MEASURES OF THE GLOTTAL AIRFLOW WAVEFORM, EGG, AND ACOUSTIC SPECTRAL SLOPE FOR FEMALE VOICE

Eva B. Holmberg, Robert E. Hillman, Joseph S. Perkell
Massachusetts Eye and Ear Infirmary and Massachusetts Institute of Technology

ABSTRACT
Comparisons were made among aerodynamic, electroglottographic, and acoustic spectral measures for syllable production and sustained vowel phonation in comfortable and loud voice of 20 women with normal voices. Measures differed significantly between tokens having harmonic energy versus noise in the F3 region. Spectral measures added useful information to glottal waveform data about abrupt versus gradual vocal fold closing.

INTRODUCTION
The objective of this study was to examine whether acoustic spectral measures of voice production could be used to supplement measures obtained from the inverse filtered oral flow waveform that are especially sensitive to technical difficulties [1]. We focused on measures that have been found salient for vocal intensity [2] and glottal aperture [3]. Our flow-based inverse filtering technique is not without problems: 1) An accidental air leak between the subject's face and the transducer mask lowers amplitude-based flow measures that include the DC flow; 2) Unsuccessful inverse filtering that results in formant residual superimposed on the glottal waveform can make time-based measures unreliable; 3) Low-pass filtering, used in our inverse filtering algorithm, has the undesirable effect of rounding of waveform discontinuities, such as at the instant of vocal fold closure.

Methods
Detailed descriptions of recording procedures, signal processing, data extraction and analyses procedures are presented in previous publications [7, 8]. In brief: twenty American females, age 20 to 43 years, with healthy voices served as subjects. They produced two different speech tasks in comfortable and loud voice: (1) strings of repeated /pa:/ syllables and sustained phonation of /æ/ vowels. Our intention is to combine measures from the two speech tasks.

METHODS
Detailed descriptions of recording procedures, signal processing, data extraction and analyses procedures are presented in previous publications [7, 8]. In brief: twenty American females, age 20 to 43 years, with healthy voices served as subjects. They produced two different speech tasks in comfortable and loud voice: (1) strings of repeated /pa:/ syllables and sustained phonation of /æ/ vowels. Our intention is to combine measures from the two speech tasks.

RESULTS
SPL was higher for the vowel in the syllable strings than in the sustained phonation. Analysis of covariance (ANCOVA, p<0.05, Bonferroni corrections, p=0.0045) with SPL as the covariate showed that there were no significant differences in other parameters between /pa:/ and /æ/ productions after adjustment for SPL.

Relationships between Flow- and EGG-Adduction Quotients
Pearson product moment correlations calculated between the flow- and EGG-adduction quotients showed strong relationships (r>0.70) for individual speakers, whose signals were strong and noise free. The results suggest that the quotients measured at the amplitude levels of 30% (flow) and 65% (EGG) were highly related.

Relationships between Glottal Airflow Measures and the Spectral Slope
Pearson product moment correlations between measurements of glottal airflow waveforms and spectral slope showed a relatively strong relationship between flow-adduction quotient and AH1-AH2 (r=0.69). The results suggest that the degree to which the glottal waveform has a sinusoidal shape, and inversely, the degree of glottal adduction, was reflected relatively well in AH1-AH2. AH1-AH2 was also relatively strongly correlated with SPL (r=0.69). The relationship between flow-adduction quotient and AF1-AF3 was significant for tokens with predominantly F3 noise, but not significant for tokens with predominantly F3 harmonic energy. Other relationships were weak in the group data.

The relationships between acoustic spectral measures and glottal airflow measures were examined also for each individual speaker. A majority of the individual speakers displayed strong relationships (r>0.70) between flow-adduction quotient and all the spectral measures, with the exception of AF1-AF3. The data suggest that the degree to which the glottal waveform had a sinusoidal shape was reflected in ratios that included the amplitude of the fundamental, but not in the ratio which included only the higher frequency region.
Differences in Acoustic and Glottal Waveform Measures between Tokens with F3 Spectral Energy and F3 Noise.

Analyses of variance (p<0.001) showed that tokens with predominantly noise in the F3 region were associated with significantly lower SPLs; larger values of AH1-AH2, AH1-AF3, and AF1-AF3; smaller values of AH1-AF1; lower subglottal air pressure; smaller adduction quotients (flow and EGG); and higher DC flow.

**DISCUSSION**

The finding of higher SPL for /ae/ in the /pa:/ syllable strings than for /ae/ in sustained phonation was most likely due to differences in location of the mid-vowel data extraction; in the syllables the mid-vowel point occurs shortly after initiation of vocalization, while in sustained phonation mid-vowel occurs well into the vowel, at a point where SPL was stabilized and somewhat lower than at the beginning of the sustained vowel.

The literature has suggested that the EGG waveform contains a number of interesting features and events that could be useful for a better understanding of the underlying vocal-fold vibration pattern [6]. However, we seldom find such clear events in the EGG waveforms. In addition, we have experienced difficulties in recording EGG, for example weak and noisy signals for women, and intermittent disruptions of the signal due to gross movements of larynx that accompany changes in vocal effort. However, for speakers with strong EGG signals and clean glottal waveforms (without formant residuals), the flow and EGG quotients, measured at 30% and 65% amplitude criteria levels respectively, were highly correlated (r=0.85). This finding suggests that quotients from clean samples from one signal can complement the other if necessary.

A particular focus of this study was relationships between flow and acoustic parameters that have been found reflective of glottal aperture, with the goal of cross validating these measures. Gradual closing movements of a somewhat abducted membranous portion of the vocal folds should result in relatively sinusoidal glottal waveforms and small adduction quotients [3]. The result of relatively strong (negative) correlation between flow-adduction quotient and AH1-AH2 suggests that measurements of adduction quotient at the 30% level criterion was sensitive enough to differentiate among waveforms with gradual and abrupt closing, and that AH1-AH2 could be used as a substitute for adduction quotient in case of unsuccessful inverse filtering. The strong (negative) relationships between flow-adduction quotient on the one hand and AH1-AF1, AH1-AF3, and SPL on the other for the individual speakers suggest that gradual vocal fold closures resulted in an increased amplitude of the first harmonic, reduced amplitude of the first formant, steeper overall spectral slope, and reduced SPL in agreement with previous research [2].

Ideally, high vocal fold closing velocities and abrupt reduction of the airflow should result in glottal waveforms with sharp corners between the closing and closed portions. However, a detrimental effect of low-pass filtering at 1100 Hz is a “rounding” of waveform discontinuities, which could have an influence on waveforms associated with high vocal fold closing velocities. Waveforms which result from more gradual closing movements (therefore already rounded) would be relatively uninfluenced by the low-pass filtering rounding effect. These waveforms, with F3 excited by noise, have a significant relationship between flow-adduction quotient and AF1-AF3. In contrast, waveforms with F3 excited by harmonic energy have a non-significant relationship between flow-adduction quotient and AF1-AF3. The filter-induced rounding of waveforms with more abrupt closures may account for the lack of significant correlation between adduction quotient and AF1-AF3. In other words, the sharp corners in those waveforms were obscured, and reliable adduction quotient measurements were precluded by the low-pass filtering at 1100 Hz. However, the effects of the sharp discontinuities were preserved in the AF1-AF3 value derived from the full-bandwidth acoustic spectra. These results suggest that the spectral measurement of AF1-AF3 may serve as a useful complement to the flow-based adduction quotient, especially when there are high vocal fold closing velocities.

In normal phonation, DC flow in the glottal waveform has been assumed to reflect airflow that passes through a posterior glottal “chink”. A large chink would increase the subglottal coupling with reduced high frequency energy and reduced SPL. However, neither the relationship between underlying physiology and the DC flow [11], nor the acoustic effect of the DC flow is completely understood. None of the acoustic measures varied systematically with DC flow. Thus, DC flow data must be interpreted with caution.

**CONCLUSIONS**

The following conclusions could be drawn from the results of this study:

1) Comparisons between measures obtained from the vowel in /pae/ syllables and those obtained from sustained /ae/ phonation can be made, as long as SPL differences are controlled for.

2) Adduction quotient, measured at a 30% level on the glottal waveform is sensitive enough to differentiate among waveforms with gradual and abrupt closing portions in data for individual subjects.

3) Measurements of the amplitude difference between the two first harmonics (AH1-AH2) may be used as a substitute measure for flow-adduction quotient, in cases of unsuccessful inverse filtering that make measurements of adduction quotient unreliable.

4) The flow- and EGG adduction quotients, measured at 30% and 65% levels respectively, may serve to complement one another.

5) AF1-AF3 may serve as a useful complement to measurements of maximum flow declination rate, especially in voices with very high closing vocal fold velocities that cannot be reflected accurately in a flow waveform that is low-pass filtered at 1100 Hz.

**ACKNOWLEDGMENTS**

This study was supported by a grant from The National Institutes of Health.

**REFERENCES**


