Estimation of Source-Filter Interaction Regions Based on Electroglottography

*,†Anil Palaparthi, *Lynn Maxfield, and *,†,‡Ingo R. Titze, *†Salt Lake City, Utah, and ‡Iowa City, Iowa

Summary: Source-filter interaction is a phenomenon in which acoustic airway pressures influence the glottal airflow at the source (level 1) and the vibration pattern of the vocal folds (level 2). This interaction is most significant when dominant source harmonics are near airway resonances. The influence of acoustic airway pressures on vocal fold vibration (level 2) was studied systematically by changing the supraglottal vocal tract length in human subjects with tube extensions. The subjects were asked to perform fundamental frequency (f_o) glides while phonating through tubes of various lengths. An algorithm was developed using the quasi-open quotient extracted from the electroglottograph. Regions of sudden vocal fold vibration pattern change due to source-filter interaction were inferred from contact area changes. The algorithm correctly identified 89% of male and 84.8% of female quantal changes in contact pattern associated with interactions between source harmonics and formants during ascending glides. During the descending glides, the algorithm correctly identified 84% of male and 81.1% of female quantal changes in contact pattern. These results are in comparison with those obtained from the f_o-based algorithm (Maxfield et al).

Key Words: Source-filter interaction–Electroglottography–Step detection–Quasi-open quotient–Voice production.

INTRODUCTION

Source-filter interaction (SFI) in voice production is described as the influence of the vocal tract on the glottal source. Voice source analysis,[1–3] excised larynx experiments,[4,5] physical model experiments,[6,7] human subject experiments with nonsingers[8,9] and singers[10] have all shown evidence of SFI. As long as the dominant source frequencies lie well below the resonant frequencies of the vocal tract, the filter influences the source mainly in terms of glottal airflow skewing and waveform ripple.[1,3] Although these interactions with glottal airflow occur in speech across all ages and genders, greater interactions between the source and filter are often found in speech of women and children.[11] The complexity of interactions increases when lower partials of the source frequently cross the resonances; this is most often the case in singing, where the fundamental frequency range spans several octaves.[12,13]

Interaction between the vocal source and the filter occurs at two levels.[12] In level 1 interaction, acoustic vocal tract pressures affect the *transglottal pressure*, and with it the glottal airflow (Figure 1, inner loop). At this level, vocal fold vibration can remain relatively undisturbed.[14] In the glottal airflow, however, frequencies creating harmonic distortion are produced that contribute to the source spectrum. This interaction is present in all phonation (speaking, shouting, and singing) by men, women, and children. It contributes to the spectral slope and the spectral ripple in the glottal airflow,[1,3,15] even when the spectrum of the glottal area is harmonic and no disruptions of vocal fold vibration occur.[12]

Level 2 interaction, defined as any source interaction with the vocal tract that disturbs the vibration pattern of the vocal folds, is conceptualized in the outer loop in Figure 1. This interaction occurs primarily when dominant modes of vibration of the vocal folds are affected by *intraglottal pressure,* which can be profoundly different from *transglottal pressure.* In a highly simplified approximation, the two pressures can be written as

$$P_{tg} = P_s - P_i \qquad \text{transglottal pressure} \tag{1}$$

$$P_{ig} = (P_s + P_i)/2 \qquad \text{intraglottal pressure} \tag{2}$$

where P_s is the subglottal pressure and P_i is the supraglottal pressure. Note that the supraglottal pressure (P_i) affects the driving pressure for flow (P_{tg}) in the opposite way it affects the driving pressure for the tissue (P_{ig}). Transglottal pressure decreases with P_i, whereas intraglottal pressure increases with P_i. The supraglottal pressure, P_i, can change suddenly with acoustic reactance of the vocal tract near an airway resonance.[16,17] In these instances, frequency jumps, subharmonic frequencies, chaos, and other bifurcations or instabilities can occur.[12,18]

Electroglottography (EGG), which is an approximation of the contact area between the vocal folds in vibration, can be used to isolate level 2 interaction from level 1 interaction because it does not directly involve glottal airflow. In this article, the changes in the vibratory pattern of the vocal folds due to level 2 interaction are inferred using changes in the vocal fold contact area in human subjects performing f_o glides. The more conventional way to predict the changes in vocal fold vibration is to use the microphone signal and extract an f_o contour,[13] but not every level 2 interaction is necessarily reflected in a sizable f_o change. Two different modes of vibration may have similar frequencies, but their vibration patterns may differ. A more robust way to determine the vocal fold vibration changes is to use the EGG signal. Thus, the motivation for this study was to determine the relative robustness of ΔEGG versus Δf_o for level 2 interaction. The current study was also extended to include both ascending and descending f_o glides in comparison to the Maxfield et al[13] study which considered only ascending f_o glides.

For general speech analysis and synthesis applications, it is important to know which features of the sound source are

Accepted for publication November 21, 2017.

From the *National Center for Voice and Speech, The University of Utah, Salt Lake City, Utah; †Department of Bioengineering, The University of Utah, Salt Lake City, Utah; and the ‡Department of Communication Sciences and Disorders, The University of Iowa, Iowa City, Iowa.

Address correspondence and reprint requests to Anil Palaparthi, National Center for Voice and Speech, The University of Utah, 136 South Main Street, Suite 320, Salt Lake City, UT 84101-3306. E-mail: anil.palaparthi@utah.edu

Journal of Voice, Vol. 33, No. 3, pp. 269–276
0892-1997

https://doi.org/10.1016/j.jvoice.2017.11.012

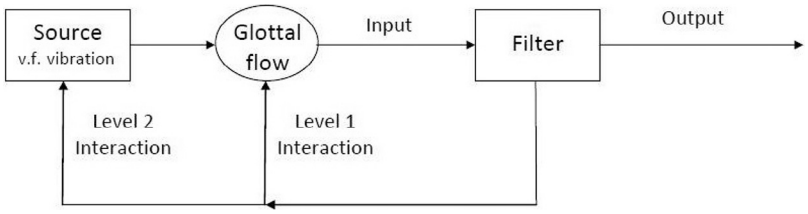

FIGURE 1. Level 1 (inner loop) and level 2 (outer loop) source-filter interaction.

intended by the speaker (controlled) and which are incidental (uncontrolled) due to interaction. In an analysis-synthesis loop, for example, if f_o is extracted and the incidental source-filter variations become part of the programmed f_o contour in synthesis, their effect may be doubled, or at least incorrectly represented. Thus, knowing where unplanned source variations are likely to occur is an important part of signal analysis and processing, especially in high f_o speech or singing where harmonics with considerable energy pass through vocal tract resonances.

Brief review of former f_o-based algorithm

In the f_o-based method of Maxfield et al[13] a rate of f_o change, $\dot{f}_o$, was computed for each glide using the equation

$$\dot{f}_o(n) = \frac{f_o(n) - f_o(n-1)}{f_o(n)\,\Delta t} \tag{3}$$

where Δt is the sampling period in s, and n is the sample number. Average rate of f_o change, $\dot{f}_p$, was computed by fitting a third-order polynomial to $\dot{f}_o$. Significant deviations of $\dot{f}_o$ from $\dot{f}_p$ at each sample were computed by setting $\dot{f}_o$ below a threshold to zero (Figure 2). The threshold τ was defined as the standard deviation of $\dot{f}_o$ from the average, $\dot{f}_p$, over the entire glide.

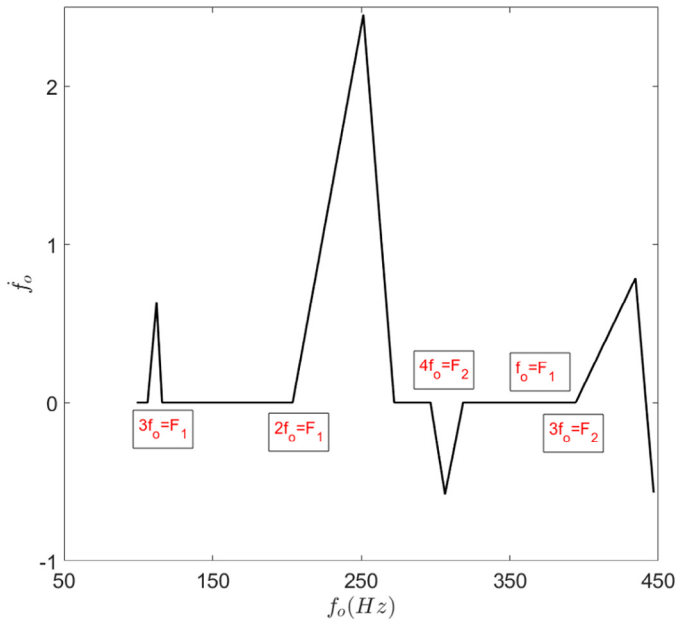

FIGURE 2. Rate of change of f_o versus f_o, indicating source-filter interaction regions during an ascending f_o glide.[13] (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$\tau = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\dot{f}_o(n) - \dot{f}_p(n)\right)^2} \tag{4}$$

Resonance frequencies and bandwidths of the vocal tract were not measured directly in this study. Instead, formants identified from spectrograms were considered as an approximation of the resonances of the vocal tract such that the SFI corresponds to the source harmonic-formant interaction. In our notation, f_n is the n-th source harmonic frequency and F_n is the n-th resonance (formant) frequency.[19]

This analysis indicated that approximately 85% of all deviations were associated with at least one of the resonance/harmonic crossing regions[13] during ascending glides. This result appeared to make f_o change a robust predictor of level 2 interaction. However, there was some evidence that subjects tended to mask (or prevent) f_o perturbations to obtain a smooth f_o glide.[20] Could this f_o smoothing be concomitant with a change in the vibration pattern? This became the primary question for the current study: Is f_o change by itself as good a predictor of Level 2 interaction as the quasi-open quotient of the EGG signal? If so, the result would be important for future signal processing of speech.

METHODS

Subject selection and data collection

Subject selection and data collection procedures were the same as those reported in Maxfield et al.[13] Here, only a brief summary of the procedure is presented. Four men and three women were asked to repeat f_o glides, beginning in vocal fry and extending to 500 Hz for men and 700 Hz for women; they then descended back to vocal fry. Between repetitions, the vocal tract was extended in length using six tubes of varying length, effectively altering the resonance frequencies and ensuring that several harmonics, including the fundamental, would be forced to pass through at least one resonance region. The tubes were fitted with a mouthpiece to ensure uniform jaw and lip placement and held between the subjects' teeth with a tight lip seal around the tube throughout the phonation. This arrangement would also limit the variation in resonances during the glide. The distal end of the tubes was attached to a handle, which was also a mount for a microphone (Countryman Isomax B3), maintaining a constant distance of 1.5 cm from the distal end of the tube. Subjects were fitted with an EGG collar (Kay Elemetrics Model 6103). The microphone signal was amplified using an FMC pre-amp (RNP model) before being captured, along with the EGG signal, using a 16-channel analog-digital converter box (ADInstruments, Powerlab 16/30). All signals were recorded using *Labchart7* data

TABLE 1.
Anticipated F_1 and F_2 Frequencies Calculated for Artificially Elongated Vocal Tracts

Tube Length (cm)		5	7	9	11	15	19
Male	F_1 (Hz)	398	365	337	313	273	243
	F_2 (Hz)	1159	1062	981	911	797	708
Female	F_1 (Hz)	437	398	337	313	291	257
	F_2 (Hz)	1275	1159	1062	981	850	750

collection software. Table 1 lists the approximate first and second formant frequencies estimated for both male and female subjects by artificially elongating the vocal tract with tubes of various lengths. The no-tube condition was deliberately not included because subjects were familiar with the resonances of their own vocal tract and the corresponding interactions. The simple formula for a quarter-wave resonance tube

$$F_n = (2n-1)\frac{c}{4L} \tag{5}$$

was used for resonance estimation (not an exact calculation with end effects), where n is an integer, c is the sound velocity, and L is the length of the tube combined with the length of the vocal tract (estimated at 17 cm for men and 15 cm for women). It can be observed that the tubes are likely to vary the first resonance frequency up to about 150 Hz, and second resonance frequency up to about 450 Hz, although the actual resonance frequencies may vary due to a nonuniform vocal tract shape behind the tube.

Data analysis

In the current study, both ascending and descending portions of an up-down f_o glide were used to study the influence of SFI on vocal fold vibration. Improved estimates of formant frequencies were obtained with linear predictive coding analysis of vocal fry, determined for each recording using the *PRAAT* voice and speech analysis software. Vocal fry was used to view the formant frequencies without the uncertainties caused by widely spread harmonics at higher frequencies. F_1, F_2, and F_3 were extracted during a 500-ms segment of stable vocal fry, both before and after the up-down glides and were averaged to obtain a single estimate of the formants.

Open quotient and quasi-open quotient

An assumption was made that a specific measure on the EGG waveform, the quasi-open quotient (QOQ), is sensitive to most changes in the vibration pattern of the vocal folds. The DECOM (DEGG Correlation-based Open Quotient Measurement) method described by Henrich et al (2004)[21] was used to compute the QOQ. It is important to clarify here that an EGG-based measure of glottal contact area change does not perfectly correlate with a visually determined measure of glottal closure. Indeed, it has been rightly suggested that the terms "open quotient" and "closed quotient" should be used only when referring to metrics derived from methods with which a definite opening and closing of the glottis can be determined (eg, endoscopic videokymography).[22] Recognizing this disparity while also maintaining accuracy to the DECOM method, the resulting derivative of the electroglottograph (DEGG)-determined metric has been labeled QOQ[23] hereafter. The DECOM method uses the DEGG signal to determine QOQ. The peaks in the DEGG signal can be related to the glottal opening and closing instants. The duration between the glottal opening event and the consecutive glottal closing event is the open time. The duration between two consecutive glottal closing events is the fundamental period. The QOQ can then be computed as the ratio of open time to the fundamental period. The QOQ was computed for every 10 ms over a window length of 30 ms.

Level-2 SFI regions determined from QOQ

The QOQ typically increases in the amount of 0.1–0.3 at modal-falsetto transitions during an f_o glide.[24] In the absence of laryngeal register transitions, QOQ changes only slightly around a mean value of 0.5 in a gradual f_o change. However, in the current study, small changes in the QOQ on the order of 0.02–0.1 were also observed to be a result of SFI. These changes in QOQ were attributed to changes in vibratory pattern of the vocal folds due to differences in acoustic pressures as harmonics passed through resonance regions.[12] These changes could be visualized as steps in the QOQ contour. Hence, an algorithm based on step detection was developed to determine the SFI regions during an f_o glide. The region between successive ascending and descending steps consists of an SFI region (Figure 3E, red portions). The algorithm was designed based on the assumption that QOQ increases during SFI. There were a few instances where QOQ decreased, especially during the descending glides, but these instances were very few in comparison to those where QOQ increased.

The QOQ contour extracted using DECOM exhibited both drift and high-frequency noise (Figure 3A). These were removed by passing the QOQ signal through a band pass filter with lower cutoff frequency of 0.1 Hz and higher cutoff frequency of 5 Hz. The resultant filtered QOQ signal can be seen in Figure 3B. The filtered QOQ signal was then passed through a median filter with cutoff frequency of 9 Hz to enhance the steps in the QOQ signal (Figure 3C). These cutoff frequencies were chosen based on the strength of the signal. There is not much energy in the signal after 9 Hz as the magnitude is reduced by 40 dB at 9 Hz. A median filter replaces each sample with the median of neighboring samples, thus enhancing the steps. Finally, the standard Potts model[25] was used for step detection in the median-filtered QOQ signal (Figure 3D, blue and green dots). The Potts model solves the nonconvex optimization problem given by

$$P_\gamma(x) = \arg\min_{x\in\mathbb{R}^N} \gamma\|\nabla x\|_0 + \|Ax-b\|_p^p \tag{6}$$

The term

$$\|\nabla x\|_0 = \#\{i : x_i \neq x_{i+1}\} \tag{7}$$

penalizes the number of steps and the term

$$\|Ax-b\|_p^p = \sum_{i=1}^{N} |Ax_i - b_i|^p \tag{8}$$

measures fidelity of steps x to data b. The parameter γ controls the tradeoff between jump sparsity and data fidelity and was set

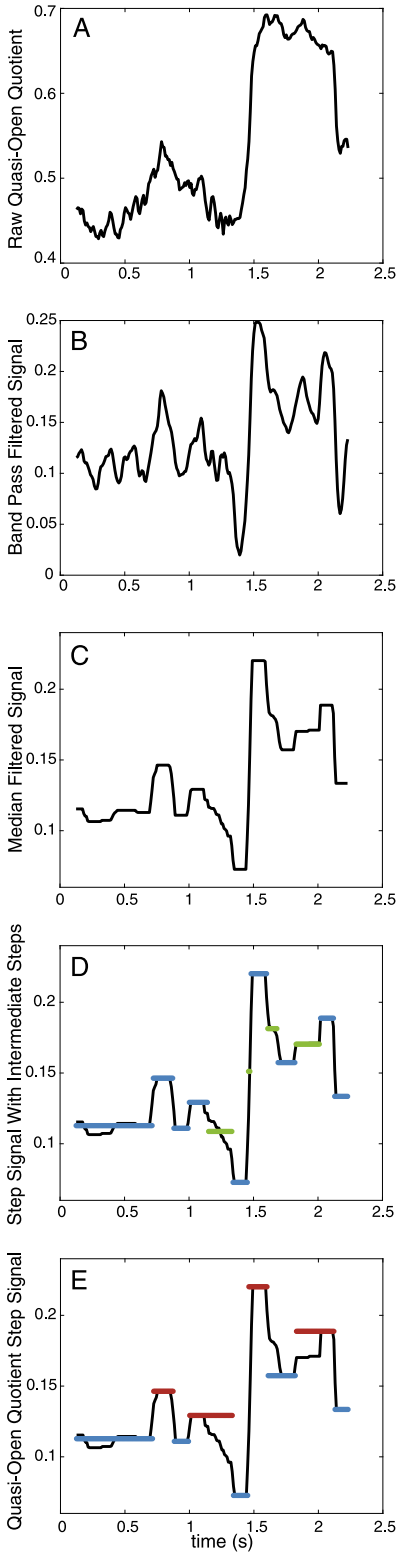

FIGURE 3. Determining level 2 SFI regions from QOQ. **A.** Raw quasi-open quotient (QOQ) contour signal. **B.** Band pass-filtered QOQ signal. **C.** Median-filtered QOQ signal. **D.** Steps in the QOQ signal detected by Potts algorithm. Intermediate regions were represented in green. **E.** Merging of intermediate regions with the extreme regions. The estimated SFI regions were represented in red in the QOQ step signal. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

to 0.1 for male and 0.05 for female data. Matrix A was considered to be an identity matrix in our study and the parameter p was set to 2. The intermediate regions obtained from the Potts model (Figure 3D, green portions) were merged with one of the adjacent extreme regions (Figure 3D, blue portions) based on their proximity to those extreme regions (Figure 3E). However, if the length of the intermediate region was larger than the adjacent lower extreme region, it was merged with the adjacent upper extreme region irrespective of the proximity. The different values of γ for men and women were chosen to reduce the number of these intermediate regions during the analysis. It should be noted that the steps in the QOQ signal start to appear after median filtering itself (Figure 3C). The additional analysis of step detection using Potts model was performed only to highlight the regions of SFI from non-SFI regions. These parameters were maintained the same for both the ascending and the descending glides.

RESULTS

A typical spectrogram of an f_o glide of a male subject is shown in Figure 4. Formants can be seen from the fry portion of the glide. It can be observed that the first three formants interact the most with the first four harmonics at various locations during the glide as f_o begins considerably lower than the first formant. However, the second and third formants interacted the most with the first four harmonics in the case of a female subject.

Ascending glide

Figure 5 shows an example of raw quasi-open quotient signal and its step signal from an ascending glide of a male subject similar to Figure 3E, along with the SFI regions determined using spectrogram (vertical bars). The interaction regions determined using the QOQ step detection algorithm were shown in red in the QOQ step signal. An interaction region in time is defined as the duration for a source harmonic to cross a 50-Hz bandwidth of a formant. Bandwidth of 50 Hz was chosen based on the nominal bandwidths of */a/* vowel formants whose minimum bandwidth was found to be 43 Hz.[26] Because of the use of a mouthpiece for formant stability, and the selection of 50 Hz bandwidth, an exact estimation of the formant frequency is not necessary for this study. For male subjects, interactions where $2f_o \approx F_1$, $f_o \approx F_1$, $2f_o \approx F_2$, $3f_o \approx F_2$, and $4f_o \approx F_2$ were of primary interest. These interactions occupied about 65% of the frequency range of a f_o glide and 54% of the glide time. As can be observed from Figure 5, the steps coincided with the beginning and end of the interaction regions. It can also be observed that some of the interactions overlap (eg, $f_o \approx F_1$ and $3f_o \approx F_2$).

Across the four male subjects, a total of 280 interaction regions, corresponding to $2f_o \approx F_1$, $f_o \approx F_1$, $2f_o \approx F_2$, $3f_o \approx F_2$, and $4f_o \approx F_2$ interactions, were identified using spectrograms. From the EGG signals, a total of 286 regions were identified using the algorithm described above. Of these regions, 250 (89%) reliably aligned with the SFI regions. Thus, the algorithm resulted in 11% false negatives (interaction present but not identified by the QOQ algorithm) and 12% false positives (QOQ regions not associ-

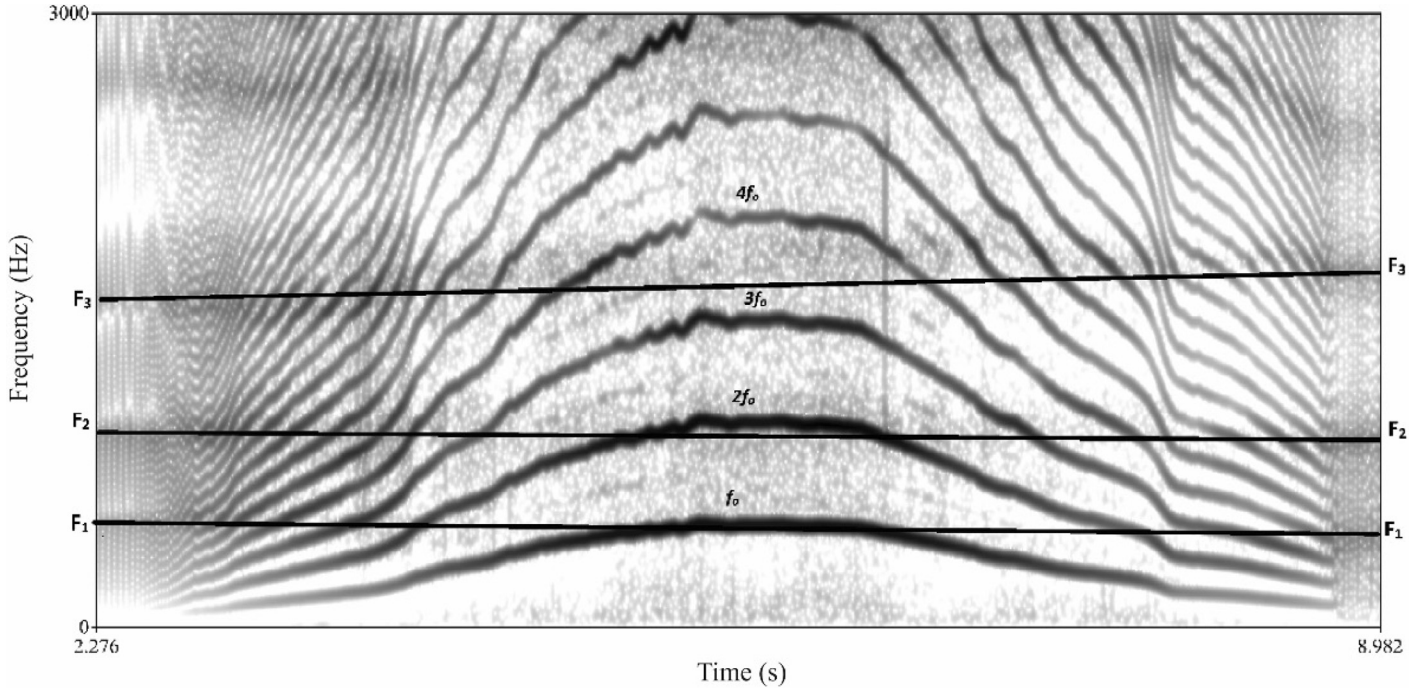

FIGURE 4. Spectrogram of a male subject performing f_o glide with approximate formant locations marked.

ated with any interaction region). The number of correctly identified interactions are listed with respect to individual interactions in Table 2. The first row shows the percentages for male subject data during an ascending glide. Interactions in the

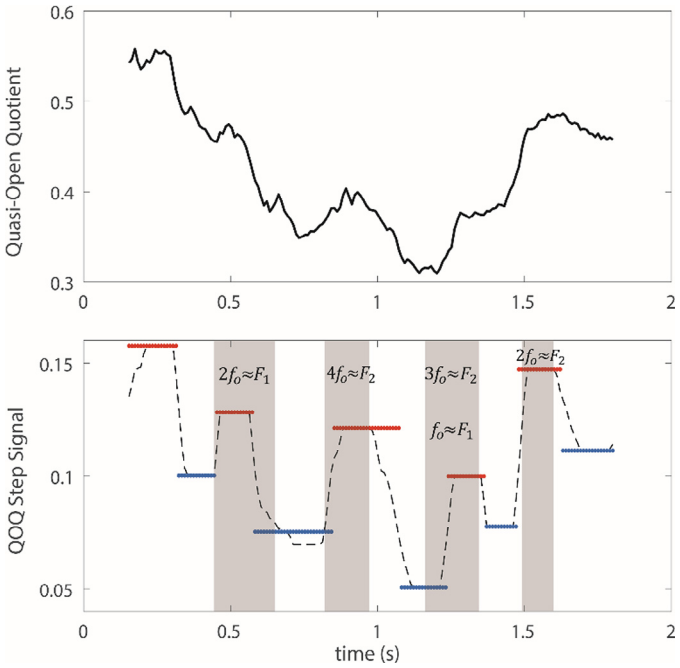

FIGURE 5. (Top) Raw quasi-open quotient (QOQ) waveform. (Bottom) QOQ step signal along with source-filter interaction (SFI) regions determined using spectrogram (vertical bars) for a male subject during ascending glide. SFI regions determined using QOQ step detection algorithm were shown in red in the QOQ step signal. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$2f_o \approx F_1$ and $3f_o \approx F_2$ regions were identified more reliably than the other interactions. The $2f_o \approx F_2$ interaction was the least identifiable for male subjects as it generally occurred during the end of the ascending glide.

For female subjects, the important $2f_o \approx F_1$ interaction was not well identified because the first resonance frequency with tube extensions was very low (257–437 Hz; Table 1). Female subjects quickly transitioned from vocal fry to f_o values in their normal speech range (around 200 Hz), which put $2f_o$ near or above 400 Hz, providing few points for $2f_o \approx F_1$ analysis. Interactions with the second and third formants were easier to determine due to this higher starting fundamental frequency. Most of the harmonics other than f_o were above the first formant. Because of the higher starting f_o, all $2f_o \approx F_1$ interactions were omitted from the study, leaving $f_o \approx F_1$, $2f_o \approx F_2$, $3f_o \approx F_2$, and $4f_o \approx F_2$ interactions, as well as $3f_o \approx F_3$, and $4f_o \approx F_3$ interactions, to be considered for female subjects. In the case of one subject, there were no $4f_o \approx F_2$ interactions because her fundamental frequency started higher than $F_2/4$.

Figure 6 shows a quasi-open quotient step signal from a female subject overlapped with the SFI regions determined using spectrogram (vertical bar) during an ascending glide. It can be observed that the female subjects have greater SFI compared to the male subjects. On an elapsed-time scale, average interaction time for female subjects during a f_o glide is about 63%, a 9% increase compared to the male subjects. Across the three female subjects, 270 interaction regions were studied. The QOQ step detection algorithm identified 249 regions. Among them, a total of 229 regions were identified reliably with the interaction regions leading to 84.8% of correct estimations, 15.2% false negatives, and 8% false positives. The second row of Table 2 provides the percentages of correctly identified interactions by the QOQ algorithm for the female subjects during ascending

TABLE 2.
Number of Correctly Identified Interactions by QOQ Step Detection Algorithm Over Total Interactions for Male and Female Subjects

Glide	Subjects	$2f_o \approx F_1$	$f_o \approx F_1$	$2f_o \approx F_2$	$3f_o \approx F_2$	$4f_o \approx F_2$	$3f_o \approx F_3$	$4f_o \approx F_3$
Ascending	Male	55/56 (98.2)	51/58 (87.9)	44/53 (83.0)	52/57 (91.2)	49/57 (85.9)	—	—
	Female	—	44/52 (84.6)	38/51 (74.5)	48/51 (94.1)	27/33 (81.8)	33/38 (86.8)	39/45 (86.6)
Descending	Male	53/63 (84.1)	54/64 (84.3)	54/62 (87.1)	50/62 (80.6)	51/62 (82.2)	—	—
	Female	—	38/43 (88.3)	40/49 (81.6)	38/47 (80.8)	15/26 (57.7)	30/36 (83.3)	37/43 (86.1)

Values in parenthesis indicate percentage of correctly identified interactions.

glides. It can be observed that the two extreme interactions (ie, $4f_o \approx F_2$ and $2f_o \approx F_2$) were the hardest to identify because the interactions occurred either shortly after the onset of the glide or near the end of the glide and were weak.

Descending glides

We report results from descending glides separately because, on visual inspection, they were not produced as consistently as the ascending glides by all the subjects. Also, level-2 SFI regions during descending glides were determined from both f_o and QOQ for completeness with the former Maxfield et al[13] study. The procedure to determine the level-2 SFI regions from f_o during descending glides was similar to the method used for ascending glides.[13] Figure 7 shows the rate of change of f_o versus f_o for a female subject, indicating SFI regions during a descending f_o glide. The descending glides yielded 584 f_o instabilities across the 4 male and 3 female subjects. Of these 584 instabilities, 508 (87%) occurred at frequencies at which one of the first four harmonics was crossing one of the first three formants. The f_o instabilities from male subjects associated with 88% of crossings, whereas those from female subjects associated with 86%. These percentages are similar to the male (88%) and female (80%) subjects reported for the ascending glides and again significantly higher than chance.

Figure 8 shows a QOQ step signal from a male subject overlapped with the SFI regions. It can be observed that the steps during descending glide were similar to those during the ascending glide. The algorithm was able to identify the start and end of the interaction regions during the descending glide. Across the four male subjects, a total of 313 interaction regions were identified using spectrograms, corresponding to $2f_o \approx F_1$, $f_o \approx F_1$, $2f_o \approx F_2$, $3f_o \approx F_2$, and $4f_o \approx F_2$ interactions. The algorithm described above was able to reliably align with 262 (84%) SFI regions. Thus, the algorithm resulted in 16% false negatives. The number of correctly identified interactions for descending glides

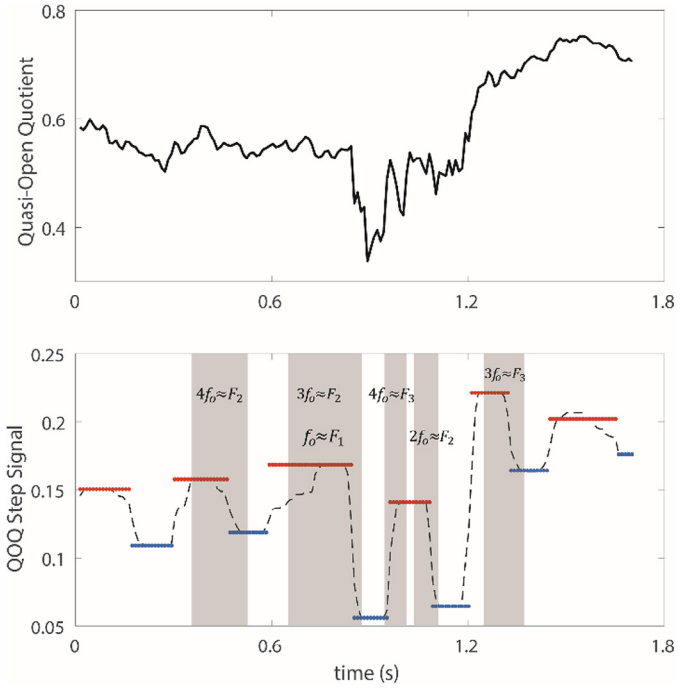

FIGURE 6. (Top) Raw quasi-open quotient (QOQ) waveform. (Bottom) QOQ step signal along with source-filter interaction (SFI) regions determined using spectrogram (vertical bars) for a female subject during ascending glide. SFI regions determined using QOQ step detection algorithm were represented in red in the QOQ step signal. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

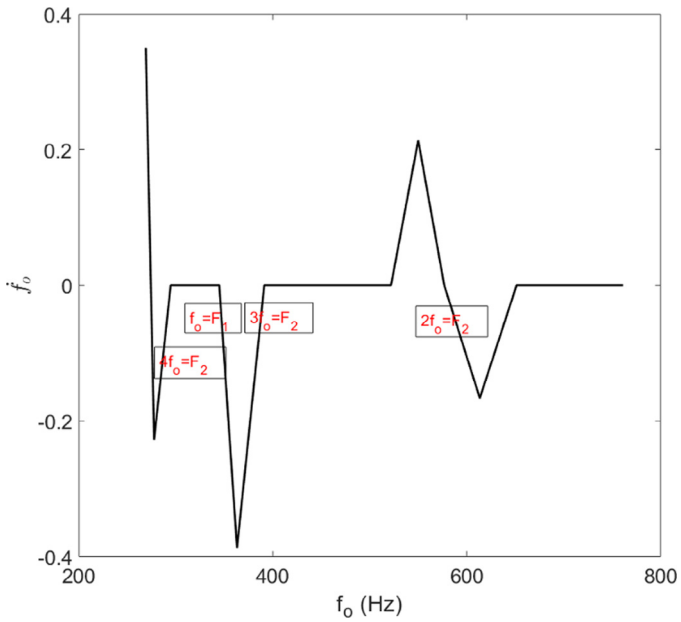

FIGURE 7. Rate of change of f_o versus f_o, indicating source-filter interaction regions for a female subject during a descending f_o glide. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

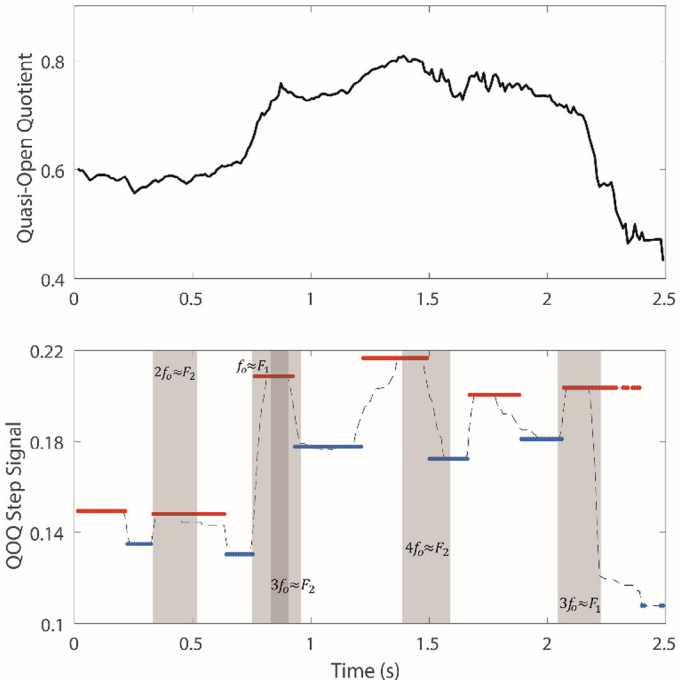

FIGURE 8. (Top) Raw quasi-open quotient (QOQ) waveform. (Bottom) QOQ step signal along with source-filter interaction (SFI) regions determined using spectrogram (vertical bars) for a male subject during descending glide. SFI regions determined using QOQ step detection algorithm were represented in red in the QOQ step signal. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

were also listed with respect to individual interactions in Table 2. The third row shows the percentages for male subject data. All the interaction regions were estimated with similar reliability unlike the ascending glides where $2f_o \approx F_1$ and $3f_o \approx F_2$ regions were identified more accurately compared to the other regions.

For female subjects, the $2f_o \approx F_1$ interaction was not well identified, even for the descending glides. The interaction regions $f_o \approx F_1$, $2f_o \approx F_2$, $3f_o \approx F_2$, $4f_o \approx F_2$, $3f_o \approx F_3$, and $4f_o \approx F_3$ were considered similar to the ascending glides. Across the three female subjects, 244 interaction regions were studied. Among them, a total of 198 regions were reliably identified with the interaction regions determined using spectrogram leading to 81.1% of correct estimations, and 18.9% false negatives. The fourth row of Table 2 provides the percentages of correctly identified interactions during descending glide for the female subjects. It can be observed that the $4f_o \approx F_2$ interaction was the least identified among the other interactions as it occurred mostly during the end of the glide. The other interactions were identified with greater accuracy, similar to the ascending glides.

DISCUSSION AND CONCLUSIONS

Results from the current study demonstrate that electroglottographic signals can be used to identify quantal changes in vocal fold contact, and therewith presumably changes in the vibratory patterns of the vocal folds, resulting from SFI. Treating the changes in the EGG signal as quantal steps allowed prediction of the start and end of the interaction regions. The finding of earlier studies that SFI is stronger in women[27] was also confirmed. It was further validated that women encounter more interaction regions (six vs. five) compared to men due to their wider f_o ranges. It was also found that higher order harmonic-resonance crossings may have a significant impact on f_o stability. However, the current study was performed only on normal healthy subjects and thus does not provide any insight into how the results change in the presence of voice disorders. In particular, when the modes of vibration of the left vocal fold are not synchronized with those of the right vocal fold, more interaction may occur.

The fidelity of the EGG signal is important to robustly predict quantal changes in contact pattern of the vocal folds from an EGG signal. Obtaining an EGG signal during f_o glides was a challenge as the subjects often tended to raise their larynx during the glide and the signal became weak. We tried to instruct the subjects to consciously be aware of the position of their larynx and to perform repetitions until a strong EGG signal was obtained. Beyond recording fidelity, the next step in accurately predicting quantal changes in contact pattern is obtaining a smooth quasi-open quotient signal from the EGG without many discontinuities. In this study, we used the DECOM method given in Reference 21, which resulted in a smooth quasi-open quotient signal in most of the cases. When the QOQ signal appeared to be inaccurate, such test cases were removed from further analysis. With a good QOQ signal, the QOQ step detection algorithm was able to estimate the regions of quantal change in the contact pattern in the vocal folds. Predicting the vibratory mode details may also be possible with the QOQ step detection algorithm. This approach was not studied here, but should be included in a future study.

Maxfield et al[13] and the current study both used data from the same subjects, and the percentages of correctly identified interaction regions were slightly higher from the QOQ step detection algorithm for ascending glides (89% vs. 88% for male subjects and 84.8% vs. 80% for female subjects) than for descending glides (84% vs. 88% for male subjects and 81.1% vs. 86% for female subjects). Overall, the difference in results between the two algorithms is not significant. Hence, it can be concluded that both f_o-based and QOQ-based algorithms perform similarly in estimating the regions of SFI. We also tried to develop an algorithm based on the slope of EGG signal during opening and closing instances to identify SFI regions. The algorithm turned out to be as complex as the QOQ-based algorithm and the results were also not much different. There are both advantages and disadvantages with f_o-based and QOQ-based algorithms. The QOQ-based algorithm is more complex compared to the f_o-based algorithm in terms of computation and analysis. On the other hand, accurately determining f_o in each time window using automated algorithms is difficult. The results would change significantly if f_o were to be determined incorrectly even in one or two time windows. In this and the previous Maxfield et al[13] studies, we used semiautomated algorithms where we manually corrected the incorrect f_o estimations after running through automated algorithms. However, the results from QOQ-based algorithm were not that strongly dependent on few errors in QOQ estimation as filtering was used.

The regions of quantal changes in vocal fold contact due to SFI were more reliably identified during the ascending glides as compared to the descending glides (89% vs. 84% for male subjects, and 84.8% vs. 81.1% for female subjects) using the QOQ-based method. The major reason for this degraded performance was the reduced energy in the descending glide signals compared to the ascending glide signals. Subjects may have depleted their usable lung pressure at the end of the descending glide, potentially creating weak signals. It became hard to identify the SFI regions as well as estimate the QOQ accurately during descending glides compared to the ascending glides. It was possible to identify f_o accurately especially using semiautomated techniques even from weak signals. Hence, the results from f_o-based method were more comparable between the ascending and descending glides.

This study could have been done using different tube lengths for male than for female subjects, thus spanning a wider and more appropriate range of F_1–F_2 for both genders. In particular, a tube shorter than 5 cm could have been used for female subjects to increase F_1. This use of the same tube lengths for both genders may have contributed to the discrepancy in results between male and female subjects. The fixed mouthpiece gave good formant control and less opportunity for the subject to predict where interactions would take place compared to if natural vowels had been selected. A future study might address the adaptation or compensation achievable when these interactions are predictable and unwanted. Future studies might also address the perceptual significance of voice quality due to sudden changes in SFI.[28]

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders Grant R01 DC012045, awarded to Ingo R. Titze, Principal Investigator.

REFERENCES

1. Rothenberg M. Acoustic interaction between the glottal source and the vocal tract. In: Stevens KN, Hirano M, eds. *Vocal Fold Physiology*. Tokyo: University of Tokyo Press; 1981:305–328.
2. Rothenberg M. Acoustic reinforcement of vocal fold vibratory behavior in singing. In: Fujimura O, ed. *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York: Raven Press; 1988:379–389.
3. Fant G. Glottal flow: models and interaction. *J Phonetics*. 1986;14:393–399.
4. Alipour F, Montequin D, Tayama N. Aerodynamic profiles of a hemilarynx with a vocal tract. *Ann Otol Rhinol Laryngol*. 2001;110:550–555.
5. Smith BL, Nemcek SP, Swinarski KA, et al. Nonlinear source-filter coupling due to the addition of a simplified vocal tract model for excised larynx experiments. *J Voice*. 2013;27:261–266.
6. Chan RW, Titze IR. Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics. *J Acoust Soc Am*. 2006;119:2351–2362.
7. Zhang Z, Neubauer J, Berry DA. The influence of subglottal acoustics on laboratory models of phonation. *J Acoust Soc Am*. 2006;120:1558–1569.
8. Titze IR, Riede T, Popolo P. Nonlinear source-filter coupling in phonation: vocal exercises. *J Acoust Soc Am*. 2008;123:1902–1915.
9. Zanartu M, Mehta DD, Ho JC, et al. Observation and analysis of in vivo vocal fold tissue instabilities produced by nonlinear source-filter coupling: a case study. *J Acoust Soc Am*. 2011;129:326–339.
10. Titze IR, Worley AS. Modeling source-filter interaction in belting and high-pitched operatic male singing. *J Acoust Soc Am*. 2009;126:1530–1540.
11. Story BH, Bunton K. Production of child-like vowels with nonlinear interaction of glottal flow and vocal tract resonance. *Proc Meeting Acoust*. 2013;19:060303.
12. Titze IR. Nonlinear source-filter coupling in phonation: theory. *J Acoust Soc Am*. 2008;123:2733–2749.
13. Maxfield LM, Palaparthi A, Titze IR. New evidence that nonlinear source-filter coupling affects harmonic intensity and f_o stability during instances of harmonics crossing formants. *J Voice*. 2017;31:149–156.
14. Koc T, Ciloglu T. Nonlinear interactive source-filter models for speech. *Comput Speech Lang*. 2016;36:365–394.
15. Fant G, Lin Q. Glottal voice source-vocal tract acoustic interaction. *Q Prog Status Rep*. 1987;1–13. Royal Institute of Technology, Stockholm. STL Q
16. Lucero JC, Louremco C, Hermant N, et al. Effect of source-tract acoustic coupling on the oscillation onset of the vocal folds. *J Acoust Soc Am*. 2012;132:403–411.
17. Hatzikirou H, Fitch W, Herzer H. Voice instabilities due to source-tract interaction. *Acta Acust*. 2006;92:468–475.
18. Giovanni A, Ouaknine M, Guelfucci B, et al. Nonlinear behavior of vocal fold vibration. *J Voice*. 1999;27:261–266.
19. Titze IR, Baken RJ, Bozeman KW, et al. Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. *J Acoust Soc Am*. 2015;137:3005–3007.
20. Davis S. Acoustic characteristics of normal and pathological voices. In: Lass NK, ed. *Speech and Language: Advances in Basic Research and Practice*, Vol. 1. New York: Academic Press; 1979:271–335.
21. Henrich N, D'Alessandro C, Doval B, et al. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J Acoust Soc Am*. 2004;115:1321–1332.
22. Herbst CT, Schutte HK, Bowling DL, et al. Comparing chalk with cheese—the EGG contact quotient is only a limited surrogate of the closed quotient. *J Voice*. 2017;31:401–409.
23. Hacki T. Electroglottographic quasi-open quotient and amplitude in crescendo phonation. *J Voice*. 1996;10:342–347.
24. Henrich N, D'Alessandro C, Doval B, et al. Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *J Acoust Soc Am*. 2005;117:1417–1430.
25. Storath M, Weinmann A, Demaret L. Jump-sparse and sparse recovery using Potts functionals. *IEEE Trans Signal Process*. 2014;62:3654–3666.
26. Titze IR, Palaparthi A, Smith SL. Benchmarks for time-domain simulation of sound propagation in soft-walled airways: steady configurations. *J Acoust Soc Am*. 2014;136:3249–3261.
27. Bozeman KW. *Practical Vocal Acoustics-Pedagogic Applications for Teachers and Singers, No. 9*. Pendragon: Press Hilsdale; 2013:32–35 Vox Musicae: The Voice, Vocal Pedagogy, and Song.
28. Samlan RA, Kreiman J. Perceptual consequences of changes in epilaryngeal area and shape. *J Acoust Soc Am*. 2014;136:2798–2806.