Influence of Glottal Fry on Acoustic Voice Assessment: A Preliminary Study

Laura W. Plexico and Mary J. Sandage, *Auburn, Alabama*

Summary: Purpose. This preliminary study examined the influence of glottal fry on measurement of speaking fundamental frequency (SF_0) and whether sex differences influence the impact glottal fry has on SF_0 and cepstral peak prominence (CPP). It was hypothesized that SF_0 and CPP would decrease as percent glottal fry in the sample increased, with larger changes observed in the female voice.

Methods. Twenty-six participants (13 men, 13 women) completed the study. Participants were recorded reading the Rainbow Passage. SF_0 and CPP were determined from the original sample. Percent glottal fry SF_0 was determined and semitone change was recalculated after removal of glottal fry from the sample. Regression analysis was used to determine the impact of glottal fry and sex on semitone change and CPP differences in SF_0.

Results. Significant differences were found for both sex and percent glottal fry on semitone change in SF_0. A significant relationship was not found between CPP and semitone change when sex was accounted for.

Conclusions. Findings from this study indicate that the measurement of SF_0 for women will decline as percent of glottal fry increases. These findings have clinical implications for interpretation of SF_0 measurement and evidence-based outcomes for voice therapy.

Key Words: Glottal fry–Speaking fundamental frequency–Cepstral peak prominence–Assessment–Vocal register.

INTRODUCTION

Glottal fry, also referred to as vocal fry or "creaky" voice, is an increasingly prevalent vocal characteristic. It is the lowest of three perceptually distinct vocal registers and is characterized by a unique vibratory pattern that is associated with a significantly reduced airflow rate,[1,2] reduced subglottal air pressure,[1,3] a single or multiple opening and closing pattern,[1,4–6] and a lower fundamental frequency when compared to the modal register.[4,7] Glottal fry is produced with a similar fundamental frequency irrespective of gender that falls approximately one octave below a male and two octaves below a female speaker's modal fundamental frequency.[8] Glottal fry typically ranges between 15 Hz and 100 Hz, with a mean of approximately 50 Hz.[1,4,8–10]

Recently, glottal fry has been observed with greater frequency in typical populations. Historically, it was considered a disordered form of phonation because characteristically aberrant vocal qualities such as hoarseness, harshness, and roughness are often associated with persistent glottal fry.[11–13] Hollien et al cautioned against considering sporadic glottal fry as a sign of pathology.[7] They indicated that sporadic glottal fry is perceived in typical voice use and is a distinct mode of laryngeal vibration (ie, register) that speakers without vocal pathology can switch out of for communicative purposes such as boundary marking. This can become clinically relevant when accounting for suprasegmental, dialectical, and language differences.[14,15] The critical distinction between pathological use and nonpathological use of glottal fry may be observed in the degree to which it is persistent or sporadic[7,16] or the overall frequency with which it occurs during discourse. Other factors including speaker discomfort as a result of an inefficient vocal technique or social factors like listener distraction resulting in communication breakdown might also be relevant when determining whether glottal fry is indicative of a communication disorder. Hollien et al characterized glottal fry as a phonational register represented by low frequencies.[7] The extent to which percentage of glottal fry in speech influences clinical determination of voice disorder is not yet well described.

Glottal fry has been found to be more prevalent in female speakers than in male speakers of American and Canadian English.[16–20] Wolk and colleagues[21] evaluated the prevalence of glottal fry in 34 female college students and found that two-thirds of the participants in their study used glottal fry. Glottal fry has been described as being characterized by a creaky and rough voice,[22] qualities of which are often associated with perception of voice disorder particularly for occupational voice users. Although it is often asserted that the phenomenon of glottal fry is more prevalent in younger women, Oliveira et al evaluated young and middle-aged women for differences in the amount of glottal fry produced and found no differences between the age groups.[23] However, they did find that glottal fry was more prevalent in the final and medial position of phrases. Previous research indicates that glottal fry was associated with a higher social status[24] and has been adopted by female speakers in an effort to appear assertive, masculine, or in a position of authority or power.[18,19,25] However, the perception of glottal fry has shifted from one of perceived authority to one with negative associations, with some studies reporting associations with indecisiveness, emotionality, laziness, and ditziness.[17,25]

Further investigation is needed to understand how the extent of glottal fry use, as measured by percent glottal fry in a given speech sample, influences acoustic assessment of voice, specifically measurement of speaking fundamental frequency (SF_0) and cepstral peak prominence (CPP), metrics that can be used in a clinical setting to describe the frequency of the voice and capture the degree to which the signal is periodic or aperiodic.[26,27] Clinically, it is expected that the presence of glottal fry in a functional speech task would decrease SF_0 and a low amplitude CPP would

Accepted for publication October 6, 2016.

From the Department of Communication Disorders, Auburn University, Auburn, Alabama.

Address correspondence and reprint requests to Laura W. Plexico, Department of Communication Disorders, Auburn University, 1199 Haley Center, Auburn, AL 36849. E-mail: lwp0002@auburn.edu

Journal of Voice, Vol. 31, No. 3, pp. 378.e13–378.e17

0892-1997

Published by Elsevier Inc. on behalf of The Voice Foundation.

http://dx.doi.org/10.1016/j.jvoice.2016.10.004

be observed.[28] The purpose of this study was to examine the relationship between glottal fry, SF_0, and CPP. Specifically, the study was designed to address the following questions:

(1) What is the incremental impact of glottal fry on SF_0?; (2) Do males and females differ in the impact glottal fry has on SF_0?; and (3) Is there a significant relationship between glottal fry and CPP? It was hypothesized that SF_0 would decrease as percent glottal fry in the sample increased, with larger changes observed in the female than in the male voice. It was also hypothesized that there would be a significant relationship between percent glottal fry and CPP, with greater percentages of glottal fry resulting in lower CPP values.

METHODS

After receipt of approval from the Auburn University Institutional Review Board, 13 male (mean age = 21.46 years, range 19–23) and 13 female (mean age = 22.8 years, range 20–29) participants were recruited for the study. All participants recruited were required to be over the age of 19 and under the age of 40 and self-reported no history of smoking, history of vocal training or voice disorder, drying medications, asthma, or reflux. All participants were not pregnant and were free of respiratory infection or illness prior to participating in this study. The task for this study was completed in a single session. In addition to self-reported absence of voice disorder, auditory-perceptual judgment of speaking voice quality ruled out voice disorder for the participants.

The participants were recorded reading the entire Rainbow Passage at their own pace[29] using a Marantz PMD 671 digital recorder (Marantz, Eindhoven, The Netherlands). The recordings were collected as .wav files with a sampling rate of 44 kHz and a quantization rate of 16 bits. A head set microphone (Isomax E6, Countryman Associates, Inc., Menlo Park, CA) was positioned at a constant distance of approximately 4 cm from the participant's mouth. The participant was seated in a chair in a sound-treated room and instructed to speak at a comfortable pitch and loudness for the task. The Computerized Speech Laboratory (Pentax Medical, Montvale, NJ) was used to analyze the data with a sampling rate of 44.1 kHz. SF_0 and CPP were determined for the entire production of the Rainbow Passage using the Computerized Speech Laboratory main program and Analysis of Dysphonia in Speech and Voice program. The waveform of the passage was analyzed cycle-by-cycle for low-frequency acoustic segments that met the criteria for glottal fry (<100 Hz for females; <80 Hz for males). Each cycle or cycle durations that met the criteria for glottal fry were tagged. Total percent glottal fry was determined for each sample and then trimmed from the sample until a glottal fry-free SF_0 was determined. After all of the glottal fry was removed, the change in speaking fundamental frequency was determined. Intraclass correlation was used to determine whether there was interrater reliability between two independent raters' analyses of glottal fry for 3 (11.5%) of the 26 samples. The intraclass correlation coefficient between the two raters was 1.00, indicating excellent agreement. To normalize the change in SF_0 between males and females, the change in SF_0 was recalculated as the semitone change (ΔST) by determining the difference in semitones between the pretrimmed and posttrimmed SF_0. Analysis of the data using semitones was deemed necessary given the nonlinear characteristic of frequency measurement in Hz throughout a functional frequency range, with smaller change in Hz occurring for a semitone change in the lower frequencies versus higher frequencies.

RESULTS

Prior to completion of the regression analysis, distribution measures and correlation analyses were completed to evaluate which independent variables should be explored as predictors in the regression analysis. Data were analyzed using SPSS version 23.0 (SPSS Inc., Chicago, IL). Distribution measures (ie, skewness and kurtosis) were used to determine what independent variables might impact the regression analysis. Predictors included sex, percent of glottal fry (%GF) in the total Rainbow Passage sample, and CPP. Distribution measures are summarized in Table 1.

To evaluate the strength and the direction of the relationship between semitone change (ΔST) and the independent variables, both Pearson's bivariate and partial correlations were evaluated. Bivariate correlations were conducted for ΔST, sex, %GF, and CPP. Table 2 provides results of the correlation analysis. Results indicate that ΔST has a moderate positive relationship with %GF and sex. Semitone change has a moderate negative relationship with CPP. Percent glottal fry has a weak relationship with both sex and CPP. CPP has a moderate relationship with sex. Partial correlation analysis, controlling for sex, demonstrated the impact that sex has on the remaining two independent variables. When sex is accounted for, the relationship between CPP and ΔST became weak. However, a positive moderately strong relationship remained between ΔST and %GF (see Table 2). Normality assumptions and the assumptions for a linear regression model were met (ie, histograms, Cook's distance, Mahalanobis distance). There was no evidence of multicollinearity or singularity of predictor variables entered into the model, with all tolerance levels above 0.999, and variance

TABLE 1.
Distribution Measures for Variables

Variables	Mean	SD	Range	Variance	Skewness	Kurtosis
Semitone change	0.32	0.41	1.50	0.17	1.86	2.96
Percent glottal fry	5.00	2.79	12.00	7.76	1.01	1.53
CPP	6.35	0.61	2.44	0.38	–0.07	–0.45

Abbreviations: CPP, cepstral peak prominence; SD, standard deviation.