.70$ ;  $p < 0.001$ ) in a multiple linear regression, and that the variation in the CPM could also be attributed to trading relationships between the HSV-based measures of vibratory phase asymmetry and glottal closure.

**Conclusions:** These initial results demonstrate that the clinical utility of cepstral-based measures can be enhanced by a better understanding of how these acoustic measures relate to underlying phonatory mechanisms. The CPM seems to integrate information about aperiodicity in vocal fold vibration, the relative speed of glottal closure, and estimates of glottal noise generation.

**Key Words:** acoustics, cepstrum, high-speed videoendoscopy, laryngology, phonosurgery.

## INTRODUCTION

In recent years, technological advances have made it possible to build high-speed videoendoscopy (HSV) systems that can acquire high-resolution images of vocal fold vibration at significantly faster rates than were previously attainable.<sup>1</sup> We have begun using the new HSV technology in conjunction with time-synchronized microphone recordings to gain a more in-depth understanding of the relationships between vocal fold vibratory parameters and the resulting acoustic characteristics of the voice. A better understanding of such relationships is critical for improving diagnostic methods for voice disorders and for guiding the development of more effective phonosurgical procedures.

In a recent study of patients who had undergone surgery for glottal cancer,<sup>2</sup> we showed that average

HSV-based measures did not correlate significantly with traditional acoustic perturbation measures; however, moderate (statistically significant) correlations were exhibited between the acoustic measures and the standard deviations of the HSV-based parameters of open quotient and vibratory asymmetry. The results corroborated the view that the most important factors contributing to improved vocal quality were adequate glottal closure and periodicity of vocal fold vibration,<sup>3</sup> and we also demonstrated that standard clinical stroboscopy is not capable of revealing the cycle-to-cycle variations in tissue vibratory behavior that were shown to be significantly correlated with the degradation of the acoustic signal.

Even though the types of acoustic perturbation measures that we examined in previous work<sup>2</sup> have

From the Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital (Mehta, Zeitels, Burns, Friedman, Hillman), and the Department of Surgery, Harvard Medical School (Zeitels, Burns, Friedman, Hillman), Boston, Massachusetts, and the Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center (Deliyski), and the Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati (Deliyski), Cincinnati, Ohio. This work was supported in part by the Eugene B. Casey Foundation, the Institute of Laryngology and Voice Restoration, and NIH grant R01 DC007640. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

**Correspondence:** Robert E. Hillman, PhD, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, 1 Bowdoin Square, 11th Floor, Boston, MA 02114.

TABLE 1. SUBJECT DEMOGRAPHICS AND DIAGNOSES

Subject No.	Sex	Age (y)	Diagnosis	Vocal Fold With Disease
1	F	25	Polyp	Right
2	F	27	Cysts	Both
3	F	35	Fibrovascular nodules	Both
4	F	24	Polyp (hemorrhagic)	Left
5	F	51	Polyp (hemorrhagic)	Left
6	F	35	Polyp (hemorrhagic)	Right
7	F	55	Nodules	Both
8	M	50	Polyp	Right
9	F	40	Polyp	Left
10	M	52	Polyp (hemorrhagic)	Left
11	M	43	Cyst	Left
12	F	22	Fibrovascular nodules	Both
13	M	75	Cancer	Right
14	M	53	Polyp (hemorrhagic)	Left
15	M	64	Polyp	Left
16	F	55	Polyp	Right
17	M	61	Polyp	Right
18	M	40	Polyps	Both
19	F	30	Fibrovascular nodules	Both
20	F	18	Polyp	Right

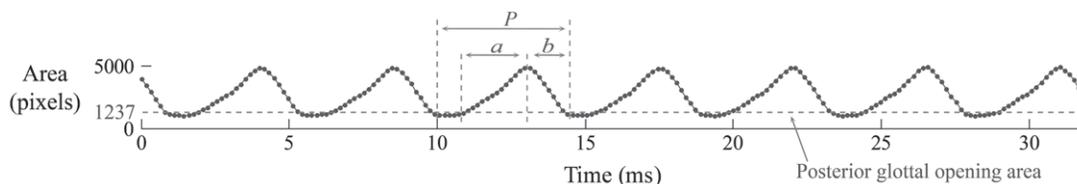
been commonly used in clinical voice assessment, such measures are known to have inherent limitations. The limitations include being restricted to sustained vowel analysis and having a diminished reliability as dysphonia increases; these limitations are mostly due to a reliance on accurate fundamental frequency ( $f_0$ ) detection.<sup>4</sup> Recent research has focused on developing improved acoustic measures that do not rely on an accurate estimate of  $f_0$  and can apply to both sustained vowels and running speech segments. Using a combination of spectral and cepstral information, Awan et al<sup>5</sup> reported a correlation of 0.96 with perceptual ratings of overall dysphonia for sustained vowels and an overall correlation of 0.81 for sentences. In particular, stepwise regression analysis revealed that cepstral-based measures accounted for most of the variance in the perceptual data. Using various versions of the cepstral peak magnitude (CPM), several studies have found strong correlations between the CPM and overall dysphonia perception,<sup>6-8</sup> as well as a strong relationship between the CPM and breathiness in subjects with voice disorders.<sup>9,10</sup> Because of their apparent good correlation with perception and signal processing advantages, measures derived from the acoustic cepstrum are beginning to be adapted for use in clinical voice assessment. However, there is a paucity of information about relationships between vocal fold vibratory parameters and the CPM measures extracted from the resulting voice acoustic signal.

In one study of subjects with organic lesions or unilateral vocal fold paralysis, 38 of 42 individuals exhibited an increase in the CPM during sustained vowel phonation after undergoing phonosurgical treatment.<sup>11</sup> A significant, but weak, correlation existed between the postoperative change in the CPM and changes in acoustic measures of high-frequency noise ( $\rho = -0.41$ ) and jitter ( $\rho = -0.39$ ). The noise and jitter measures, however, did not correlate with one another, suggesting that the changes in the CPM were due to the combined independent effects of noise turbulence at the glottis and aperiodicity of the vocal fold vibration. Acoustic analyses of synthesized vowels have linked cepstral-based acoustic measures to the harmonics-to-noise ratio,<sup>12,13</sup> jitter,<sup>12,13</sup> and spectral tilt,<sup>13</sup> which are meant to be acoustic correlates of vocal fold vibratory patterns.

The purpose of the current study was to elucidate the relationship between cepstral-based acoustic measures and physiological measures of vocal fold vibration derived from laryngeal HSV assessment of subjects before and after phonomicrosurgical treatment. With access to endoscopic images and time-synchronized acoustic recordings, we can begin to address the impact of vocal fold vibratory patterns on traditional and cepstral measures of the acoustic signal. HSV-based parameterization of the glottal area allows for the quantification of cycle-to-cycle glottal closure measures (eg, speed quotient and cycle period) that may be correlated with changes in acoustic measures. It is hypothesized that the cepstral-based CPM measures will 1) correlate positively with the HSV-based measure of speed quotient, which corresponds to more rapid postoperative glottal closure; and 2) correlate negatively with the HSV-based measure of  $f_0$  deviation, which reflects an increase in cycle-to-cycle aperiodicity.

## METHODS

The HSV and acoustic voice data were obtained from 20 subjects (8 male, 12 female) who underwent phonomicrosurgical procedures to treat organic lesions directly affecting the phonatory mucosa. At the time of their postoperative assessment, the average age of the subjects was 43 years (SD, 16 years; range, 18 to 75 years). The HSV assessment was performed during both preoperative and postoperative clinical appointments; the latter were an average of 3.5 weeks after surgery. One subject (patient 13) underwent two surgical procedures before the postoperative HSV assessment. Table 1 lists the sex, age, and diagnosis of each subject. Each subject was evaluated in a sound-treated room with transoral rigid endoscopy and a head-mounted condenser mi-



**Fig 1.** Illustration of parameterization of glottal area waveform (subject 1) to compute high-speed videoendoscopy-based measures of speed quotient and fundamental frequency deviation.  $P$  — period duration;  $a$  — opening phase duration;  $b$  — closing phase duration. Filled circles depict waveform sampling by the high-speed camera.

crophone (MKE104, Sennheiser Electronic GmbH, Wennebostel, Germany) positioned approximately 4 cm from the corner of the subject's mouth. The subjects were instructed to produce the vowel /i/ at a comfortable pitch and loudness for 2 to 4 seconds. The most stable phonatory segment of 320-ms duration was selected for each subject for subsequent analysis.

**Phon microsurgery.** All subjects underwent microlaryngeal surgery under general anesthesia for treatment of their vocal fold lesions using cold instruments<sup>14</sup> and/or 532-nm pulsed potassium titanyl phosphate (KTP) laser photoablation<sup>15</sup> (malignant lesions). In all cases, a pretreatment subepithelial vocal fold infusion of saline solution<sup>16</sup> was used to lift basement membrane lesions away from the superficial lamina propria (nodules, polyps) and to distend and define the existing superficial lamina propria for potentially deeper lesions (cysts, cancer). Pretreatment infusion of saline solution was compared to an immediate posttreatment infusion of saline solution to ensure that the microlayer was maximally preserved. After surgery, all subjects were placed on voice rest (up to 2 weeks, depending on surgeon preference) and received prophylactic anti-reflux medication.

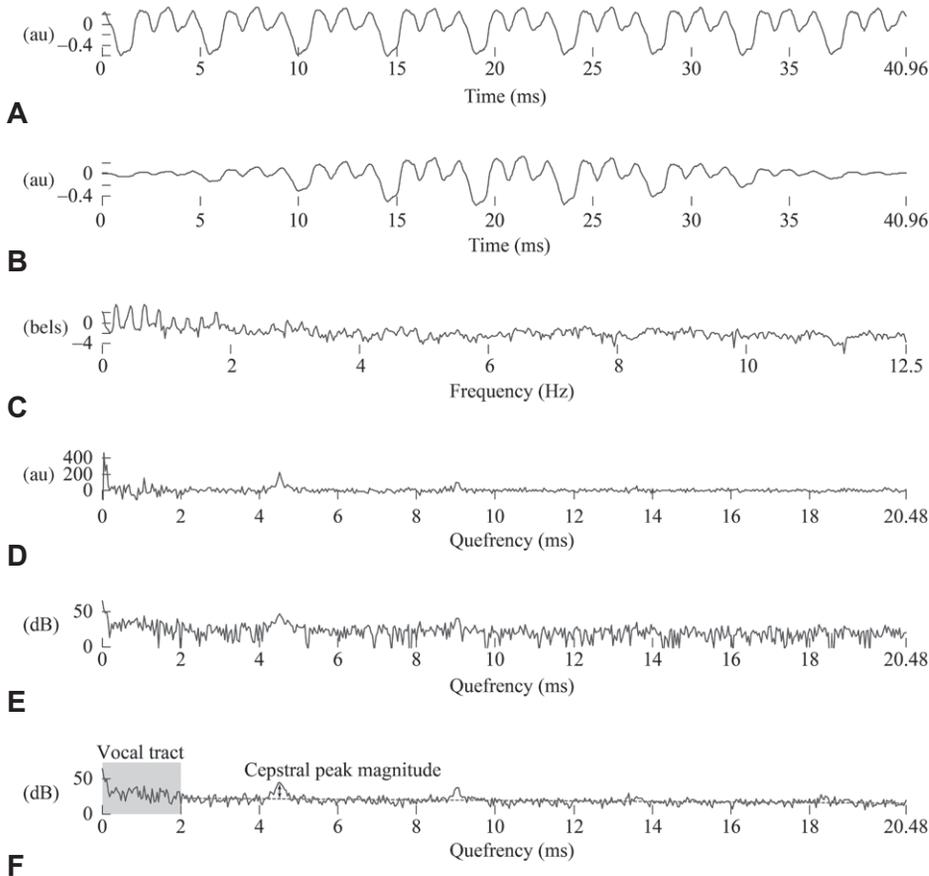
**High-Speed Videoendoscopy.** The HSV data were acquired with the high-speed color video camera (Phantom v7.3, Vision Research Inc, Wayne, New Jersey) used in a previous investigation.<sup>2</sup> The camera lens was coupled to a transoral rigid endoscope and a xenon light source containing 3 glass infrared filters for thermal energy reduction. The video sampling rate was set to 6,250 images per second with maximum integration time, and the spatial resolution was 320 horizontal  $\times$  352 vertical pixels (approximately 1.5 cm<sup>2</sup>). Time synchronization of the HSV data and the acoustic signal was enabled by a common high-frequency clock generated by the data acquisition board, whose settings were controlled in software (MiDAS DA, Xcitex Corporation, Cambridge, Massachusetts). A digital signal output from the camera was used to mark the time of a recorded image to within 11  $\mu$ s.

Properties of the vocal fold vibratory pattern were

derived from the glottal area waveforms and lateral displacement waveforms that were automatically extracted from the HSV recordings by custom digital image processing routines created in MATLAB (The MathWorks, Inc, Natick, Massachusetts). Although not spatially calibrated to physical units in the current setup, the normalized waveforms provided for the quantification of relative measures of cycle-to-cycle behavior. Intensity-based threshold detection was used to obtain the glottal area, in pixels, from successive images.<sup>17</sup> The measures of vocal fold vibratory asymmetry reflected kinematics at the mid-musculomembranous glottis and included left-right phase asymmetry, left-right amplitude asymmetry, and axis shift during glottal closure. The details of obtaining these asymmetry-related measures can be found in previous articles.<sup>2,17</sup>

Figure 1 illustrates the parameterization of the glottal area waveform to determine cycle-to-cycle measures of the speed quotient and  $f_0$  deviation. The speed quotient was defined as the ratio, in percent, between the opening phase duration  $a$  and the closing phase duration  $b$ . Larger values of the speed quotient reflect a skewed glottal area shape due to a relatively more rapid glottal closing phase. After taking into account any constant area offset due to posterior glottal openings during closure, we defined the opening and closing phase durations on either side of the waveform peak by counting the number of images during the opening and closing phases, respectively. The  $f_0$  deviation was computed, in hertz, as the standard deviation of the reciprocal of period  $P$  over all cycles. The period was derived from the number of images between successive instances of glottal closure.

**Acoustic Measures.** Because of the common clock signal, the acoustic sampling rate of the microphone signal was always an integer multiple of the video sampling rate. The acoustic sampling rate was thus set at 100 kHz (16 samples per image), and the 3-dB cutoff frequency of the antialiasing filter was set at 30 kHz. The microphone signal was shifted 600  $\mu$ s into the past relative to the HSV data to compensate approximately for the larynx-to-microphone acoustic propagation time.



**Fig 2.** Illustration of intermediate processing for computing cepstral peak magnitude. See text for corresponding equations. au — linear arbitrary units. **A)** Waveform. **B)** Windowed waveform. **C)** Power spectrum. **D)** Real cepstrum. **E)** Power cepstrum. **F)** Averaged power cepstrum.

The standard acoustic perturbation measures of shimmer and jitter were computed from the synchronously recorded acoustic voice signal by use of the Multi-Dimensional Voice Program (MDVP, model 5105, KayPENTAX, Montvale, New Jersey). An estimate of the harmonics-to-noise ratio was obtained with Praat.<sup>18</sup>

In addition, the cepstral-based measure CPM was computed from the segment. Because the cepstrum can be calculated in many different ways, we selected a procedure reported in a recent study on cepstral correlates of perception.<sup>5</sup> After resampling of the waveform to 25 kHz, the signal was windowed into 40.96-ms frames (1,024 samples) with 75% overlap. After multiplication by a Hamming window, two 1,024-point discrete Fourier transforms (DFTs) were computed in succession with a logarithmic transformation between them. Figure 2 illustrates the intermediate processing applied to each frame  $x[n]$  (Fig 2A) to yield the CPM, in decibels:

Definition of DFT:  $X[k] = \sum_{n=0}^{N-1} x_w[n] e^{-j2\pi kn/N}$ ,  $0 \leq k \leq N-1$

Power spectrum (Fig 2C):  $X_p[k] = \log_{10}|X[k]|^2$

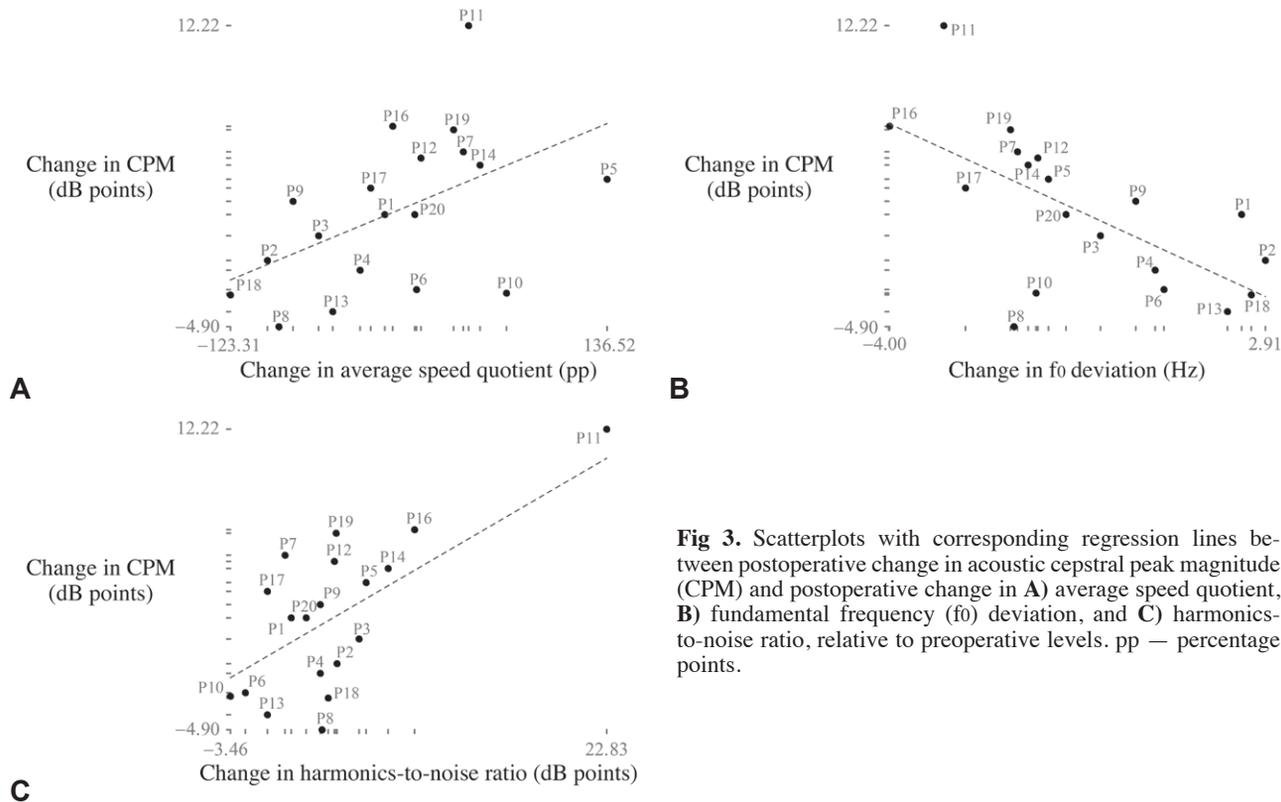
Real cepstrum (Fig 2D):  $\hat{x}[q] = \sum_{k=0}^{N-1} X_p[k] e^{-j2\pi qk/N}$ ,  $0 \leq q \leq N-1$

Power cepstrum (Fig 2E):  $\hat{x}_p[q] = 10 \log_{10} |\hat{x}[q]|^2$

where  $x_w[n]$  is the Hamming-windowed frame (Fig 2B),  $X_p[k]$  is the power spectrum in bels,  $\hat{x}[q]$  is the real cepstrum, and  $\hat{x}_p[q]$  is the power cepstrum in decibels. The discrete indices for time, frequency, and quefrency are  $n$ ,  $k$ , and  $q$ , respectively.

A 7-frame cepstral smoothing was performed by averaging the power cepstrum with those of the 3 frames before and after a given frame. Because of a bias in the decibel power cepstrum, a regression line was computed over quefrencies greater than 2 ms, which corresponded to a quefrency range assumed to be minimally affected by vocal tract-related information. Finally, the CPM was defined as the difference, in decibels, between the magnitude of the highest peak and the baseline regression level in the averaged power cepstrum (Fig 2F). The peak search was limited to quefrencies between 3.3 ms and 16.7 ms, corresponding to  $f_0$  values of 83 Hz and 417 Hz, respectively.

**Statistical Analysis.** The subject data were divided into three groups: preoperative measures, postoperative measures, and the differential change in measures from preoperative to postoperative assessment. The averages and standard deviations of the HSV-based measures were taken over all cycles in a subject's phonatory segment. Acoustic measures



**Fig 3.** Scatterplots with corresponding regression lines between postoperative change in acoustic cepstral peak magnitude (CPM) and postoperative change in **A)** average speed quotient, **B)** fundamental frequency ( $f_0$ ) deviation, and **C)** harmonics-to-noise ratio, relative to preoperative levels. pp — percentage points.

of perturbation and harmonics-to-noise ratio were computed from the time-synchronized acoustic segment. Because of the limited number of samples in the three groups, the Spearman rank-order correlation coefficient  $\rho$  was used to determine any significant pairwise relationships between HSV-based measures and acoustic measures. A Spearman statistic was considered significant if the  $p$  value was less than 0.05. In addition, linear multiple regression was used to further determine whether combining 2 or more measures could be used to improve predictive power. Scatterplots were generated for visual assessment of marginal and joint distributions of the measures of interest.

## RESULTS

For the preoperative data, statistically significant pairwise correlations were found between the CPM and two HSV-based measures: average speed quotient ( $\rho = 0.55$ ;  $p = 0.016$ ) and  $f_0$  deviation ( $\rho = -0.57$ ;  $p = 0.012$ ). An indirect relationship between the CPM and glottal noise was obtained by correlating the CPM with the acoustic-based harmonics-to-noise ratio ( $\rho = 0.68$ ;  $p = 0.001$ ). Acoustic jitter correlated significantly with HSV-based measures of the standard deviation of left-right phase asymmetry ( $\rho = 0.54$ ;  $p = 0.023$ ) and the standard deviation of left-right amplitude asymmetry ( $\rho = 0.56$ ;  $p = 0.015$ ).

In the postoperative data, acoustic jitter only correlated significantly with the standard deviation of left-right phase asymmetry ( $\rho = 0.50$ ;  $p = 0.031$ ). The CPM only correlated significantly with the acoustic-based harmonics-to-noise ratio ( $\rho = 0.57$ ;  $p = 0.009$ ), and not with any HSV-based measure.

For the differential change in measures from the preoperative to the postoperative assessments, significant correlations were exhibited between the CPM and the average speed quotient ( $\rho = 0.56$ ;  $p = 0.015$ ) and between the CPM and the  $f_0$  deviation ( $\rho = -0.63$ ;  $p = 0.009$ ). Combining the two HSV-based measures in a parametric multiple regression yielded a correlation ( $r$ ) with the CPM of 0.70 ( $p < 0.001$ ). Adding the acoustic-based harmonics-to-noise ratio to the regression raised this correlation to 0.84. Figure 3 displays scatterplots of the change in the CPM versus the three measures in this multiple regression. Table 2 displays raw values of these acoustic and HSV-based measures for the preoperative and postoperative assessments.

## DISCUSSION

This study documented the first attempt at defining relationships between physiological measures derived from HSV assessment of vocal fold vibration and acoustic cepstral-based measures before and after phonosurgery. A better understanding of such relationships is critical to improving di-

TABLE 2. PREOPERATIVE AND POSTOPERATIVE ACOUSTIC MEASURES OF CPM AND HNR AND HSV-BASED MEASURES OF AVERAGE SQ AND  $f_0$  DEVIATION

Subject No.	CPM (dB)		HNR (dB)		Average SQ (%)		$f_0$ Deviation (Hz)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1	23.3	24.8	19.6	16.7	149.4	132.5	2.43	4.91
2	22.4	21.2	19.7	21.5	188.1	90.4	2.60	5.51
3	20.2	20.5	20.2	19.3	133.7	71.4	5.59	5.47
4	22.0	20.3	20.4	18.3	93.3	59.2	4.97	5.85
5	18.8	22.3	16.8	20.4	60.6	197.1	5.21	4.13
6	24.7	21.9	22.6	17.7	121.0	126.0	4.06	5.10
7	21.6	26.6	18.9	18.1	96.1	133.4	4.85	3.22
8	28.3	23.4	17.7	17.8	190.3	100.5	4.38	2.68
9	22.2	24.4	23.7	29.9	160.1	80.1	4.07	4.60
10	27.4	24.4	18.0	17.3	129.0	196.3	5.34	4.03
11	13.5	25.7	19.5	17.9	73.6	114.5	6.50	2.50
12	15.0	19.7	15.0	25.4	160.0	167.8	5.80	4.53
13	22.6	18.5	20.9	23.7	143.2	90.2	3.21	5.42
14	17.7	22.0	15.0	19.2	52.6	101.5	2.91	1.46
15	13.9	20.6	9.9	17.3	*	92.6	*	2.76
16	17.8	24.2	15.3	18.1	146.7	135.2	6.65	2.66
17	19.8	22.7	17.6	23.9	126.8	100.4	5.62	3.02
18	25.3	22.2	18.6	18.8	244.4	121.1	3.05	5.71
19	18.3	24.6	19.5	24.7	88.3	119.3	11.12	9.35
20	20.8	22.3	19.7	21.1	103.1	107.1	9.02	8.26

CPM — cepstral peak magnitude; HNR — harmonics-to-noise ratio; HSV — high-speed videoendoscopy; SQ — speed quotient;  $f_0$  — fundamental frequency.

\*Preoperative HSV-based measures for subject 15 were not applicable because of period irregularities and biphonation.

agnostic methods for voice disorders and to guiding the development of more effective phonosurgical procedures.

It was hypothesized that the CPM would be affected by changes in the HSV-based measures of the aperiodicity and the speed quotient of vocal fold vibration. In addition, even though a direct measurement of glottal turbulent noise was not possible with the current recording setup, its potential contribution to the CPM measure was estimated by use of the acoustic harmonics-to-noise ratio.

The results of this study lend empirical support to the idea that the magnitude of the first cepstral peak in the acoustic signal — ignoring low-frequency, vocal tract-related information — reflects contributions from multiple features of the glottal vibratory sound source. A significant portion of the variance in the postoperative change in the CPM ( $r^2 = 0.49$ ) was explained by combining the HSV-based measures of average speed quotient and  $f_0$  deviation. The sensitivity of the CPM to the speed of glottal closure was expected; the increases in the speed quotient (more rapid glottal closure in terms of glottal area) signifi-

TABLE 3. POSTOPERATIVE CHANGE IN CPM AND PA DEVIATION AND VISUAL VERIFICATION OF COMPLETE OR INCOMPLETE GLOTTAL CLOSURE DURING CLOSED PHASE

Subject No.	Change in CPM (dB Points)	Change in PA Deviation (pp)	Complete Closure?
1	1.48	0.211	Yes
2	-1.16	-0.504	No
3	0.25	-0.503	Yes
4	-1.69	-0.686	No
5	3.47	0.193	Yes
6	-2.82	-5.938	No
7	5.03	-1.687	Yes
8	-4.90	-0.071	No
9	2.23	-0.186	Yes
10	-3.00	0.008	Yes
11	12.22	0.220	Yes
12	4.69	0.706	No
13	-4.09	-1.214	No
14	4.25	-0.460	Yes
15	6.70	-9.430	Yes
16	6.45	0.179	No
17	2.96	-1.905	Yes
18	-3.14	0.686	Yes
19	6.28	-0.712	Yes
20	1.45	-0.851	Yes

PA — left-right phase asymmetry; pp — percentage points.

cantly correlated with increases in the CPM. Adding an estimate of the harmonics-to-noise ratio to the multiple regression increased the explained variance to 71%. The impact of glottal turbulent noise on the CPM could be better defined in future work by obtaining simultaneous estimates of the unmodulated component of the glottal airflow.<sup>19</sup>

Previously documented correlations between the standard deviations of asymmetry measures and acoustic perturbation<sup>2</sup> were corroborated by the data from the preoperative group of the current study. Furthermore, postoperative decreases in the standard deviation of left-right phase asymmetry were hypothesized to accompany postoperative increases in the CPM due to the attenuation of vocal fold aperiodicity, and vice versa. In practice (Table 3), we observed this expected relationship for only 10 of the 20 subjects when we examined the change in measures from preoperative to postoperative assessments. In 5 subjects, the CPM decreased even though there was a decrease in the standard deviation of the left-right phase asymmetry. Case-by-case qualitative examination of the HSV recordings showed incomplete glottal closure during the closed phase in these 5 cases. A possible trading relation was thus exhibited, whereby postoperative changes in the CPM appear to be effected by a balance between the extent to which glottal closure was com-

plete (with associated variation in turbulent noise generation) and the degree of vocal fold aperiodicity. In 3 subjects, better postoperative glottal closure appears to have overridden minor perturbation in terms of improving the CPM. On the other hand, 2 subjects exhibited postoperative increases in the CPM due, in part, to improved postoperative vocal fold entrainment, even though the glottis was not closing completely during phonation.

Even though statistically significant correlations were found between two HSV-based measures and the CPM, the strengths of the correlations were moderate at best. A substantial amount of the variation in the CPM was still not accounted for by the HSV measures computed. All of the usual caveats of using HSV assessments apply to the current study, including limited spatial resolution, 2-dimensional image capture, and endoscopic motion compensation.<sup>1</sup> Further investigation is thus warranted to explore these relationships in a larger subject sample with the addition of glottal airflow measurement

and, ultimately, the development of methods for extracting information about the 3-dimensional motion of the vocal folds during phonation.

## CONCLUSIONS

Even though cepstral-based measures are being adapted for clinical voice assessment, there is a paucity of information about how these measures are related to the underlying physiology of vocal fold vibration. The current study used time-synchronized HSV and microphone recordings for 20 patients before and after phonosurgery to begin defining these relationships. Our initial results indicate that the CPM seems to integrate information about aperiodicity in vocal fold vibration, the relative speed of glottal closure, and estimates of glottal noise generation. These initial results demonstrate that the clinical utility of cepstral-based measures can be enhanced by a better understanding of how these acoustic measures relate to underlying phonatory mechanisms.

## REFERENCES

1. Deliyski DD, Petrushev PP, Bonilha HS, Gerlach TT, Martin-Harris B, Hillman RE. Clinical implementation of laryngeal high-speed videoendoscopy: challenges and evolution. *Folia Phoniatri Logop* 2008;60:33-44.
2. 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.
3. Woo P, Casper J, Colton R, Brewer D. Aerodynamic and stroboscopic findings before and after microlaryngeal phonosurgery. *J Voice* 1994;8:186-94.
4. Mehta DD, Hillman RE. Voice assessment: updates on perceptual, acoustic, aerodynamic, and endoscopic imaging methods. *Curr Opin Otolaryngol Head Neck Surg* 2008;16:211-5.
5. Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgments from the CAPE-V. *Clin Linguist Phon* 2010;24:742-58.
6. Heman-Ackah YD, Heuer RJ, Michael DD, et al. Cepstral peak prominence: a more reliable measure of dysphonia. *Ann Otol Rhinol Laryngol* 2003;112:324-33.
7. Dejonckere PH, Lebacqz J. Acoustic, perceptual, aerodynamic and anatomical correlations in voice pathology. *ORL J Otorhinolaryngol Relat Spec* 1996;58:326-32.
8. Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice* 2010;24:540-55.
9. Heman-Ackah YD, Michael DD, Goding GS Jr. The relationship between cepstral peak prominence and selected parameters of dysphonia. *J Voice* 2002;16:20-7.
10. Hillenbrand J, Houde RA. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J Speech Hear Res* 1996;39:311-21.
11. Dejonckere PH, Wieneke GH. Spectral, cepstral and aperiodicity characteristics of pathological voices before and after phonosurgical treatment. *Clin Linguist Phon* 1994;8:161-9.
12. de Krom G. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *J Speech Hear Res* 1993;36:254-66.
13. Murphy PJ. On first harmonic amplitude in the analysis of synthesized aperiodic voice signals. *J Acoust Soc Am* 2006;120:2896-907.
14. Zeitels SM, Hillman RE, Desloge R, Mauri M, Doyle PB. Phonosurgery in singers and performing artists: treatment outcomes, management theories, and future directions. *Ann Otol Rhinol Laryngol Suppl* 2002;111(suppl 190):21-40.
15. Zeitels SM, Burns JA, Lopez-Guerra G, Anderson RR, Hillman RE. Photoangiolytic laser treatment of early glottic cancer: a new management strategy. *Ann Otol Rhinol Laryngol Suppl* 2008;117(suppl 199):1-24.
16. Zeitels SM, Vaughan CW. A submucosal true vocal fold infusion needle. *Otolaryngol Head Neck Surg* 1991;105:478-9.
17. 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.
18. Boersma P, Weenink D. Praat: Doing phonetics by computer [Computer Program]. Version 5.1.07. Amsterdam, the Netherlands: University of Amsterdam, 2009. <http://www.fon.hum.uva.nl/praat/>.
19. Holmberg EB, Hillman RE, Perkell JS, Guiod PC, Goldman SL. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *J Speech Hear Res* 1995;38:1212-23.