Quantitative High-Speed Laryngoscopic Analysis of Vocal Fold Vibration in Fatigued Voice of Young Karaoke Singers

***Edwin M.-L. Yiu, *Gaowu Wang, *Andy C.Y. Lo, *Karen M.-K. Chan, *Estella P.-M. Ma, †Jiangping Kong, and *Elizabeth Ann Barrett,** **Pokfulam, Hong Kong, and †Beijing, China*

Summary: Purpose. The present study aimed to determine whether there were physiological differences in the vocal fold vibration between nonfatigued and fatigued voices using high-speed laryngoscopic imaging and quantitative analysis.

Methods. Twenty participants aged from 18 to 23 years (mean, 21.2 years; standard deviation, 1.3 years) with normal voice were recruited to participate in an extended singing task. Vocal fatigue was induced using a singing task. High-speed laryngoscopic image recordings of /i/ phonation were taken before and after the singing task. The laryngoscopic images were semiautomatically analyzed with the quantitative high-speed video processing program to extract indices related to the anteroposterior dimension (length), transverse dimension (width), and the speed of opening and closing.

Results. Significant reduction in the glottal length-to-width ratio index was found after vocal fatigue. Physiologically, this indicated either a significantly shorter (anteroposteriorly) or a wider (transversely) glottis after vocal fatigue.

Conclusion. The high-speed imaging technique using quantitative analysis has the potential for early identification of vocally fatigued voice.

Key Words: Vocal fatigue–High-speed imaging–Amateur singing.

INTRODUCTION

Vocal fatigue

Vocal fatigue is a common complaint found in teachers, sales professionals, singers, and individuals who constantly use their voice for a prolonged period. It is often described as an increased effort in voicing, harshness, strained voice quality, dryness, and sensation of pain in the throat.[1,2] Some authors considered vocal fatigue as one of the symptoms of voice disorders.[3] Others considered it as an isolated phenomenon.[2,4,5] Stemple et al[2] reported that subjects complained about a dry sensation in the throat and effortful speaking after reading aloud for 2 hours. Teachers with vocal fatigue rated their voice with increased harshness, breathiness, and strain after a day of teaching.[1]

There is yet a consensus on the definition for vocal fatigue. It is generally regarded as vocal tiredness after voice overuse, misuse, or abuse.[2,6] It could happen in speakers with or without any voice problem. Chronic vocal fatigue, however, could be an indicator of subsequent voice disorder.[2,3]

Researchers have attempted to investigate vocal fatigue using aerodynamic, acoustic, and laryngoscopic evaluations. For example, phonation threshold pressure was found to increase after prolonged reading in both women[4] and men.[5] Solomon and DiMattia[4] found that the phonation threshold pressure increased under low-hydration condition. They argued that the viscosity of the vocal folds increases when fatigued. Without sufficient hydration to the vocal folds during prolonged voice use, the stiffness of vocal folds increased, and more effort was required to initiate the vibration of vocal folds. This would lead to an increase in phonation threshold pressure. Stemple et al[2] reported a significant increase in the fundamental frequency after 2 hours of reading. Gelfer et al[7] reported that untrained singers demonstrated an increase in aperiodicity (jitter) and noise levels. Inefficient vocal functions, as indicated by increased airflow rate and reduced maximum phonation time, were found in subjects with chronic laryngeal fatigue, but the fundamental frequency and jitter values remained within normal limit.[8] Indeed, the literature showed mixed results when acoustic measures were used to examine vocal fatigue. The inconsistent results could have been attributed to the different methodologies used in different studies.

With the use of videolaryngostroboscopic examination, anterior glottal chinks[2,8] and abnormal spindle-shaped closure[4,8] have been reported in speakers with vocal fatigue. The presence of chinks or incomplete closure correlated with the perceptual finding of increased breathiness and airflow. Stemple et al[2] hypothesized that the thyroarytenoid muscles would become strained and weak in a fatigue state. Such weakness would cause the bowing of the vocal folds and would lead to an incomplete closure. Prolonged strained contraction of the muscles would give rise to a sensation of pain and increased effort in voicing. In a study by Mann et al,[9] a significant increase in vocal fold edema was observed after a 5-day vocally demanding training. The investigators contended that vocal fold tissues were damaged after an extended vocally demanding task. Gelfer et al[10] found an increase in the amplitude of glottal opening after 1 hour of loud reading. They[10] suggested that the participants might have adopted the loud speaking mode even during the endoscopy task.

The studies reviewed the aforementioned laryngostroboscopic technique, which provided only a pseudo slow motion analysis of the vibration. Intracycle vocal fold vibration pattern

---

Accepted for publication June 13, 2013.

From the *Voice Research Laboratory, Division of Speech & Hearing Sciences, The University of Hong Kong, Pokfulam, Hong Kong, China; and the †Department of Chinese, Peking University, Beijing, China.

Address correspondence and reprint requests to Edwin M.-L. Yiu, Voice Research Laboratory, Division of Speech & Hearing Sciences, The University of Hong Kong, 5/F Prince Philip Dental Hospital, 34 Hospital Road, Pokfulam, Hong Kong, China. E-mail: eyiu@hku.hk

Journal of Voice, Vol. ■, No. ■, pp. 1-9
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2013.06.010

would not be possible with this technique because of the limited camera recording rate. The camera records around 24–30 frames per second. This recording rate could not have captured even one complete vibratory cycle, with the vocal folds vibrating at 80–300 cycles per second. The stroboscopic technique produces images of vibratory motion by sequencing different image frames from different phases across many glottal periods,[11] and hence, videolaryngostroboscopic images do not show a complete full glottal cycle.

High-speed laryngoscopy

High-speed laryngoscopic imaging system emerged in the last 10 years. It can capture up to 8000 frames per second, compared with the ordinary camera system that captures around 24–30 frames per second. High-speed imaging technique has gained popularity over the past decades because of reduced cost and improved resolution of the equipment. With the high-speed system using digital technology, it is possible to examine complete cycle-by-cycle vocal fold vibratory patterns.[11] Cycle-to-cycle visualization allows any irregular vibratory cycle or phase asymmetry to be identified. In a study that compared the usefulness of high-speed imaging and videolaryngostroboscopy in identifying the vibratory features in dysphonic voice,[12] it was found that the use of the high-speed imaging system achieved a 100% identification rate for dysphonic voices compared with the use of videostroboscopy that could only identify 37% of the dysphonic voices correctly. This finding indicates that the videolaryngostroboscopy would not be effective in capturing pathological voices with aperiodic signals. It is, therefore, reasonable to expect that the use of high-speed laryngoscopic imaging system would provide more relevant intracycle information for investigating the physiological changes of fatigued vocal folds, as intracycle aperiodic vibration has been reported as one of the features of vocal fatigue.[7]

Both qualitative and quantitative data of normal and pathological high-speed laryngoscopic images have been reported in the literature. Qualitative methods have been found to be useful in identifying specific vibratory patterns of the vocal folds. For example, specific patterns of glottal closure were reported in diplophonic phonation using kymography, which is a high-speed line scanning technique.[13] A more recent qualitative high-speed laryngoscopic study by Inwald et al[14] reported the glottal closure, degree of mucosal wave, asymmetry, and the amount of mucus deposit in individuals with voice disorders. These qualitative studies were based on perceptual evaluations. Reliability is always an issue in perceptual measurement as it is a subjective rating process. Interrater reliability in perceptual evaluation of laryngoscopic images is usually no better than 70% (eg, Patel et al[12]).

Quantitative measurement (eg, quantifying glottal area in pixels), on the other hand, is a more objective and less variable method that facilitates data summarization and interpretation. Quantitative measures of the glottal area, glottal width, and glottal length have been reported to be useful for studying normal and dysphonic phonations (eg, Yiu et al[15]). The study by Inwald et al[14] also reported the use of asymmetry and perturbation measures extracted quantitatively from the high-speed images. These investigators[14] contended that the combined method of qualitative and quantitative evaluations could best differentiate between dysphonic and normal voices. Temporal measure using frame-to-frame analysis of the high-speed images has also been suggested to be useful in describing phonation. For example, fundamental frequency and open quotient (ie, ratio of the open phase to the glottal period) can be determined by examining the number of frames with open and nonopen glottis.[16] Whether these measures are useful for describing vocal fatigue voice is still open to investigation.

It should be noted that quantitative methods are not always useful to describe different types of phonations. For example, Mehta et al[17] quantified the amount of left-right displacement waveforms of normal and pressed phonations using kymography. They found no significant difference in the asymmetry between these two types of phonations. In another study, Koster et al[18] quantified and analyzed the change in glottal area and glottal width during different modes of voice onset. They were also not able to find any significant differences among the different modes of voice onset.

The present study used the high-speed video processing (HSVP) program developed at the University of Hong Kong[15,19] to investigate vocal fold vibration in vocally fatigued voices. As reviewed previously, the observation of stiffness of vocal folds,[4] glottal chinks, abnormal spindle-shaped glottic closure,[4,8] and strained thyroarytenoid muscles that might cause bowing of the vocal fold[2] in fatigue voices suggested that quantitative analysis of the glottal configuration and vocal fold vibratory pattern would be a logical choice of assessment direction. The present study aimed to examine the glottal configuration and vocal fold vibratory pattern of fatigued voice, induced by prolonged singing,[20] using quantitative analysis of high-speed laryngoscopic images. It was hypothesized that fatigued and nonfatigued voices would demonstrate different glottal configurations and vibratory patterns because of changes in vocal fold physiology.

METHOD

Participants

Ten males and 10 females were recruited from the University of Hong Kong through the social circle of the second (G.W.) and third (A.C.Y.L.) authors. The participants were 18–23 years old (mean, 21.2 years; standard deviation [SD], 1.3 years). All participants reported to be free of any voice or general health problems, nonsmokers, nonalcoholic drinkers, and had no prior voice training. All the participants were Cantonese speakers who, at the time of the study, were attending or had completed their tertiary education. All the participants were further evaluated perceptually by the third author (A.C.Y.L.) to have normal voice quality at the time of the study. Participants were excluded if they reported to have respiratory disease, such as sore throat or flu, 1 day before the examination.

Procedure

Singing task inducing vocal fatigue. All participants were asked to undertake karaoke singing for a minimum of

95 minutes without rest and without drinking water to induce vocal fatigue. The singing time used in the present study was based on the mean plus two SDs singing time that resulted in vocal fatigue, as reported in a separate study by Yiu and Chan.[20] The loudness level of the background music was set at around 60 dB sound pressure level (SPL) for all participants, and the participants were required to sing at least above 80 dB SPL measured with a sound pressure meter (TES 1530A) at a distance of 30 cm from their mouth.

After singing for 95 minutes, all the participants reported feeling tired. They were further asked to continue singing until they felt they could not sing anymore. The purpose of this procedure was to ensure that all the participants had indeed achieved a vocal fatigue condition after the minimum recommended time (at least 95 minutes) of singing. The final mean singing time among the participants was 103.8 minutes (SD, 7.2 minutes; range, 95–115 minutes).

Participants' self-ratings of vocal conditions. Before the singing task, each participant was asked to rate his or her own vocal conditions on the level of discomfort, dryness in the throat, and effort used in voicing, using an 11-point rating scale (0 = normal and 10 = most severely affected). After the singing task, each participant also rated his or her vocal conditions again.

High-speed laryngoscopic and voice recordings. High-Speed camera 5562 digital high-speed imaging system (Richard Wolf GmbH, Knittlingen, Germany) was used to record the laryngoscopic images before and after the singing task. This was performed by the fourth author (K.M.-K.C.), a qualified speech pathologist who had more than 10 years of experience in conducting laryngoscopy. Synchronized voice signals (in WAV format) were also recorded by a microphone attached to the endoscope at approximately 10 cm from the participant's mouth. The participants were asked to sustain an /i/ phonation for as long as possible at their most comfortable pitch and loudness with their tongue protruded. Two seconds of the sustained /i/ phonation with the onset and offset excluded were captured by the digital high-speed imaging system. Synchronized voice signals were recorded by a microphone attached to the endoscope, which was approximately 10 cm from the mouth. A total of 8192 frames (in AVI format) were recorded for the 2-second time span in each recording. Each participant produced three /i/ phonations for recording before and also after the singing task. Therefore, a total of six recordings were collected for each participant.

Preparation of high-speed laryngoscopic images and synchronized voice samples

The high-speed image recordings were analyzed using the HSVP program developed by the Voice Research Laboratory, The University of Hong Kong.[15,19] A number of ratio indices and temporal measurements could be extracted using the HSVP program. The present study, however, focused on four measures only: fundamental frequency, length-to-width ratio index of the glottis (based on the maximum opening of 100 vibratory cycles), open quotient, and speed quotient.

The third author (A.C.Y.L.) first carried out five manual steps to select and prepare the images for the automatic analysis by the HSVP program:

1. *Selection of a comparable pair of presinging and postsinging recordings*. From among each of the three recordings produced by each participant before and after the singing task, a pair of presinging and postsinging recordings with comparable quality and pitch level was selected for each participant. Each pair of presinging and postsinging recordings was perceptually judged to be similar in pitch and loudness by the third author (A.C.Y.L.). Subsequent analysis of the fundamental frequency and intensity of these selected presinging recordings (mean, 240.3 Hz; SPL, 90.4 dB, respectively) and postsinging phonations (mean, 241.2 Hz; SPL, 88.8 dB) showed no significant difference between them (Wilcoxon signed ranked test—fundamental frequency: $Z = -0.09$, $P = 0.93$; intensity: $Z = -1.36$; $P = 0.17$).
2. *Extraction of image frames for analysis*. A minimum of 100 vibratory cycles is considered to be necessary[19] for the HSVP analysis. Because the fundamental frequency for each participant was different, approximately between 1000 and 2000 stable frames of images that contained 100 cycles were extracted from the 8192 frames of each raw video recording according to the frequency produced by each of the participant. Frames that did not contain view of a full glottis were eliminated from the extraction process.
3. *Postextraction processing of images—resizing and gray-scale conversion of the images*. The HSVP program, at the time of the study, was designed to analyze gray-scale images with a resolution of 120 × 256 pixels (although the most current version of HSVP is now capable of analyzing color images of 256 × 256 pixels). Therefore, the size of the raw video was cropped into 120 × 256 pixels and converted into gray-scale image.
4. *Fine adjustment of image quality*. The processed images were rotated manually so that the longitudinal axis of the glottis aligned with the vertical axis (Figure 1). Manual zooming feature and brightness and contrast controls were also available for clearer visualization. The HSVP program has a build-in motion compensation function, which allows tracking the dynamic movement of the images because of the movement of the endoscope. Corresponding adjustment was made to keep the glottis to remain at the relative position across the frames using the automatic motion compensation function if there was endoscopic movement during the recordings.
5. *Delineation of glottis for analysis*. Once the structures in the image were clearly visualized using the fine adjustment described in step 4 previously, an analysis window (Figure 1) was added to the image to enclose the glottis. The window was placed on the left, right, anterior and posterior edges of the glottis.

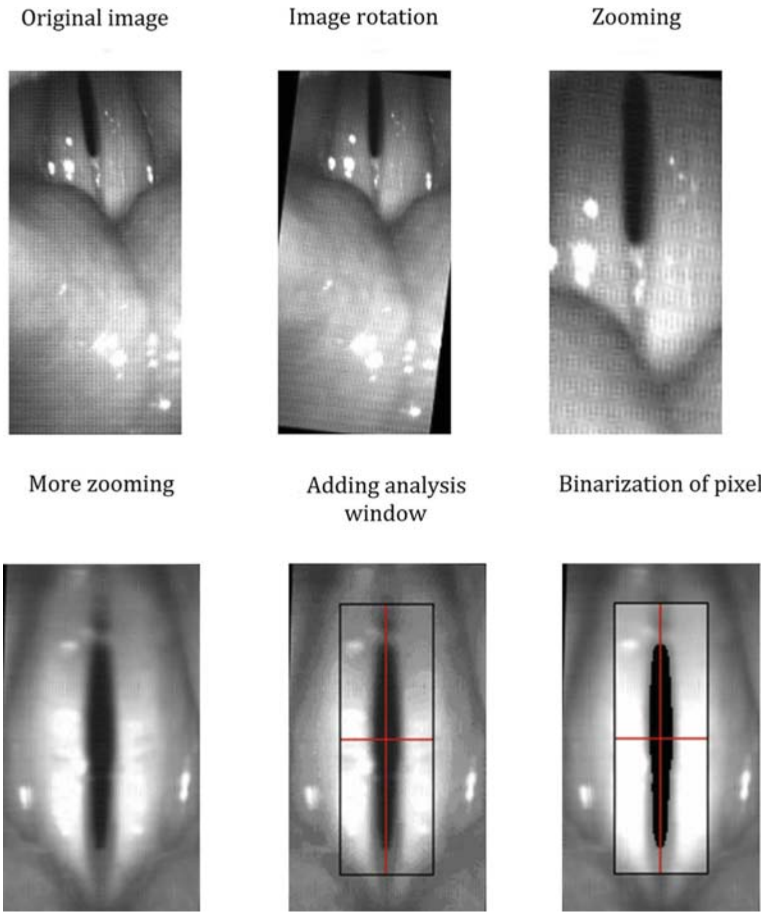


FIGURE 1. Procedures in analyzing high-speed laryngoscopic images.

Data analysis

Interrater and intrarater reliability.

Because the preparation of the high-speed images required some subjective judgment, it was therefore necessary to determine the reliability in preparing these images. Eight video samples, which represented 25% of the total analyzed samples, were randomly selected for reanalysis to determine the reliability of frames selection and the effect on the calculation of the glottal measures. Intrarater reliability was carried out with the third author (A.C.Y.L.) analyzing the images 2 weeks later. Interrater reliability was carried out by comparing the analyses undertaken by the third author with those of the second author (G.W.), who worked as an independent rater.

Acoustic analyses.

Fundamental frequency and the intensity level of the synchronized voice signals were extracted using the *Praat* software.[21] This was carried out by the third author (A.C.Y.L.).

Extraction of the glottal measures.

A procedure to determine the vibratory (fundamental) frequency was carried out on each extracted video image using the HSVP program by the third (A.C.Y.L.) author. The HSVP program calculated the vibratory (fundamental) frequency by transforming the number of frames into a time function based on the sampling rate at 4000 frames per second. For example, if 200 completed vibratory cycles were identified within a 4000-frame video sample, it will be equivalent to 200 cycles per second (ie, fundamental frequency is equal to 200 Hz). The glottal length-to-width ratio index, open quotient, and speed quotient were extracted by analyzing the pixels of the glottis. The HSVP program automatically binarized the pixels into black and white within the analysis window. The black pixels represented the size of the glottis. As it was impossible to determine the actual size of the glottis because of the unknown magnification factors (ie, the distance between the vocal folds and the laryngoscope), a length-to-width ratio index of the glottis based on the maximum opening of 100 vibratory cycles was calculated. A higher index indicated the shape of the glottis to be longer (anteroposteriorly) and/or narrower (transversely) during the maximum opening. The open quotient (the ratio of the glottal opening over one vibratory cycle—calculated by dividing the duration of the open phase by the glottal period) and the speed quotient (the symmetry between the open phase and the close phase—calculated by dividing the duration of opening by the duration of the closing within the open phase) were extracted automatically by the HSVP program. A high open quotient indicated a longer glottal opening in a given cycle, whereas a high speed quotient suggested a longer glottal opening and shorter glottal closing in a given cycle.

TABLE 1.
Mean (SD) for Self-Ratings of Vocal Conditions on an 11-Point Rating Scale (0–10)

	Mean (SD)	
Vocal Conditions	Presinging	Postsinging
Discomfort in the throat*	0.90 (0.91)	6.80 (1.74)
Dryness in the throat*	0.95 (1.00)	7.70 (1.78)
Effort in voicing*	0.55 (1.05)	6.95 (1.99)

* The asterisk symbol indicates significant difference ($P < 0.05$) between singing.

RESULTS

Participants' self-ratings on vocal conditions

Table 1 lists the participants' mean self-rating of vocal conditions before and after the singing task. Bonferroni adjustment was used because three tests were carried out. The alpha was set at .017 (.05/3). The subjects complained of significantly more discomfort and dryness in the throat as well as more effortful in voicing (Wilcoxon signed ranked test for each of the conditions: $Z = -3.93$, $P = 0.001$).

Intrarater and interrater reliability measures

Table 2 lists the interrater and intrarater agreement measures in extracting image frames for the final analysis. For the intrarater agreement measure, 75% of the reanalyzed videos were within ±500 frames of the first analysis. The interrater agreement was lower, with 50% of the videos agreed within ±1000 frames between the two raters.

To determine whether two different samples selected by the same rater or the different raters produced significantly different results in the glottal measures, Wilcoxon signed rank tests were used to determine whether there were significant differences in the glottal measures between the analyses. No significant differences ($P > 0.09$) were found in the glottal measures between the two samples in the interrater and intrarater procedures.

Acoustic and high-speed glottal measures

Table 3 lists the mean acoustic measures (fundamental frequency and intensity) and the mean glottal frequency, mean glottal length-to-width ratio index, mean open and speed quotients before and after the singing tasks from the extracted high-speed images. Analyses were carried out using the combined data from the two gender groups and data from each gender group separately. Bonferroni adjustment with the alpha set at .017 (.05/3) was also used because three tests were carried out for each measure.

Gender differences. No significant differences were found between the two gender groups in the mean voice intensity, glottal length-to-width ratio index, and speed quotient in both presinging and postsinging conditions ($P > 0.05$). The female group, however, demonstrated a significantly higher open quotient when compared with the male group before singing ($Z = -3.78$, $P < 0.0001$) but not after singing ($P > 0.05$).

Changes after vocal fatigue. None of the fundamental frequency, intensity, open quotient and speed quotient measures showed any significant changes after singing ($P > 0.05$, Table 3). The glottal length-to-width ratio index, however, showed a significant reduction after the singing task both in the male and female groups ($P \leq 0.01$, Table 3). Although the mean reduction in the glottal length-to-width ratio index in the male group was bigger in magnitude (−1.67) than that in the female group (−0.56), no statistical significance was found ($Z = -1.96$, $P = 0.052$). The variability of the glottal length-to-width ratio index in the male subjects was indeed rather large, with an SD (1.61) larger than the mean (1.22, Table 3). Therefore, a closer examination of the individual data was conducted.

It was noted that nine males and six females (total = 15) showed a reduction in the ratio index after singing, whereas one male and four females (total = five) demonstrated an increase in the ratio index after singing. For those that demonstrated a reduction in the ratio index after vocal fatigue, the change in the ratio indices ranged from −0.28 to −4.33 and −0.03 to −3.30 in the male and female groups, respectively. For those who demonstrated an increase in the ratio index, the change in the ratio index was 0.63 in the male and ranged from 0.09 to 0.90 in the female groups.

DISCUSSION

The aim of this study was to examine the glottal configuration and vocal fold vibratory pattern after vocal fatigue using high-speed laryngoscopic imaging. Vocal fatigue was induced using a prolonged karaoke singing task. The subjects' self-perception of the vocal conditions was analyzed, and the glottal configuration and vibratory patterns were quantified using three primary glottal measures: glottal length-to-width ratio index, open quotient, and speed quotient.

Self-ratings on vocal condition before and after singing task

After a mean singing time of 103.8 minutes (SD, 7.2 minutes; range, 95–115 minutes), all the participants reported to have

TABLE 2.
The Interrater and Intrarater Agreement in the Extraction of Image Frames for Analysis

Agreement	±250 Frames	±500 Frames	±1000 Frames	±2000 Frames	±4000 Frames
Intrarater	37.5% (3/8)	75% (6/8)	87.5% (7/8)	87.5% (7/8)	100% (8/8)
Interrater	0% (0/8)	37.5% (3/8)	50% (4/8)	50% (4/8)	100% (8/8)

TABLE 3.
Acoustic Measures and Extracted High-Speed Glottal Measures for the Combined Group and Separate Gender Groups

Measures	Mean (SD) Presinging	Mean (SD) Postsinging	Wilcoxon Signed Rank Test
Frequency (Hz) measured acoustically	240.25 (71.13)	241.20 (71.97)	$Z = -0.09$, $P = 0.93$
Male	179.80 (17.66)	181.20 (23.31)	$Z = -0.10$, $P = 0.9$
Female	300.70 (47.41)	301.20 (48.91)	$Z = -0.15$, $P = 0.88$
Intensity (dB)	90.40 (6.75)	88.80 (6.46)	$Z = -1.36$, $P = 0.17$
Male	89.80 (7.36)	87.50 (8.41)	$Z = -1.30$, $P = 0.2$
Female	91.00 (6.43)	90.10 (3.69)	$Z = -0.54$, $P = 0.59$
High-speed extracted glottal frequency (Hz)	239.85 (69.67)	242.90 (72.13)	$Z = -1.09$, $P = 0.27$
Male	181.00 (19.25)	183.60 (25.02)	$Z = -0.46$, $P = 0.65$
Female	298.70 (46.69)	302.20 (50.44)	$Z = -0.89$, $P = 0.37$
Glottal length-to-width ratio index*	2.79 (1.12)	1.68 (1.52)	$Z = -2.65$, $P = 0.006$
Male	2.89 (1.39)	1.22 (1.61)	$Z = -2.49$, $P = 0.01$
Female	2.69 (0.81)	2.13 (1.35)	$Z = -2.70$, $P = 0.007$
Open quotient (%)	69.50 (9.6)	67.10 (11.9)	$Z = -1.07$, $P = 0.28$
Male	63.40 (9.24)	61.80 (12.22)	$Z = -0.59$, $P = 0.55$
Female	75.60 (5.10)	72.40 (9.35)	$Z = -0.76$, $P = 0.44$
Speed quotient (%)	106.40 (20.5)	106.60 (29.4)	$Z = -0.56$, $P = 0.57$
Male	98.60 (13.51)	103.30 (21.94)	$Z = -0.53$, $P = 0.59$
Female	114.10 (23.95)	109.90 (36.27)	$Z = -0.15$, $P = 0.88$

* The asterisk symbol indicates significant difference at 0.017 level.

vocal fatigue. The participants self-rated their level of discomfort in the throat, dryness, and voicing effort to be significantly worse after the singing task (Table 1). These three features are common signs of vocal fatigue (eg, Stemple et al[2]; Hunter and Titze[22]) and could be considered as cardinal perceptual signs for identifying vocal fatigue in speakers.

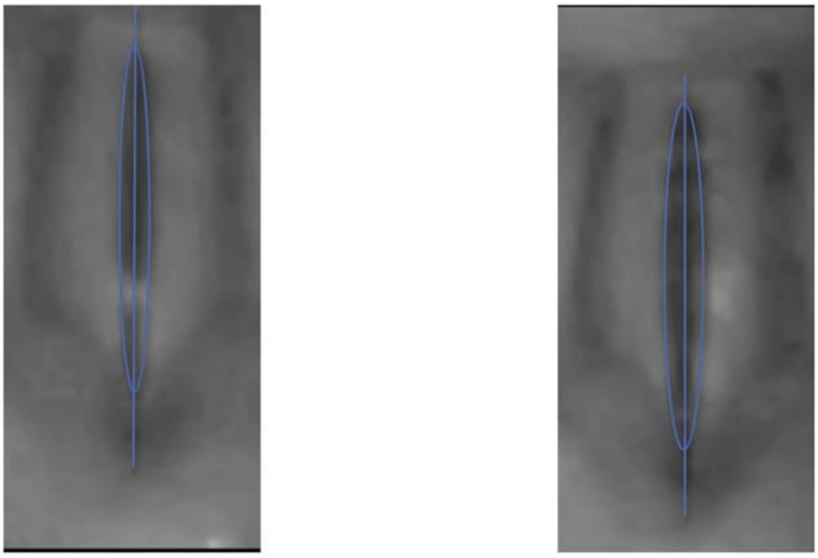
Reliability of the data processing

The reliability in extracting the high-speed glottal measures is dependent on the precision of the manual extraction procedures carried out by the investigator. The intrarater agreement within 1000 frames was more than 87%, whereas the interrater agreement was only 50%. Nevertheless, statistical results showed no significant impact on the final extracted measurements despite the discrepancy in the frame selection. The use of a large number of frames and hence averaging out the measurements could have possibly reduced the impact on the final extracted data, despite the relatively moderate interrater agreement. It is essential that these reliability data should be considered carefully in any imaging processing studies, and investigators should ensure that every effort has been made to obtain the highest reliability or agreement.

High-speed glottal measures

The mean glottal length-to-width ratio index decreased significantly from 2.79 to 1.68 after the prolonged singing task (Table 3). This suggested that the glottis in the fatigued voice demonstrated a relative shortening of the vocal folds anteroposteriorly or widening of the glottis transversely. Whether it was a shortening in the anteroposterior dimension or widening in the transverse dimension was not possible because of the limitation of the derivation of the ratio index. Figure 2 illustrates the shape of the glottis of one of the subjects who showed a typical reduced glottal length-to-width ratio. The magnification of the two images was adjusted accordingly based on the reference landmarks using the anterior and posterior ends of the vocal folds identified in the two images. It can be seen that the glottis after singing (fatigue) became relatively wider when compared with that of before singing. Such widening of the glottis could be interpreted as an increase in vibratory amplitude. Gelfer et al,[10] who also found an increased amplitude in fatigued vocal fold vibration, argued that this might have been an adaptation or compensatory effect to the already fatigue condition. Indeed, the compensatory hypothesis has also been proposed by Linville,[23] who found in a study that the amount of glottal closure increased after 15 minutes of loud reading. Both Linville[23] and Gelfer et al[10] contended that their participants generalized the loud reading phonatory mode to the posttask evaluations, resulting in the endoscopic observation of increase in vocal fold contact and greater amplitude of vocal fold excursion.

On the other hand, Stemple et al[2] reported an increase in the presence of incomplete glottal closure after vocal fatigue. They hypothesized that the incomplete glottal closure was because of thyroarytenoid muscle weakness, causing bowing at the edge of the vocal folds. The present study, however, found no change in the glottal closure pattern before and after the singing task.



Presinging Postsinging

FIGURE 2. An illustration of the shape of glottis in a participant who showed a typical reduced glottal length-to-width ratio after prolonged singing task.

Indeed, 17 of the subjects showed a posterior glottal chink before and after the singing task. One subject showed an incomplete glottal closure before and after the singing task, whereas only two subjects showed complete glottal closure before and after the singing task. Hence, the present study did not have adequate evidence to support the thyroarytenoid weakness hypothesis. It was likely that the participants in the present study might have also adopted a hypertensive mode so that the medial compression increased after the singing task resulted in a decrease in the mean glottal length-to-width ratio index.

It should be noted, however, that not all subjects showed similar changes in the glottal shape. A closer examination of individual data revealed that five (one male and four females) of the 20 subjects showed a reversed pattern, that is, an increase in the glottal length-to-width ratio after singing. Nevertheless, the magnitude of increase (mean, 0.56; range, 0.09–0.90) was relatively smaller than that of the magnitude of reduction (mean, −1.67; range, −0.03 to −4.33). Such small magnitude changes that did not conform to the general trend could have been attributed to individual differences in response to vocal fatigue. This observation, however, appeared to be similar to a single case study reported by Boucher and Ayad.[24] They found reduced lateral cricoarytenoid muscle activities in their subject during vocal fatigue. At the same time, the muscular activities of thyroarytenoid and cricothyroid muscles increased to compensate for the decrease in activity in the lateral cricoarytenoid muscles. The increase in thyroarytenoid muscle activities served to tense and stretch the vocal folds to stabilize the adduction force.

No significant differences were found in the open and speed quotients before and after the singing task. This result contradicted with that reported by Lauri et al,[25] who found a higher speed quotient and a lower closing quotient using electroglottography (EGG) after vocal fatigue. They hypothesized that these changes were because of an increase in adductory force, reflecting a hyperfunctional vocal adjustment. It should be noted that the temporal parameters like open quotient generated from high-speed imaging did not necessarily correlate with that obtained using EGG.[26] There are two possible explanations for such a discrepancy in the findings. First, it might have been because of the technical limitation of the high-speed imaging. The sampling rate of the high-speed imaging in the present study was only 4000 frames per second (4 kHz), whereas EGGs are analog signals that are usually sampled at above 20 kHz or even up to 44 kHz. Second, it might well be that vocal fatigue does not necessarily cause any changes in the vibratory pattern in the temporal dimension, as measured by open and speed quotients, whereas, the spatial configuration of the glottis could be affected. To conceptualize this, one might like to consider an analogy in the physics of a swinging pendulum. In a free-swinging pendulum, the amplitude of the swings (spatial configuration) gradually reduces, whereas the frequency (temporal dimension) remains the same. Although this second hypothesis seems more plausible, further studies are required to determine whether the temporal dimensions (open and speed quotients) are preserved in vocally fatigued voice.

The results in this study should not be interpreted without some cautions in mind. First, the participants who took part in this study were relatively young (mean age, 21.2 years). Therefore, the physiological reactions to vocal fatigue might not be the same in individuals who are older. Second, the vocal fatigue-inducing task used in the present study was a singing task, which might have produced different vocal fatigue pattern that was induced by an extended talking task. The pitch level during the recordings was controlled in this study, and the voice onset and offset were excluded from the analysis. Therefore, the effect of vocal fatigue on one's pitch, and the pattern of voice onset and offset could not be determined. Further high-speed imaging investigations should take into consideration the voice onset and offset. In addition, high-speed imaging should be accompanied with other instrumental measurements, such as phonation threshold pressure, electroglottograph, and stroboscopy.

Vocal fatigue is often treated as one of the symptoms of voice disorders.[3] Early identification on vocal fatigue, especially for those occupational voice users, who are prone to developing chronic fatigue, would help preventing from the development of a chronic condition that could result in voice disorders. The present study demonstrated the quantitative analysis of high-speed images using the HSVP program. The program was able to detect the difference on vocal vibration pattern between nonfatigue and fatigue vocal conditions. The findings suggested that the participants might have developed a compensatory behavior after vocal fatigue. More studies using multiple measurements will be needed for a better understanding of the physiological changes in vocal fold under vocal fatigue condition.

Methods in studying vocal fatigue

A final concluding remark on the methodology is warranted here for consideration by investigators on vocal fatigue in the future. It should be noted that studies that investigated vocal fatigue have used different procedural methods to induce vocal fatigue. The present study employed amateurs using a singing task. Whether the use of a reading or talking aloud task would make any difference in the effect on vocal fatigue is not known. Furthermore, it would be interesting to determine if amateur and professional singers would have similar or different response to vocal fatigue in terms of the glottal configuration.

Acknowledgment

This study was supported in part by a grant from the Hong Kong Research Grant Council General Research Fund (#HKU757811). The authors also acknowledge Professor William Wei, Dr Elaine Kwong, Dr Cate Madill, and Samantha Wurhurst, who have contributed to the discussion of the research topic in the initial stage of the project planning.

REFERENCES

1. Gotaas C, Starr CD. Vocal fatigue among teachers. *Folia Phoniatr (Basel)*. 1993;45:120–129.
2. Stemple J, Stanley J, Lee L. Objective measures of voice production in normal subjects following prolonged voice use. *J Voice*. 1995;9: 127–133.
3. Colton RH, Casper JK, Leonard RL. *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2006.
4. Solomon NP, DiMattia MS. Effects of a vocally fatiguing task and systemic hydration on phonation threshold pressure. *J Voice*. 2000;14: 341–362.
5. Solomon NP, Glaze LE, Arnold RR, van Mersbergen M. Effects of a vocally fatiguing task and systemic hydration on men's voice. *J Voice*. 2003;17: 31–46.
6. Welham NV, Maclagan MA. Vocal fatigue: current knowledge and future directions. *J Voice*. 2003;17:21–30.
7. Gelfer MP, Andrews ML, Schmidt CP. Effects of prolonged loud reading on selected measures of vocal function in trained and untrained singers. *J Voice*. 1991;5:158–167.
8. Eustace CS, Stemple JC, Lee L. Objective measures of voice production in patients complaining of laryngeal fatigue. *J Voice*. 1996;10:146–154.
9. Mann EA, McClean MD, Gurevich-Uvena J, Barkmeier J, McKenzie-Garner P, Paffrath J, Patow C. The effects of excessive vocalization on acoustic and videostroboscopic measures of vocal fold condition. *J Voice*. 1999;13:294–302.
10. Gelfer MP, Andrews ML, Schmidt CP. Documenting laryngeal change following prolonged loud reading. A videostroboscopic study. *J Voice*. 1996; 10:368–377.
11. Deliyski DD, Petrushev PP, Bonilha HS, Gerlach TT, Martin-Harris B, Hillman RE. Clinical implementation of laryngeal high-speed videoendoscopy: challenges and evolution. *Folia Phoniatr Logop*. 2008;60: 33–44.
12. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008;117:413–424.
13. Larsson H, Hertegard S, Lindestad P-A, Hammarberg B. Vocal fold vibrations: high-speed imaging, kymography, and acoustic analysis: a preliminary report. *Laryngoscope*. 2000;110:2117–2122.
14. Inwald EC, Dollonger M, Schuster M, Eysholdt U, Bohr C. Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *J Voice*. 2011;25:576–590.
15. Yiu EM-L, Kong J, Fong R, Chan KMK. A preliminary study of a quantitative analysis method for high speed laryngoscopic images. *Int J Speech Lang Pathol*. 2010;12:520–528.
16. Hess M, Gross M. High-speed imaging of vocal fold vibrations and larynx movements within vocalizations of different vowels. *Ann Otol Rhinol Laryngol*. 1996;105:975–981.
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. Koster O, Marx B, Gemmar P, Hess MM, Kunzel HJ. Qualitative and quantitative analysis of voice onset by means of a multidimensional voice analysis system (MVAS) using high-speed imaging. *J Voice*. 1999;13: 355–374.
19. Kong JP, Yiu EM-L. Quantitative analysis of high-speed laryngoscopic images. In: Ma EP-M, Yiu EM-L, eds. *Handbook of Voice Assessments*. San Diego, CA: Plural Publishing; 2011.

20. Yiu EM-L, Chan RMM. Effect of hydration and vocal rest on the vocal fatigue in amateur karaoke singers. *J Voice*. 2003;17:216–227.
21. Boersma P, Weenink D. Praat: doing phonetics by computer [computer program]. Version 5.3.39. 2013. Available at: http://www.praat.org/.
22. Hunter EJ, Titze IR. Quantifying vocal fatigue recovery: dynamic vocal recovery trajectories after a vocal loading exercise. *Ann Otol Rhinol Laryngol*. 2009;118:449–460.
23. Linville SE. Changes in glottal configuration in women after loud talking. *J Voice*. 1995;9:57–65.
24. Boucher VJ, Ayad T. Physiological attributes of vocal fatigue and their acoustic effects: a synthesis of findings for a criterion-based prevention of acquired voice disorders. *J Voice*. 2010;24:324–336.
25. Lauri ER, Alku P, Vilkman E, Sala E, Sihvo M. Effects of prolonged oral reading on time-based glottal flow waveform parameters with special reference to gender differences. *Folia Phoniatr Logop*. 1997;49:234–246.
26. Echternach M, Dippold S, Sundberg J, Arndt S, Zander MF, Richter B. High-speed imaging and electroglottography measurements of the open quotient in untrained male voices' register transitions. *J Voice*. 2010;24:644–650.