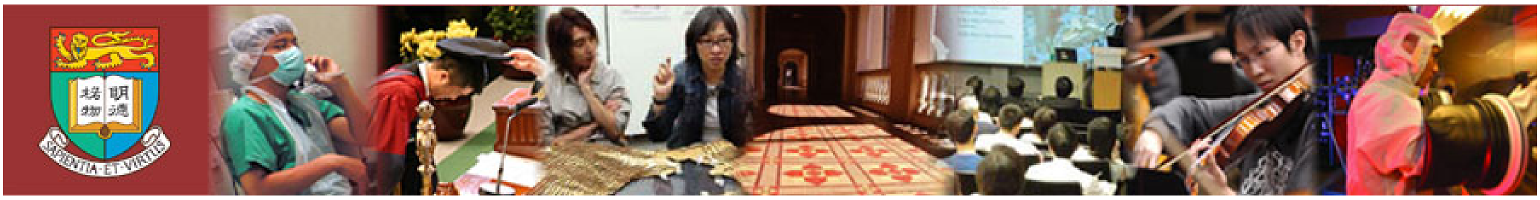


| | |
|---|---|
| Title | High speed laryngoscopic study of vocal fold vibratory patterns in normal and dysphonic subjects |
| Other Contributor(s) | University of Hong Kong |
| Author(s) | Fong, Raymond |
| Citation | |
| Issued Date | 2005 |
| URL | http://hdl.handle.net/10722/56204 |
| Rights | Creative Commons: Attribution 3.0 Hong Kong License |

# High speed laryngoscopic study of vocal fold vibratory patterns in normal and dysphonic subjects

Raymond Fong

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, May 6, 2005.

# Abstract

Video images from high speed laryngoscopy have been used to study vibratory patterns qualitatively. This study used a computer program (High Speed Video Processing Program) to obtain quantitative data from video images. Fifteen subjects with normal voice use and 2 hyperfunctionally dysphonic subjects were recruited. High speed images (2000 frames/second) of the production of three common vocal registers (modal, falsetto, fry) and pitch gliding were recorded for each subject. Nine ratio indices that represented the glottal opening area, width and length were computed using HSVP. High variability was found in all these measures across subjects. The results revealed no significant differences in the quantitative measures between high speed images of modal and falsetto register while those of vocal fry were significantly different from those of the other two registers. The glottal opening area ratio index was found to be a more distinctive indicator of dysphonia in using high speed images than other ratio indices.

Instrumental analysis is often used in evaluating dysphonia. The most direct method is the visual examination of the vocal fold movements. Videostroboscopy is one of the most common instruments used to investigate the pathophysiology of vocal fold movement in voice production. However, videostroboscopic images are not easy to interpret when the vocal fold vibration is irregular or aperiodic (Eysholdt, Tigges, Wittenberg, & Proschel 1996). Stroboscopic images are formed from a combination of different phases of vibratory cycles. Some of the phases are lost in the stroboscopic images. Therefore, irregular vibration patterns are not easily captured by stroboscopic method. Pathological voice production is often due to irregular vocal fold vibration (Eysholdt et. al, 1996). Therefore, videostroboscopy is limited in investigating dysphonic voice.

Electroglottography (EGG) is an indirect measurement of laryngeal movements. The degree of contact of the vocal folds during instances of opening and closure can be obtained indirectly by passing a current through the vocal folds (Eysholdt et. al, 1996). However, this method is not suitable for all subjects. Individuals with incomplete glottal closure, or thick neck tissue which do not allow EGG to pass through are not suitable for this method (Colton, & Casper, 1996).

High speed laryngoscopy has been developed to overcome the limitations of various objective measurements (e.g. stroboscopy and EGG) on pathological voice production (Hirose, 1988). Current technology makes use of a digital camera head attached to an endoscope and a computer that processes and renders videos with as many as 8000 frames per second. This frame rate is affected by the resolution of the video obtained in that the higher the resolution, the lower the frame rate. The images can be analyzed using quantitative methods.

High speed laryngoscopy has been used in three major areas:

1) General study of vocal fold vibration; 2) Study of the underlying physiology of specific voice production mechanisms; and 3) Investigation of the underlying pathophysiology of dysphonic

voice quality. The three areas will be discussed in the following sections

*General study of vocal fold vibration*. High speed digital imaging of vocal fold vibration is still very much in the exploratory phase. Valid and reliable clinical applications have not been established yet due to a lack of specific, reliable and valid measurements that can be used to identify abnormal vibratory patterns. Therefore, much of the studies have been on the application of this tool on studying normal vocal fold vibratory patterns. As early as 1988, Hirose (1988) used high speed laryngoscopy to investigate vocal fold vibration. Early application tended to focus only on irregularity of vocal fold vibration (Hirose, 1988). Hirose (1988) found that incomplete glottal closure and asymmetrical vibratory patterns were easily identified using high speed laryngoscopy. Later studies in the mid-1990s were able to make use of imaging processing techniques to determine the pattern of changes in the glottis (Eysholdt et. al, 1996). Koster, Marx, Gemmar, Hess and Kunzel (1999) used a multidimensional voice analysis system (MVAS) to analyze high speed images that synchronized with EGG signals and acoustic signals. The parameters used were glottal area and widths under different voice onset conditions. They found high intra- and inter-subject variability in the vibratory patterns. They concluded that phonation was a highly complex and variable phenomenon (Koster et. al, 1996).

*Study of phonatory physiology.* High speed laryngoscopy has also been used to study the underlying physiology of specific type of phonation and articulation. Maurer, Hess and Gross (1996) used high speed imaging to investigate the vocal fold vibration and larynx movements in the production of different vowels. An “open-quotient” (number of frames that contained opened glottis/total number of frames in a cycle) measurement was used. The investigators found that different vowels showed distinct patterns of open-quotients. The study also found no simple correlation between the EGG data and the high speed images in terms of the “open-quotient”

(Maurer et. al, 1996). Lindestad, Sodersten, Merker and Granqvist (2001) used high speed imaging as one of the evaluating tools to investigate voice source characteristics in Mongolian "Throat Singing" technique. Kymograms were obtained from high speed images and evaluated qualitatively. Investigators found ventricular fold phonation present in this singing technique.

*Study of the pathophysiology of dysphonic voice.* A number of studies have used high speed imaging to investigate the underlying pathophysiology of deviant voice quality. Vocal fold vibration in diplophonia was investigated by Kiritani, Hirose and Imagawa (1993). The phase difference between the left and right vocal folds were evaluated in successive high speed images by the investigators. They found that when the phase difference between two out-of-phase vocal folds reached a maximum value (threshold), it reset and the two vocal folds were in phase again (Kiritani et. al, 1993). In other words, asymmetry of vocal fold vibration will reach a maximum then the vocal fold vibration will be symmetrical again.

Vibration of the pharynoesophageal segment after total laryngectomy has been investigated by van As, Wittenberg, Op de Coul, Eysholdt and Hilgers (1999). They used high speed imaging to assess the presence of saliva; visibility, size and shape of the neoglottic and the regularity of vibration. Neoglottic vibration is often non-periodic and traditional stroboscopy is not able to investigate aperiodic vibratory patterns. Thus, high speed imaging is able to overcome this limitation and provides potential information regarding the vibratory patterns of the pharynoesophageal segment (van As et. al, 1999).

Verdonck-de Leeuw, Festen and Mahieu (2001) used qualitative measurements to investigate videokymographic images obtained from dysphonic patients with various pathologies that included laryngeal nerve paresis, tumor and cyst. They found that the presence of mucus and phase difference between the vocal folds, as observed in high speed laryngoscopy, contributed to the perception of roughness. Ventricular fold co-vibration in a dysphonic subject was investigated by using both high speed imaging and perceptual analysis by Lindestad, Hammearberg, Larsson and Granqvist (1998). They found that although the vibratory patterns of ventricular folds remained constant, the voice quality varied. They concluded that the ventricular fold co-vibration affected the quality of the voice by changing the overall laryngeal vibratory patterns or by damping the voice produced at the glottal level.

### *Image processing*

A number of studies explored specifically the data processing aspect of high speed imaging using mathematical technique (Tigges, Wittenberg, Mergell & Eysholdt, 1999; Mergell, Herzel & Titze 2000; Granqvist, & Lindestad, 2000). Tigges et. al (1999) defined an algorithm to compensate for the constant relative movement of the glottis within the images. Granqvist and Lindestad (2000) applied Fourier analysis to high speed imaging which allowed the display of laryngeal oscillation in one single diagram. This is known as kymography. Kymograms represent vibratory cycles with multiple images across time in a single picture.

Phonation is a highly complex and variable phenomenon (Koster et. al, 1996). High speed laryngoscopy can provide more precise and finer measurements across a limited time frame than traditional stroboscopic analysis. It is able to provide a substantial amount of quantitative data that previous instruments were not able to. A number of published studies correlate perceptual voice analysis with qualitative high speed image analysis. For example, Niimi and Miyaji (2000) investigated vocal fold symmetry, regularity of vibration and phase differences between vocal folds using qualitative measurements. Qualitative investigation of kymograms generated by high speed laryngoscopy was also reported by Lindestad et. al (2001) and Verdonck-de Leeuw et. al (2001). A number of studies investigated the quantitative measurements obtained from high speed laryngoscopy. Koster et. al (1999) used glottal area and widths while Maurer et. al (1996) used “open-quotient” to study phonatory physiology. However, the scope of these two studies was limited to some specific areas. Koster et. al (1999) focused on voice onsets while Maurer et. al (1996) focused on the production of different vowels.

### *Objective of the study*

The study aimed at determining the quantitative measurements in high speed laryngoscopic images that can be used to describe normal voices. The study made use of a specially designed program that computed measurements (glottal area, length and width) throughout a speech sample with a maximum of 2.184 seconds long. Three vocal registers and pitch gliding were produced by each subject and analyzed accordingly.

Production of different vocal registers is generally regarded as one of the most perceptually salient representation of vocal quality (Laver, 1991). The three common registers include: vocal fry; modal and falsetto. These three registers cover the low, mid and high frequency of the human voice (Colton & Casper, 1996). Pitch gliding was included in the study in addition to the three common vocal registers. It was hypothesized that for the subject to glide from the highest pitch to the lowest pitch possible within a short period of time (i.e. a maximum of 2.184 seconds), the physiological boundaries of the laryngeal structures for phonation were stretched to a maximum. Data obtained from this transition from high to low pitch could then be used to construct a profile of a subject’s phonation in terms of changes of glottal area, width and length, and the rate of changes of these measurements.

*Research Question.* Can quantitative data from high speed laryngoscopy be used to describe vocal fold vibration in normal subjects with modal, falsetto, fry registers and pitch gliding?

# Method

## *Subjects*

A total of 15 non-dysphonic subjects (mean age=21.3 years, standard deviation=1.2 years, range=19-23 years) and two dysphonic subjects (mean age=21 years, standard deviation=0 years) took part in this study. All subjects were non-smokers and non-drinkers. All subjects were not required to use their voice for their daily occupation. All subjects were able to produce modal, falsetto and fry registers, and also pitch gliding (falling) following demonstrations by the investigator (RF).

The non-dysphonic subjects were students from the University of Hong Kong. They had no history of voice disorder. Rigid laryngoscopy was performed by a qualified speech therapist prior to the recording task to confirm no vocal pathologies were present.

The dysphonic subjects were recruited from the Voice Research Laboratory at the University of Hong Kong between December 2004 and January 2005. They were diagnosed by an ear, nose and throat surgeon as having voice disorders with visible vocal fold pathologies. One had a unilateral vocal nodule and one had a unilateral cyst. The presence of vocal pathologies was also confirmed at the time of recording task using rigid laryngoscopy performed by the speech therapist.

## *Equipment*

The high speed images were taken by Kay’s High Speed Video System (Kay; model number: 9700). The system included a rigid laryngoscope (KAY; model number: 9106) attached to a halogen light source. The camera head was attached to a Windows 2000-based computer via an adapter. The images were captured, at 2000 frames per second, by the Kips High Speed Video (KHV) program. Synchronized sound files were recorded through a microphone attached to the camera head (placed about 5 cm from the subject's lips when the images were taken). The microphone was connected to a Windows XP-based computer and the recording was done using Syntrillium's Cool Edit (version 2000).

### *Procedures*

Each subject was asked to produce modal, falsetto and fry register using the vowel /i/ and pitching gliding from high to low pitch (falling: /i1/→/i4/) following the investigator's demonstration. The production by each individual subject was always in the following order: modal, falsetto, fry and pitch gliding.

### *Raw data processing*

Video images (in AVI format with a resolution of 120 X 256 pixels) and synchronized speech signals file (in WAV format at a sampling rate of 10000 Hz, 16 bit; stereo format) were recorded for each of the three vocal registers and pitch gliding produced by each subject. A total of 4183 frames covering a time span of 2.184 seconds were available for analysis in each video file. The video clips and sound files were then processed by a specially designed beta version of the High Speed Video Processing (HSVP) program.

*HSVP program.* The program is designed by Professor Kong Jianping, Beijing University and Dr. Edwin Yiu, University of Hong Kong and it is still under development. It is designed to analyze laryngoscopic video images and speech signals. The program, in its beta version, includes the following editing functions for preparing the images for later analysis:

1. Cropping the images and speech signals: This function enables the user to extract the most representative segment for further analysis.
2. Adjusting contrast and brightness of the images: The program allows the contrast and brightness of the images to be adjusted such that the vocal folds and the surrounding

anatomical landmarks can be visually identified.

3. Rotating the images: The image can be rotated so that the vocal folds are aligned in a vertical manner for later analysis.
4. Adding an analysis “window” to delineate the area surrounding the glottis: After the vocal folds are aligned vertically, a rectangular box with a cross in the middle can be added to surround the glottis. The edges of the window will fall on the four laryngeal landmarks: 1) Anterior and 2) posterior tip and the 3) left and 4) right lateral edges of the glottis.

Figure 1 illustrates 1) first rotating the image frame (Figure 1a) until the vocal folds are aligned vertically (Figure 1b); 2) Adding the “analysis window” (Figure 1b).

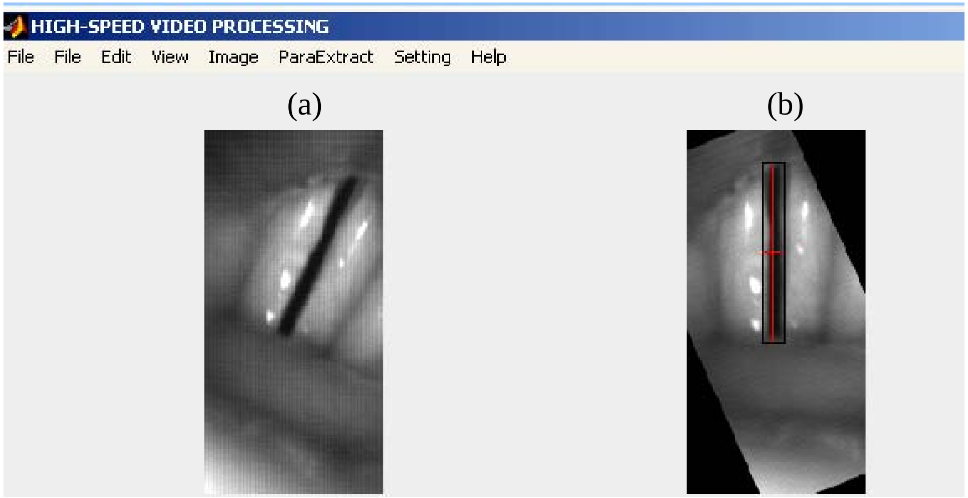


*Figure 1.* Diagram illustrating (a) rotating the image and (b) adding the “analysis window”

After the images and the corresponding sound file containing the speech signals were extracted, edited, adjusted for brightness and contrast, rotated appropriately so that the vocal folds were aligned vertically and the “analysis window” was added, the following measurements were extracted in each selected frame:

1. Total glottal area (GA), left (lGA) and right glottal area (rGA): The unit for these three parameters was square-pixel. The left/right glottal area was the area of the glottis on the

left/right hand side of the vertical midline of the “analysis window” added earlier.

2. Length of the anterior (aLen) and posterior (pLen) of the glottis; width of the left (lLen) and right (rLen) glottis: The unit for these four parameters was pixels. The anterior/posterior length of the glottis referred to the length of the glottis measured from the center of the “window” to the anterior/posterior end of the glottis. The left/right width of the glottis referred to the width of the glottis measured from the center of the “window” to the left/right lateral tip of the glottis.

Based on the glottal area measured from (1), each frame was determined by HSVP to belong to either a maximum glottal opening category, minimum glottal opening category or neither of the two. The HSVP program automatically traces the position of the glottis in each image frame and performs calculations on the measurements outlined above even when the position of the glottis changes in the video image due to the movement of the laryngoscope.

*Selection of representative frames for analysis.* Each image consisted of 4183 frames covering a span of 2.184 seconds. After a preliminary inspection of 15 video files, the investigator determined that it was sufficient to extract approximately 0.5 seconds (1000 frames) of the image files with the anterior and posterior tip and the lateral edges of the vocal folds clearly visible. Therefore, it was decided that 1000 stable frames were first to be extracted from each video file. In these 1000 stable frames, 100 vibratory cycles were used for analysis. If there were less than 100 cycles in the 1000 frames (usually this happened in male modal and fry production as a result of low fundamental frequency), 1500 stable frames were first selected for the initial procedure.

1. Frames that did not have the four landmarks (anterior and posterior tip and the lateral edges of the vocal folds) clearly visible were excluded.

2. Frames with the vocal folds that were more than 5mm away from the midline of the frame were excluded.
3. The remaining frames (i.e. images with the four laryngeal landmarks clearly visible and the vocal folds within 5 mm from the midline) were examined once again using steps 1 and 2 described above. Then the middle 1000 frames from the frames selected after step 3 were extracted. Figure 2 is a schematic diagram illustrating the selection procedure.
4. From the 1000 selected frames, the frames that contained the middle 100 cycles were used for analysis. They therefore contained 100 frames of complete glottal closure and 100 frames of maximum glottal opening.

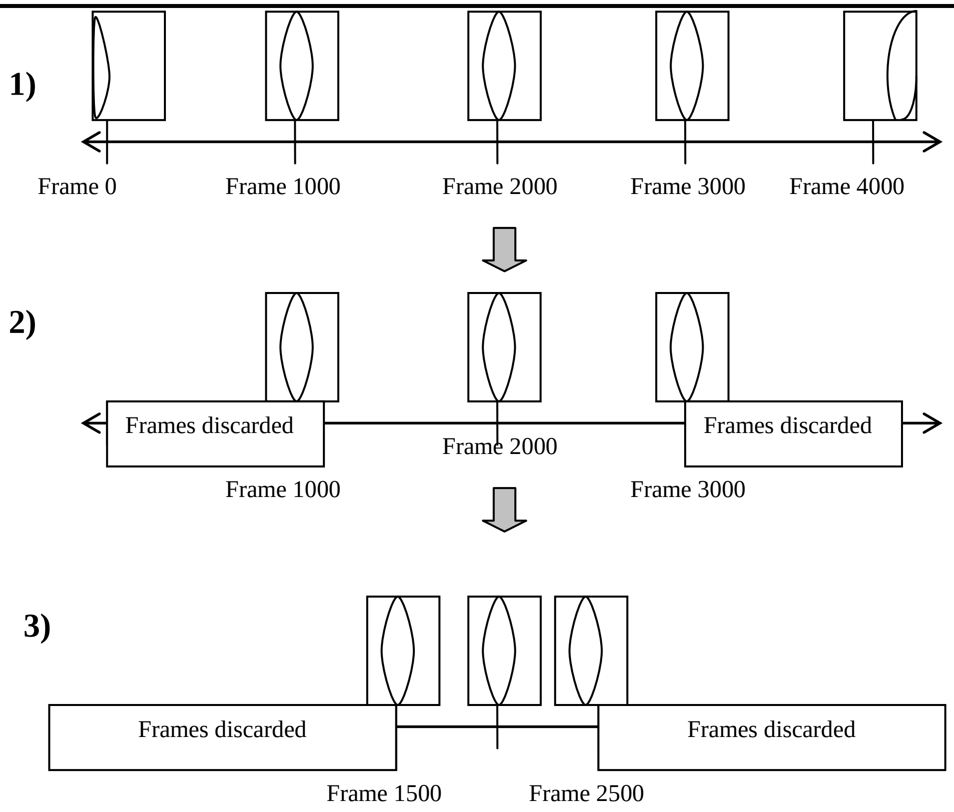


*Figure 2.* Procedures for selecting 1000 representative frames

## *Data Analysis*

*Data analysis for data from the three vocal registers.* As the distance between the tip of the laryngoscope and the vocal folds was unknown, absolute measurements of the glottal area and lengths were not possible because of the unknown magnification factor. Ratio index between the minimum and maximum glottal opening measures was used to represent the glottal measurements for each vibratory cycle. Therefore, only frames with the minimum glottal opening and the maximum glottal opening frames were used in the analysis. The other frames were not used in the data analysis procedure because the number of frames per cycle in each subject was different, making it difficult to draw comparisons across subjects if frames other than the maximum opening and closing frames were used.

The Ratio Index was calculated as follows:

$$\frac{\text{Average of the measurement from the minimum glottal opening (100 cycles)}}{\text{Average of the measurement from the maximum glottal opening (100 cycles)}}$$

Theoretically, frames containing smallest glottal opening (i.e. closed glottis) should demonstrate zero area and zero length. However, as the program used pixels for measurement, noise (hot pixels) were included and therefore, the pixel would never be zero in image processing.

The maximum value for each index should be approaching 1 as the measurement of the minimum glottal opening is always smaller than the measurement of the maximum glottal opening. A smaller index represents a larger glottal opening in each vibratory cycle. Nine ratio indices were calculated for each video image set. Details of each index are given below:

1. Glottal opening area ratio index: This index represented the glottal area of the maximum opening frames in relation to the corresponding minimum opening frames.
2. Left/Right glottal opening area ratio index: This index represented the left/right glottal area measured from the vertical midline to the lateral edges of the glottis. Asymmetry of the vocal folds vibration can be inferred from significant differences among these two indices.

3. Glottal opening width ratio index: This index represented the width of the glottis of the maximum opening frames in relation to the corresponding minimum opening frames.
4. Left/Right glottal opening width ratio index: This index represented the left/right glottal width measured from the vertical midline to the lateral edges of the glottis. Asymmetry of the vocal folds vibration can be inferred from significant differences among these two indices.
5. Antero-posterior glottal opening ratio index: This index represented the length of the glottis of the maximum opening frames in relation to the corresponding minimum opening frames.
6. Anterior/Posterior glottal opening ratio index: This index represented the anterior/posterior glottal length measured from the horizontal midline to the anterior/posterior edges of the glottis. Significant differences of these two indices would indicate the anterior and posterior parts of the vocal folds were asymmetrical when the glottis opening is at its maximum.

*Data analysis for data from the pitch glides.* The change of frequency across the pitch glides was analyzed acoustically using the PRAAT program (Institute of Phonetic Sciences, University of Amsterdam; version 4.3.04). A preliminary inspection of the 15 audio files of pitch glides showed that 33% (5/15) of the subjects were not able to glide one octave below their highest frequency within the 2.184 seconds limit of the images. Therefore, it was determined that the suitable length of the images for further image analysis should cover half an octave decrease from the highest frequency.

The images were then processed by the HSVP program and the data of three selected measurements: 1) glottal area; 2) width of the glottis; 3) length of the glottis was obtained. The corresponding ratio indices were calculated for each cycle: 1) glottal opening area ratio index; 2) glottal opening width ratio index; 3) Antero-posterior glottal opening ratio index. As the indices changed across the half octave decrease in a non-linear manner, regression lines using quadratic and cubic models were used to represent the changes across the half octave decrease. The lines of regression were compared across subjects. Figure 3 is the flowchart detailing the procedures of the analysis of pitch glides.

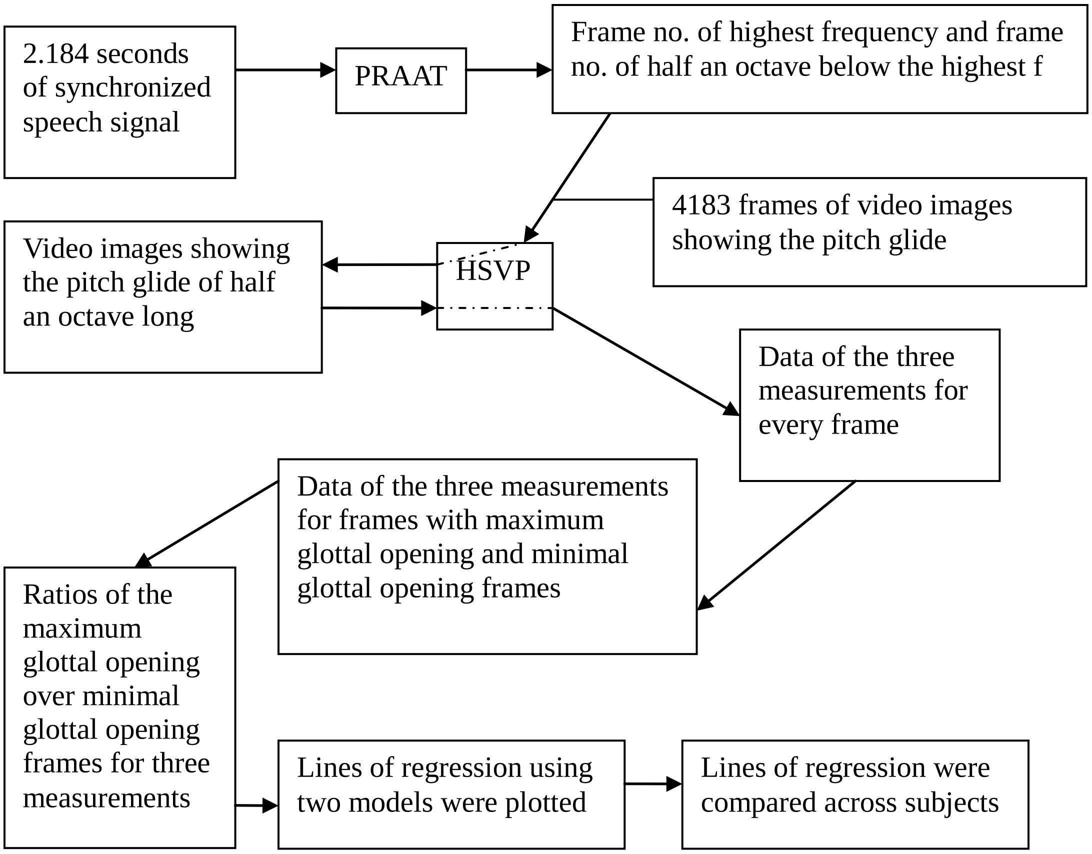


*Figure 3.* Flowchart illustrating the data processing of the pitch glides

*Inter/Intra-rater reliability*

As the ratio indices calculated were dependent on extracting 1000 stable frames and placing the “analysis window” using the HSVP program, reliability in placing the analysis window was determined by re-analyzing 15% (8) of the video images. Intra-rater reliability was established by re-analyzing 15% of the video images by the investigator (RF) after two weeks. Inter-rater reliability was established by re-analyzing 15% of the video images by a second person (a speech and language pathology student).

# Results

## *Reliability on frame selection (1000 stable frames) and placing the analysis window*

Table 1 shows the intra- and inter- rater reliability measures on frames selection. For placing the "analysis window", significant moderate correlation (>0.8) was observed for intra-rater reliabilities in eight out of nine indices, and three out of nine indices had significant correlation (>0.7) for inter-rater reliabilities (see Table 2).

Table 1. *Intra-rater and inter-rater reliability on frames selection*

| | Same segment | ±500 frames | ±1000 frames | Total |
|---|---|---|---|---|
| Intra-rater | 87.50% (7/8) | 12.50% (1/8) | 0.00% (0/8) | 100% (8/8) |
| Inter-rater | 62.50% (5/8) | 12.50% (1/8) | 25.00% (2/8) | 100% (8/8) |

Table 2. *Intra-rater and inter-rater reliability on nine indices (Valid cases=8)*

| Ratio indices | Intra-rater reliability Spearman's Rho | Inter-rater reliability Spearman's Rho |
|---|---|---|
| Glottal opening area ratio index | 1.00** | 0.85** |
| Left glottal opening area ratio index | 0.90** | 0.93** |
| Right glottal opening area ratio index | 1.00** | 0.71* |
| Glottal opening width ratio index | 0.61 | 0.27 |
| Left glottal opening width ratio index | 0.93** | 0.58 |
| Right glottal opening width ratio index | 0.60 | 0.26 |
| Antero-posterior glottal opening ratio index | 0.95** | 0.48 |
| Anterior glottal opening ratio index | 0.95** | 0.48 |
| Posterior glottal opening ratio index | 0.86** | 0.40 |

** $p<0.01$ (2-tailed) * $p<0.05$ (2-tailed)

### *Mean ratio indices for the three vocal registers (modal, falsetto, fry) - normal subjects*

Table 3 lists the mean ratio indices. In all subjects (N=15), high speed images of the modal and falsetto registers were taken. However, for fry, only 7 subjects had high speed images successfully taken. For most ratio indices, modal register had the smallest values and fry had the largest values, values from the falsetto register usually fell in between modal and fry register.

A non-parametric test, Wilcoxon signed rank test, was used to compare the indices from the three vocal registers. Wilcoxon signed rank test was chosen because the normality for parametric tests was violated, as the indices were not in a normal distribution. The modal and falsetto register showed similar ratio indices in most cases. The anterior glottal opening ratio index was however significantly smaller in the modal register when compared to the falsetto register. The modal register showed significantly smaller values in all ratio indices except with the left glottal opening width ratio index than the fry register. The vocal fry register also demonstrated significantly higher values in six of the ratio indices than the falsetto register (see Table 3).

### *Symmetry measurements of left-right (area and width) and antero-posterior (length) indices*

The left-right symmetry of the glottis was investigated by comparing the left and right glottal opening area/width ratio indices using Wilcoxon signed rank test. The left and right glottal opening area ratio indices showed no significant differences in all three registers. However, the left and right glottal opening width ratio indices showed a significant difference in the modal register (See Table 4).

The antero-posterior symmetry of the glottis was investigated by comparing the anterior and posterior glottal opening ratio indices using Wilcoxon signed rank test. In all three vocal registers, the two indices showed no significant differences (See Table 4).

Table 3. *Mean, standard deviation and Wilcoxon signed ranked tests for the nine indices of the three vocal registers (modal, falsetto, fry) from normal subjects*

| | Mean and Standard Deviation (SD) | | | Wilcoxon signed rank test | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Modal (n=15) | Falsetto (n=15) | Fry (n=7) | Modal (n=15)-Falsetto (n=15) | | Modal (n=7) -Fry (n=7) | | Falsetto (n=7) -Fry (n=7) | |
| Ratio indices | Mean (SD) | Mean (SD) | Mean (SD) | Z | p | Z | p | Z | p |
| Glottal opening area ratio index | 0.27 (0.15) | 0.36 (0.16) | 0.57 (0.21) | -1.42 | 0.16 | -2.37 | 0.02* | -2.20 | 0.03* |
| Left glottal opening area ratio index | 0.27 (0.15) | 0.36 (0.17) | 0.57 (0.24) | -1.31 | 0.19 | -2.37 | 0.02* | -2.20 | 0.03* |
| Right glottal opening area ratio index | 0.27 (0.18) | 0.35 (0.18) | 0.60 (0.20) | -1.59 | 0.11 | -2.37 | 0.02* | -2.37 | 0.02* |
| Glottal opening width ratio index | 0.27 (0.26) | 0.33 (0.20) | 0.64 (0.36) | -1.08 | 0.28 | -2.20 | 0.03* | -2.03 | 0.04* |
| Left glottal opening width ratio index | 0.30 (0.27) | 0.31 (0.23) | 0.44 (0.45) | -0.40 | 0.69 | -1.35 | 0.18 | -1.01 | 0.31 |
| Right glottal opening width ratio index | 0.25 (0.31) | 0.32 (0.22) | 0.86 (0.16) | -1.31 | 0.19 | -2.20 | 0.03* | -2.20 | 0.03* |
| Antero-posterior glottal opening ratio index | 0.36 (0.16) | 0.41 (0.16) | 0.63 (0.23) | -1.82 | 0.07 | -2.37 | 0.02* | -1.86 | 0.06 |
| Anterior glottal opening ratio index | 0.33 (0.16) | 0.50 (0.22) | 0.73 (0.24) | -1.99 | 0.05* | -2.37 | 0.02* | -2.20 | 0.03* |
| Posterior glottal opening ratio index | 0.40 (0.30) | 0.44 (0.24) | 0.59 (0.28) | -0.63 | 0.53 | -2.03 | 0.04* | -1.86 | 0.06 |

* $p<0.05$ (2-tailed)

Table 4. *Wilcoxon signed rank test of left-right and antero-postero indices*

| | Modal | | Falsetto | | Fry | |
|---|---|---|---|---|---|---|
| | Z | p | Z | P | Z | p |
| Area (Left-Right) | -0.11 | 0.91 | -0.62 | 0.53 | -0.51 | 0.61 |
| Width (Left-Right) | -2.27 | 0.02* | -0.47 | 0.64 | -1.57 | 0.12 |
| Length (Antero-posterior) | -0.85 | 0.39 | -0.74 | 0.46 | -1.18 | 0.24 |

* $p < 0.05$ (2-tailed)

### *Comparison of individual dysphonic subjects with the normal subjects group*

The number of dysphonic subjects included in the study was too few (N=2) to allow the conduction of inferential statistical tests to investigate indices that could be used to distinguish normal and dysphonic subjects. Data of the dysphonic subjects were compared individually using the 95% confidence interval level (see Tables 5 and 6) of the nine indices from normal subjects. In both dysphonic subjects, 12 out of 27 (3 registers with 9 indices each) indices laid outside the 95% confidence interval of normal subjects.

Table 5. *95% confidence interval level of the nine ratio indices from normal subjects*

| | Modal | | Falsetto | | Fry | |
|---|---|---|---|---|---|---|
| | Lower Boundary | Upper Boundary | Lower Boundary | Upper Boundary | Lower Boundary | Upper Boundary |
| Glottal opening area ratio index | 0.12 | 0.41 | 0.17 | 0.34 | 0.33 | 0.80 |
| Left glottal opening area ratio index | 0.15 | 0.36 | 0.17 | 0.31 | 0.29 | 0.85 |
| Right glottal opening area ratio index | 0.06 | 0.48 | 0.11 | 0.45 | 0.37 | 0.82 |
| Glottal opening width ratio index | 0.01 | 0.31 | 0.05 | 0.36 | 0.21 | 1.01 |
| Left glottal opening width ratio index | 0.07 | 0.36 | -0.01 | 0.33 | 0.04 | 0.99 |
| Right glottal opening width ratio index | -0.16 | 0.58 | 0.02 | 0.39 | 0.70 | 1.03 |
| Antero-posterior glottal opening ratio index | 0.13 | 0.53 | 0.24 | 0.47 | 0.37 | 0.90 |
| Anterior glottal opening ratio index | 0.17 | 0.54 | 0.23 | 0.73 | 0.44 | 0.98 |
| Posterior glottal opening ratio index | 0.05 | 0.60 | 0.14 | 0.33 | 0.33 | 0.93 |

Table 6. *Nine ratio indices of the two dysphonic subjects*

| | Dysphonic subject 1 | | | Dysphonic subject 2 | | |
|---|---|---|---|---|---|---|
| | Modal | Falsetto | Fry | Modal | Falsetto | Fry |
| Glottal opening area ratio index | 0.42* | 0.35* | 0.91* | 0.16 | 0.17 | 0.25* |
| Left glottal opening area ratio index | 0.37* | 0.44* | 0.86* | 0.16 | 0.10* | 0.24* |
| Right glottal opening area ratio index | 0.51* | 0.25 | 0.92* | 0.16 | 0.26 | 0.27* |
| Glottal opening width ratio index | N/A | N/A | 0.98 | 0.04 | 0.02* | 0.01* |
| Left glottal opening width ratio index | N/A | N/A | 1.05* | 0.04 | 0.01 | 0.02* |
| Right glottal opening width ratio index | N/A | N/A | 0.96 | 0.04 | 0.03 | N/A |
| Antero-posterior glottal opening ratio index | 0.64* | 0.30 | 0.91* | 0.33 | 0.39 | 0.21* |
| Anterior glottal opening ratio index | 0.49 | 0.37 | 0.95 | 0.37 | 0.18* | 0.25* |
| Posterior glottal opening ratio index | 0.81* | 0.26 | 0.88 | 0.26 | 0.59* | 0.17* |

| | |
|---|---|
| * | Laid outside the 95% confidence interval of normal subjects |
| N/A | The measurements could not be computed by the HSVP program |

### *Descriptive data from pitch gliding*

The half an octave decrease represents a mean decrease of 118.2 Hz in frequency (Standard deviation=28.4 Hz; range=66.4-165.1 Hz) from the maximum frequency of each individual normal subject (N=15). Lines of regression using quadratic and cubic models were plotted to represent the changes of the three selected indices (glottal opening area ratio index, glottal opening width ratio index and antero-posterior glottal opening ratio index) for each subject (See Appendices 1-15).

From the scatterplots of each subject, the changes of the three indices within each individual subject were similar. The change of the glottal opening area ratio index was especially similar to the antero-posterior glottal opening ratio index in most subjects.

Across subjects, no dominating pattern of changes of these three indices was found. In six of the subjects (Subject 5, 8, 9, 12, 13, 15), the three indices were generally increasing as they pitch glided downwards; four of them (Subject 2, 3, 5, 6) had indices that were generally decreasing. Subject 14 (RC) was a special case; the three indices were almost constant throughout the pitch glide. For the remaining four (Subject 1, 4, 10, 11) the changes of indices were in a highly variable manner throughout the pitch glide.

## Discussion

The aim of the study was to investigate quantitative measurements in high speed laryngoscopic images that can be used to describe normal voices. In general, the ratio indices of the fry register were significantly different from that of modal and falsetto registers.

### *Reliability*

High intra-rater reliability (Spearman's Rho > 0.8 for seven of the nine ratio indices, see Table 2) was found for the quantitative data generated from the high speed images. The inter-rater reliability was only moderate as only three out of the nine indices had a Spearman's Rho greater than 0.7 (see Table 2). The reliabilities were determined for two procedures: 1) Selection of the 1000 stable frames; 2) Precision of manually adding the "analysis window" to surround the glottis. The selection of 1000 frames should be less critical as the position of the glottis and the clarity of the images was quite uniform throughout the 2.184 seconds of the original video images. Therefore, although the frames chosen might differ across raters, as long as the view of the glottis was not obstructed throughout the 1000 frames, the effect on the index measurements generated from frames from different raters should not be too variable.

The procedure for adding the "analysis window" was more critical. To add the analysis window precisely, the rater needed to place the window surrounding the glottis at a frame of maximum glottal opening. When the images were not clear, the rater would need to make some judgment on the lateral edges and the anterior/posterior tips of the glottis subjectively in order to add the analysis window. Since the unit of the measurements was “pixel”, even a slight deviation of the placement of the window would lead to a big change in the final pixel measurement.

Reliability measures show that intra-rater reliabilities were relatively better than inter-rater reliabilities. To increase the inter-rater reliabilities, more training perhaps should have been given for identifying the landmarks of the glottis and also in consistency in placing the analysis window. A zooming function could also be added to future versions of the HSVP program so that the placement of the window can be more precise.

### *Quantitative data from high speed images on the three common vocal registers*

Comparison between ratio indices of the three vocal registers showed the vocal fry register was significantly different from the other two, while indices from the modal and falsetto registers were not significantly different (see Table 3).

*Modal-Falsetto Comparison.* Laver (1991) stated that in falsetto voice production, medial compression of the glottis is large and the glottis becomes thin but also extended longitudinally, resulting in a similar glottis size between modal and falsetto register. This physiological correlate is shown by insignificant difference between the three glottal opening area ratio indices of modal and falsetto registers. The glottal opening width ratio index should reflect the difference of medial compression between modal and falsetto registers. From the study, the difference was not significant (Wilcoxon signed rank test, p=0.28, 2-tailed). The inability of the index to reflect this physiological correlate could be due to the high variability in adding the analysis window on the lateral plane of the glottis. The values for these three glottal opening width ratio indices were dispersed (see Table 3). For modal register, the standard deviation of the above ratio indices was almost as big as the mean, while for falsetto register, the standard deviation was about 2/3 of the values of the mean. This suggests that the three ratio indices were variable and was probably a direct result from the variable procedure of adding the “analysis window”. Moreover, inter- and intra-rater reliabilities were especially low for these three indices (see Table 2).

*Fry-Modal; Fry-Falsetto comparison.* The vocal fry register is characterized by strongly compressed vocal folds and the ventricular folds are also adducted in most cases of fry production (Laver, 1991). In other words, the glottal area, width and length should be significantly smaller in fry production than in the other two vocal registers. Z-scores (Z=-2.37 for modal-fry comparison; Z=-2.20 for falsetto-fry comparison) obtained in comparisons of the glottal opening area ratio indices were larger in magnitude than the other six indices (i.e. length and width indices). Thus, the three glottal opening area ratio indices appeared to be the most salient indices to represent this physiological correlate.

Strong Medial compression and ventricular folds adduction during the production of fry should lead to significant differences in the glottal opening width ratio indices when fry-modal and fry-falsetto comparisons were drawn. However, the high variability across individual subjects in the glottal opening width ratio index of fry (mean=0.64; standard deviation=0.36) led to insignificant differences when comparisons were made. Glottal opening width ratio indices were therefore not a good indicator for the different vibratory patterns across vocal registers.

In fry production, adduction of the ventricular folds is likely to block a clear view of the glottis longitudinally. As a result, the antero-posterior indices of fry production may not represent the true length of the glottis, but a representation of the length of the gap created by the adduction of the ventricular folds. Therefore, these three indices may not be suitable in describing the glottis length in fry production and consequently in using them to compare fry and other registers.

It should be noted that only seven subjects had the fry register successfully recorded.

Moreover, the quality of the video images was considered as less clear, with the glottis less visible to inspection for fry production. The difficulty in recording a highly visible glottis in fry production was due to the lowered position of the vocal folds in fry production, and also the medial compression of supraglottic structures during vocal fry production. Therefore, the interpretations of data from the fry register warrant cautions.

### *Symmetry between the left and right glottis*

In normal vocal fold vibratory patterns, the left and right glottis should be symmetrical, any asymmetry between the left and right glottis may indicate a phase difference between the vibratory patterns of the two vocal folds. A comparison between the left and right glottal area/width ratio indices of the three registers show that the ratio indices were not significantly different from each other except in the glottal opening width ratio index of the modal register. More investigation on the correlation between asymmetrical vocal fold vibration and this index could be done on dysphonic subjects.

### *Two case studies of dysphonic subjects*

No single ratio index was considered to be representative and predictive in distinguishing dysphonic subjects and normal subjects. The three glottal opening area ratio indices across registers were most salient in identifying dysphonic subjects than the glottal opening width/length ratio indices. In the two dysphonic subjects, 12 out of 18 glottal opening area ratio indices (2 subjects X 3 indices X 3 registers) laid outside the 95% confidence interval (C.I.) of normal subjects. For the glottal opening width ratio indices of both subjects, only 4 out of 12 indices (6 indices were not able to be computed from the HSVP) fell outside the 95% C.I., while for antero-posterior, anterior and posterior glottal opening ratio indices, 8 out of 18 indices were outside the 95% C.I. This suggests that the glottal area at maximum opening is a more salient indicator of dysphonia in using high speed images.

An interesting finding was found when the ratio indices from dysphonic subjects were compared with the 95% C.I. For subject 1 (NC), all indices were greater than the 95% C.I of normal subjects. While for subject 2 (CTT), all indices except for the posterior glottal opening ratio index, were smaller than the 95% C.I. of normal subjects. This shows that dysphonic subjects do not necessarily have absolute larger/smaller measurements of the glottis when compared with normal subjects. Two physiological correlates could account for this finding. For NC, the indices were greater than normal subjects. This suggests that the glottal area at maximum opening were smaller. The reduced size of the glottis leads to insufficient air turbulence passing through the glottis, resulting in reduced pitch and possible hoarseness. For CTT, smaller indices than normal subjects may suggest that the incomplete glottal closure at minimum glottal closing frames. Incomplete adduction of the vocal folds during the closed phase would lead to perceptual correlate of breathiness.

### *Descriptive data on pitch gliding*

The changes of the three selected indices (glottal opening area ratio index, glottal opening width ratio index and antero-posterior glottal opening ratio index) were highly variable among normal subjects. In the pitch glide production task, the subject should start phonating with high pitch and then gradually glide down at least half an octave. Depending on their starting frequency, some terminated their pitch glide with phonation of the modal register and some with the fry register (i.e. in the two male subjects). In other words, the glottal opening area ratio index, theoretically, should show a rather constant and relatively small value throughout the first portion of the glide (since the glottal area is similar in falsetto and modal phonation), if fry production is present, the glottal opening area ratio index should reach a larger value towards the end. In the 15 subjects, only three subjects showed a trend approximating this hypothesis (i.e. Subject 9, 10 and 12; see Appendices 9, 10, 12).

For the glottal opening width ratio index, theoretically, the value should be large at the beginning because medial compression of the glottis is large (thus width is small) for production of high pitch. As the subject glided down the pitch, the value should decrease as the glottal width becomes larger when compared to that of high pitch. If vocal fry was produced, the value should become larger still as medial compression of the glottis and adduction of ventricular folds are present in fry production. Two of the subjects showed such a trend (i.e. Subject 1 and 9; see Appendices 1 and 9).

For the antero-posterior glottal opening ratio index, theoretically, the value should be small at the beginning because the vocal folds are stretched in production of high pitch, creating a longitudinally extended glottis. As the subject pitch glided down to lower pitch, the glottal length should decrease across time, which would result in an increasing value of the index. Only three subjects displayed this trend (i.e. Subject 3, 7 and 15; see Appendices 3, 7, 15). In some subjects, the reverse trend was demonstrated (i.e. Subject 2, 3 and 7; see Appendices 2, 3, 7).

From the results, it is concluded that pitch gliding is also a highly variable phenomenon. Although the physiology, as reflected by the ratio indices representing the changes of the glottis size, may vary greatly among subjects, the production of the pitch glides were perceptually similar. In addition to the variable nature of pitch gliding, the variability in adding the "analysis window" may also contribute to the variability in results obtained.

However, the data analysis procedure applied on analyzing pitch glides is favorable in representing the changes of the indices of the three common registers across a specific timeframe. This method of analysis can be implemented when data from dysphonic subjects are analyzed for aperiodicity in vocal fold vibration. Plotting the changes of the indices across time may allow investigators to discover subtle variances in glottal dimensions during phonation.

### *Limitations and additional research*

The number of subjects who had their production of the vocal fry register recorded were fewer than modal and falsetto registers. Although significant differences were found between the vocal fry register and the other two registers, the small number of subjects hindered the validity of the interpretation of this finding. Further studies could focus on studying the unique nature of the vocal fry register using high speed laryngoscopy through a flexible laryngoscope.

The number of dysphonic subjects who had a complete profile of high speed images taken were too few (i.e. two). Therefore, no inferential statistics could be carried out. In order to establish normative data to distinguish the two groups of subjects, more data from dysphonic subjects are needed to carry out valid comparison. Further studies could focus on establishing normative data for dysphonic subjects.

Due to subject availability, only two male subjects were recruited in the study. Further study would be required to evaluate if gender difference exists in the ratio indices for both normal and dysphonic subjects.

In addition, the HSVP program is still in its early development stage. Improvements could be made to the program (i.e. adding a zooming function) in order to increase the inter- and intra-rater reliabilities in carrying out data analysis of the high speed images.

### *Clinical Implications*

The development of high speed laryngoscopy for clinical use is still at an exploratory stage. Qualitative investigation, as proposed by previous researchers (Lindestad et. al, 2001; Verdonck-de Leeuw et. al, 2001), would be more feasible in clinical use at this stage as the time taken for post-recording analysis is not as long as quantitative analysis. If the parameters are well-established and the normative data of normal and dysphonic subjects are present in the future, a computer program can be written to compute the various indices from the high speed images. However, this would require a large pool of normal and dysphonic subjects as the variability across subject is great and the inter-rater reliability is not at an acceptable level at this stage of development.

Moreover, during the stage of data collection, it was observed that only about one in every three person recruited had a complete profile of high speed images (three vocal registers and pitch glide). Some unsuccessful subjects showed a sensitive gag reflex while some had prominent supraglottic structures, limiting a clear view of the glottis. In dysphonic clients, compressed pathway to the glottis is more probable as their supraglottic and surrounding laryngeal structures may be swollen due to hyperfunctional vocal misuse and abuse. Thus, the difficulty in obtaining a complete profile of high speed images is predicted to be even greater for dysphonic subjects.

### *Conclusions*

Quantitative measures on analyzing high speed laryngoscopic images were established. Tentative conclusion was made about statistically significant difference found between ratio indices from fry and two other registers: modal and falsetto. No single ratio index was representative in differentiating dysphonic subjects from normal subjects as reflected from the two case studies. Moreover, data from pitch glides suggestted that this method of phonation is highly variable across subjects. The three indices selected in the analysis of pitch glides did not show similar changes across the half an octave decrease. Until sufficient normative data are established, high speed laryngoscopy should be confined to qualitative analysis in clinical use.

# References

Berry, D. A., Montequin, D. W., & Tayama, N. (2001). High-speed digital imaging of the medial surface of the vocal folds. *Journal of Acoustical Society of America, 110,* 2539-2547.

Colton, R. H., & Casper, J. K. (1996). *Understanding voice problems: A physiological perspective for diagnosis and treatment (2nd ed.).* Lippincott Williams & Wilkins, Maryland.

Eysholdt, U., Tigges, M., Wittenberg, T., & Proschel, U. (1996). Direct evaluation of high speed recordings of vocal fold vibrations. *Folia Phoniatrica et Logopaedica, 48,* 163-170.

Granqvist, S., & Lindestad, P. (2001). A method of applying Fourier analysis to high-speed laryngoscopy. *Journal of Acoustical Society of America, 110,* 3193-3197.

Hirose, H. (1988). High-speed digital imaging of vocal fold vibration. *Acta Otolaryngol Suppl., 458,* 151-153.

Kiritani, S., Hirose, H., & Imagawa, H. (1993). High-speed digital image analysis of vocal cord vibration in diplophonia. *Speech Communication, 13,* 23-32.

Koster, O., Marx, B., Gemmar, P., Hess, M. M., & Josef Kunzel H. (1999). Qualitative and quantitative analysis of voice onset by means of a multidimensional voice analysis system (MVAS) using high-speed imaging. *Journal of Voice, 13,* 355-374.

Laver, J. (1991). The gift of speech. *Papers in the analysis of speech and voice.* Edinburgh: Edinburgh University Press.

Lindestad, P., Hammearberg, B., & Larsson, H. (1998). Ventricular fold co-vibration in chronic laryngitis studied with high-speed imaging and perceptual-acoustic analysis- a case report. *Proceedings of IALP 1998.*

Lindestad, P., Sodersten, M., Merker, B., & Granqvist, S. (2001). Voice source characteristics in Mongolian “Throat Singing” studied with high-speed imaging technique, acoustic spectra and inverse filtering. *Journal of Voice, 15,* 78-85.

Mak, J. (2003). Quantitative measures of supraglottic activity. Bachelor of Science dissertation, The University of Hong Kong, Hong Kong.

Maurer, D., Hess, M., & Gross, M. (1996). High-speed imaging of vocal fold vibrations and larynx movements within vocalizations of different vowels. *Ann Otol Rhinol Laryngol, 105,* 975-981.

Mergell, P., Herzel, H., & Titze, I. R. (2000). Irregular vocal-fold vibration-High-speed observation and modeling. *Journal of Acoustical Society of America, 108,* 2996-3002.

Niimi, S., & Miyaji, M. (2000). Vocal fold vibration and voice quality. *Folia Phoniatrica et Logopaedica, 52,* 32-38.

Schutte, H. K., Svec, J. G.., & Sram, F. (1998). First results of clinical application of videokymography. *The Laryngoscope, 108,* 1206-1210.

Tigges, M., Wittenberg, T., Mergell, P., & Eysholdt, U. (1999). Imaging of vocal fold vibration by digital multi-plane kymography. *Computerized medical imaging and graphics, 23,* 323-330.

van As, C. J., Tigges, M., Wittenberg, T., Op de Coul, B. M. R., Eysholdt, U., & Hilgers, F. J. M. (1999). High-speed digital imaging of neoglottic vibration after total laryngectomy. *Archives of Otolaryngology-Head and Neck Surgery, 125,* 891-897.

Verdonck-de, L. I. M., Festen, J. M., & Mahieu, H. F. (2001). Deviant vocal fold vibration as observed during videokymography: The effect on voice quality. *Journal of Voice, 15,* 313-322.

## Notes for Appendices 1-15

Scatterplots and coefficients of lines of regression (quadratic and cubic models) to show the changes of the three ratio indices across half an octave decrease during pitch gliding

Note:

For lines of regression using the quadratic (Quad) model, the equation would be in the form of:

$$y=ax^2+bx+constant$$

For lines of regression using the cubic (Cub) model, the equation would be in the form of:

$$y=ax^3+bx^2+cx+constant$$

In the following tables, a, b, c and constant are listed for each lines of regression.

Key:

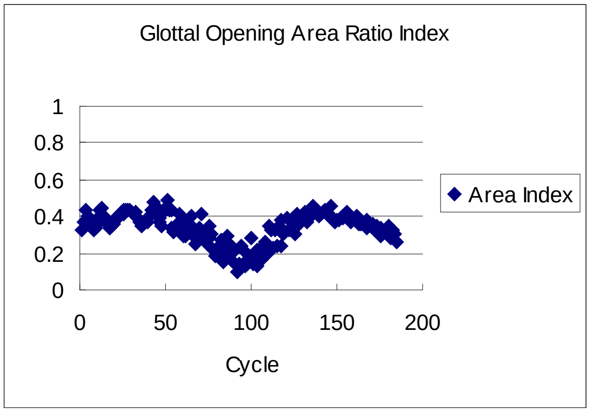
GA: Glottal opening area ratio index

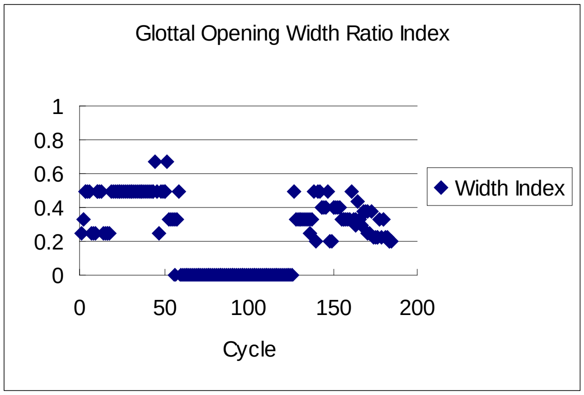
Width: Glottal opening width ratio index

Length: Antero-posterior glottal opening ratio index

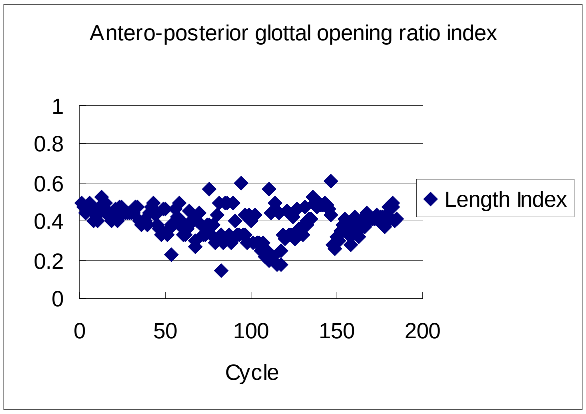
## Appendix 1

### Scatterplots and coefficients of lines of regression for Subject 1 (AC)

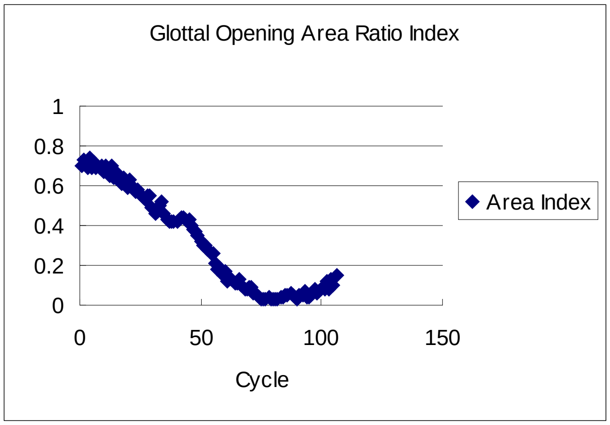


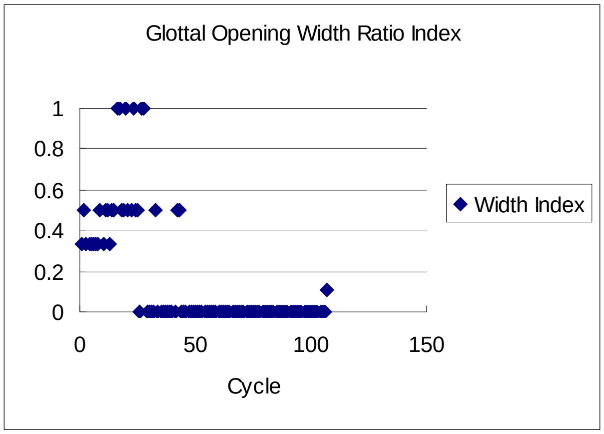


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.4338 | -0.0027 | 0.000013 | |
| GA | Cub | 0.4556 | -0.0041 | 0.000032 | -7E-08 |
| Width | Quad | 0.5891 | -0.0094 | 0.000047 | |
| Width | Cub | 0.5932 | -0.0097 | 0.00005 | -1E-08 |
| Length | Quad | 0.4777 | -0.0021 | 0.00001 | |
| Length | Cub | 0.4794 | -0.0022 | 0.000012 | -5E-09 |

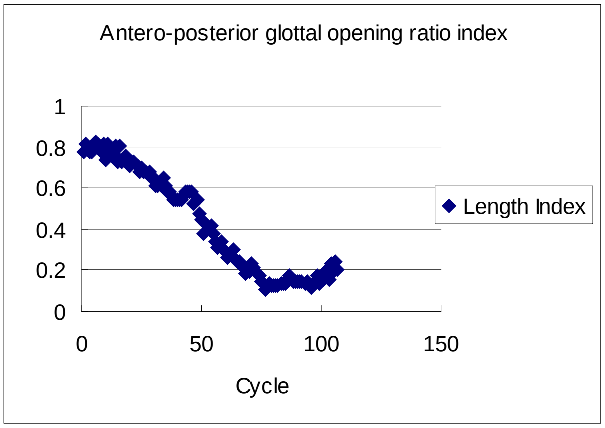


## Appendix 2
## Scatterplots and coefficients of lines of regression for Subject 2 (CT)

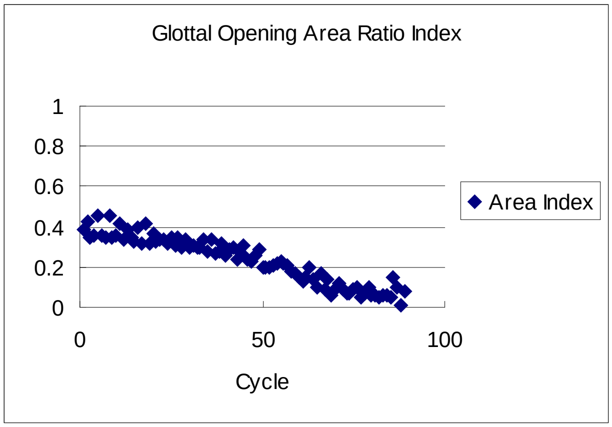


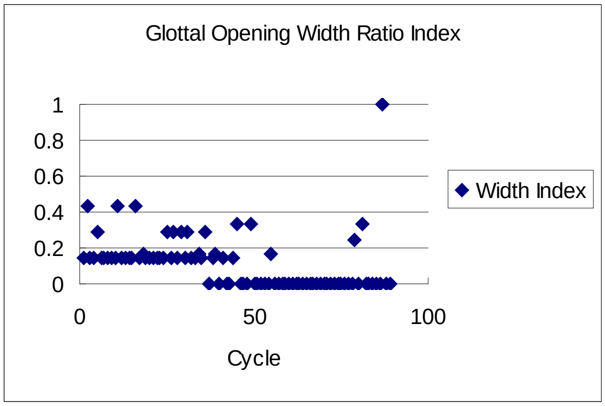


| Ratio Index | Method | constant a | b | c | |
|---|---|---|---|---|---|
| GA | Quad | 0.8426 | -0.0145 | 0.000064 | |
| GA | Cub | 0.7018 | 0.0008 | -0.0003 | 0.0000022 |
| Width | Quad | 0.6366 | -0.0153 | 0.000089 | |
| Width | Cub | 0.5174 | -0.0023 | -0.0002 | 0.0000018 |
| Length | Quad | 0.9346 | -0.0126 | 0.000044 | |
| Length | Cub | 0.7817 | 0.004 | -0.0003 | 0.0000024 |

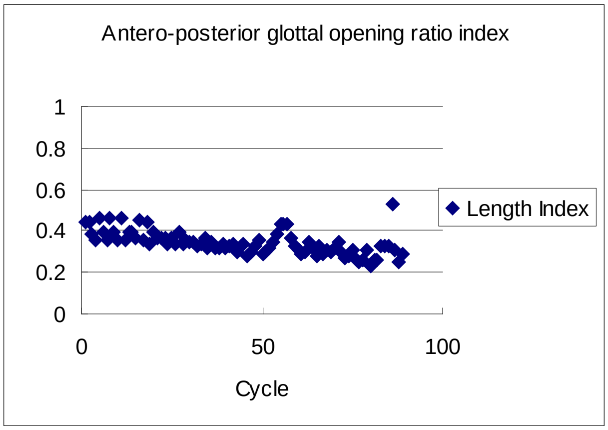


## Appendix 3
## Scatterplots and coefficients of lines of regression for Subject 3 (CA)

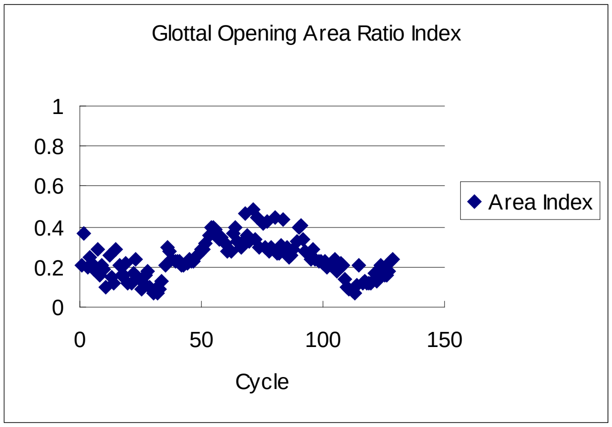


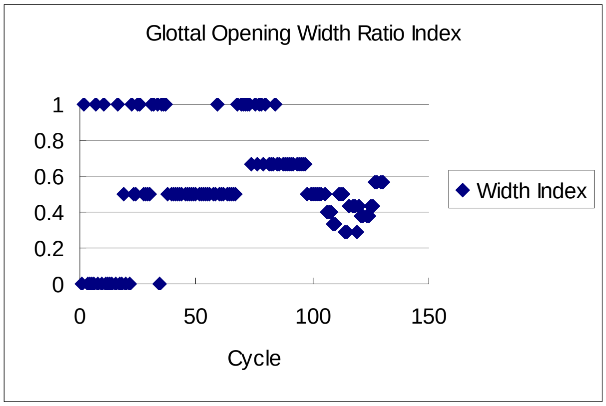


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.4095 | -0.0028 | -0.00002 | |
| GA | Cub | 0.3769 | 0.0015 | -0.0001 | 8.7E-07 |
| Width | Quad | 0.278 | -0.0069 | 0.000054 | |
| Width | Cub | 0.1578 | 0.0087 | -0.0004 | 0.0000032 |
| Length | Quad | 0.4191 | -0.0025 | 0.000014 | |
| Length | Cub | 0.4237 | -0.0031 | 0.000031 | -1E-07 |

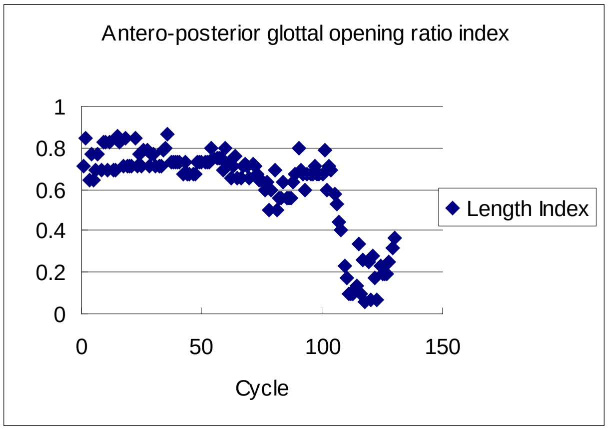


# Appendix 4
## Scatterplots and coefficients of lines of regression for Subject 4 (CF)

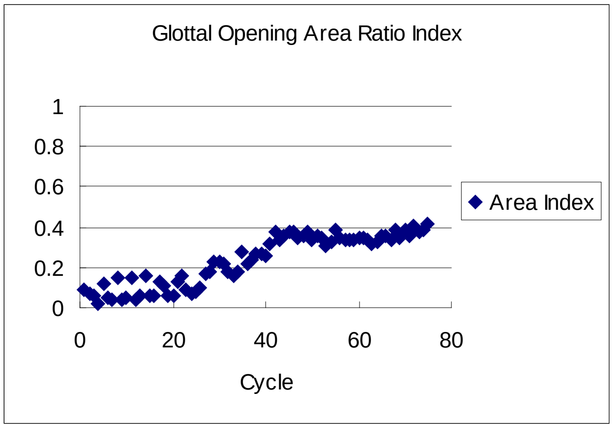


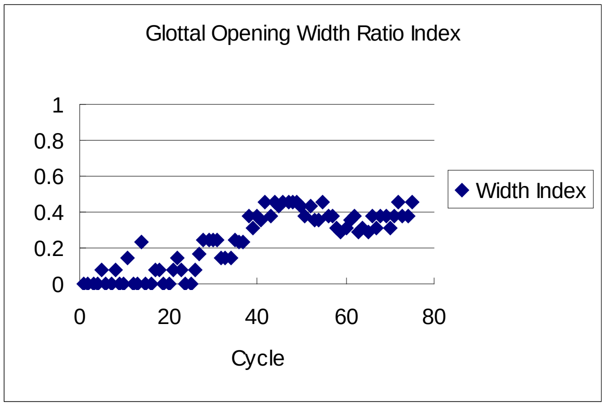


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.1028 | 0.0061 | -0.00005 | |
| GA | Cub | 0.1775 | 0.0003 | 0.00006 | -5E-07 |
| Width | Quad | 0.3935 | 0.009 | -0.00007 | |
| Width | Cub | 0.4233 | 0.0067 | -0.00003 | -2E-07 |
| Length | Quad | 0.6994 | 0.0046 | -0.00007 | |
| Length | Cub | 0.7954 | -0.0029 | 0.000067 | -7E-07 |

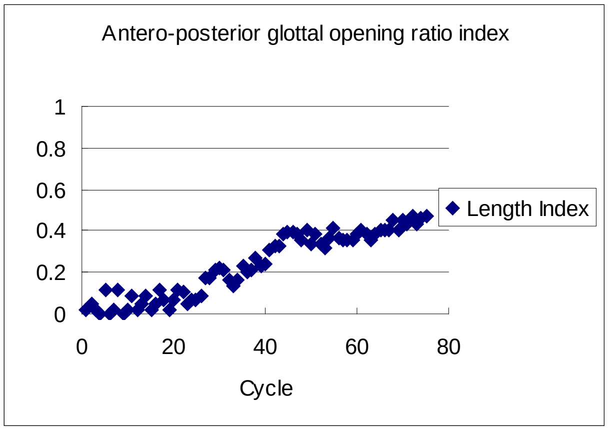


## Appendix 5
## Scatterplots and coefficients of lines of regression for Subject 5 (CN)

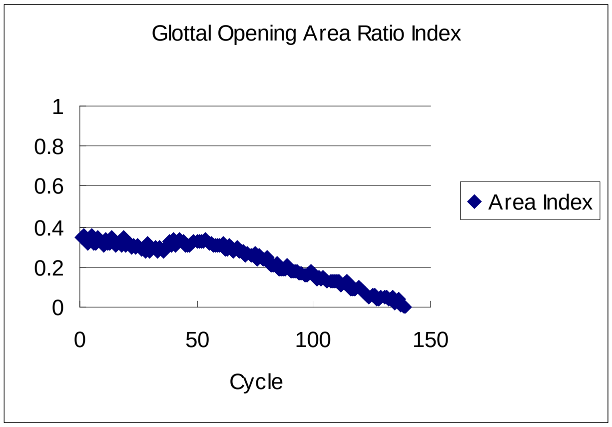


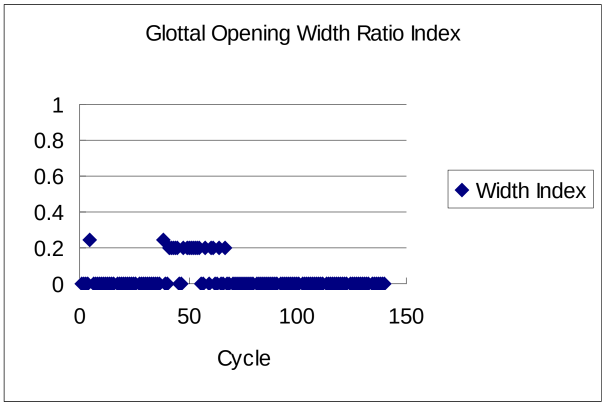


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.0002 | 0.008 | 0 | |
| GA | Cub | 0.0678 | -0.0024 | 0.0003 | 0 |
| Width | Quad | -0.0927 | 0.0134 | -0.0001 | |
| Width | Cub | 0.0003 | -0.0008 | 0.0004 | 0 |
| Length | Quad | -0.0319 | 0.0078 | 0 | |
| Length | Cub | 0.0363 | -0.0026 | 0.0003 | 0 |

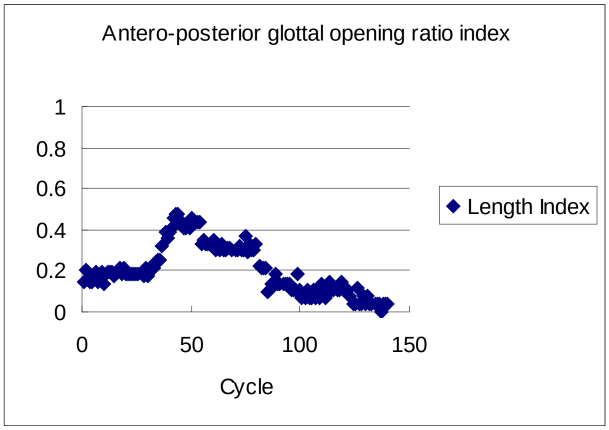


## Appendix 6
## Scatterplots and coefficients of lines of regression for Subject 6 (DC) *male



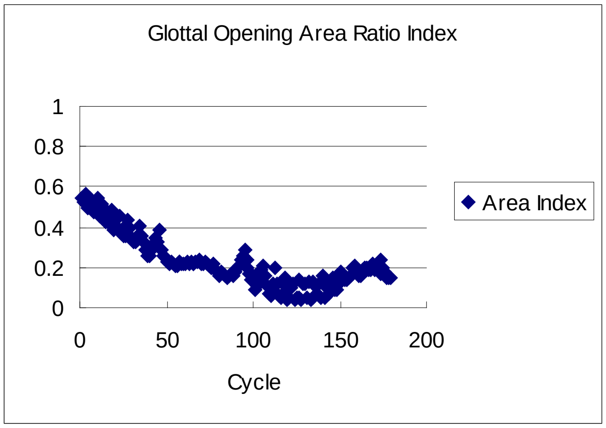


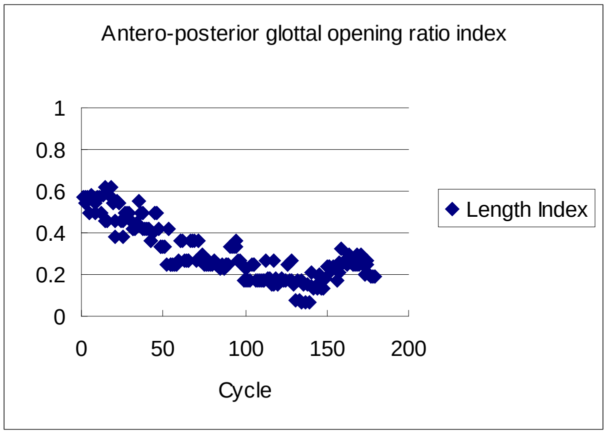
| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.3241 | 0.0006 | 0 | |
| GA | Cub | 0.3137 | 0.0015 | 0 | 7.3E-08 |
| Width | Quad | 0.0134 | 0.0015 | 0 | |
| Width | Cub | -0.0456 | 0.0064 | -0.0001 | 4.1E-07 |
| Length | Quad | 0.15 | 0.0054 | 0 | |
| Length | Cub | 0.0409 | 0.0145 | -0.0002 | 7.6E-07 |



# Appendix 7

## Scatterplots and coefficients of lines of regression for Subject 7 (DL)

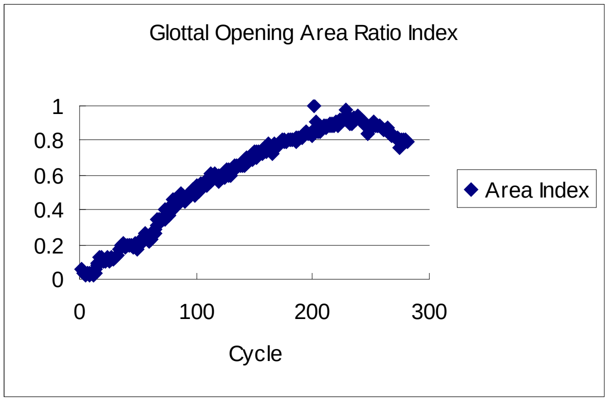


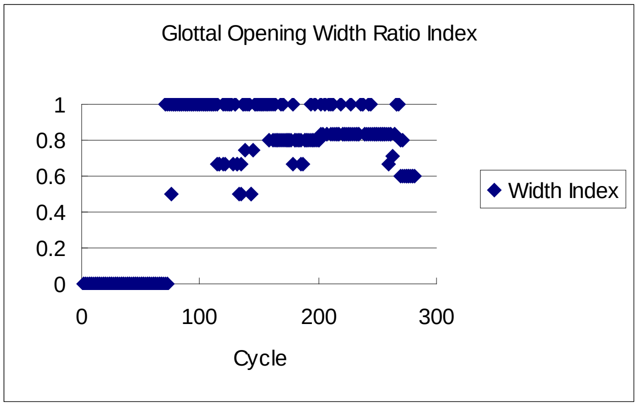


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.5571 | -0.007 | 0.000028 | |
| GA | Cub | 0.5458 | -0.0063 | 0.000018 | 3.8E-08 |
| Length | Quad | 0.6189 | -0.0064 | 0.000024 | |
| Length | Cub | 0.5863 | -0.0043 | -0.000006 | 1.1E-07 |

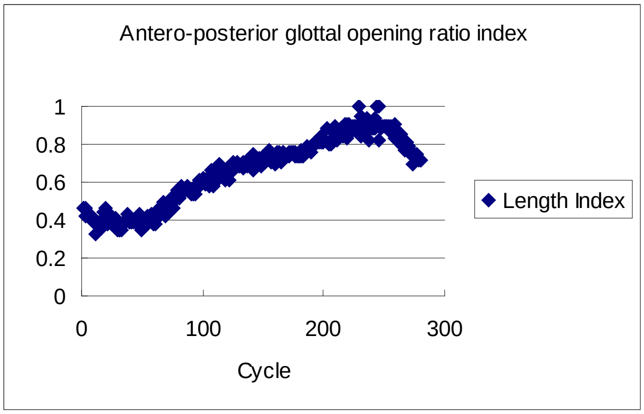
# Appendix 8

## Scatterplots and coefficients of lines of regression for Subject 8 (FL)



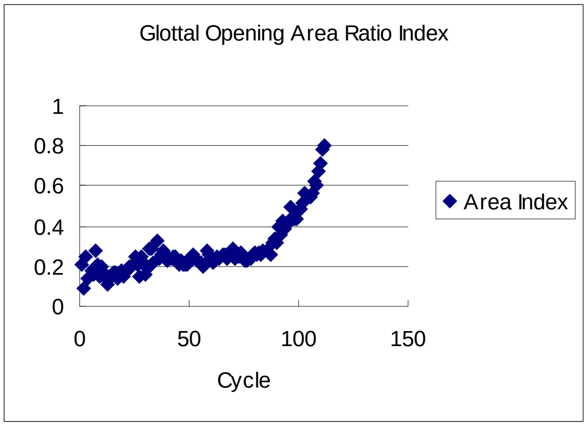


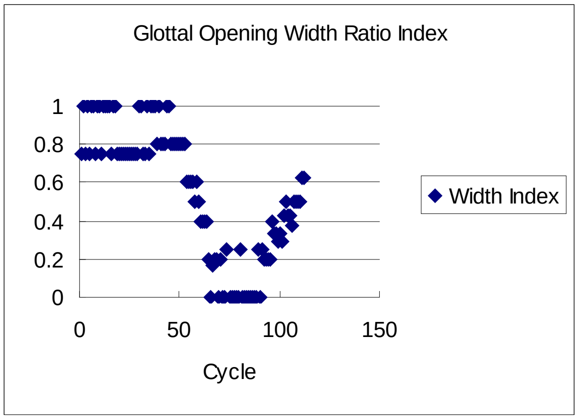
| Ratio Index | Method | constant a | b | c | |
|---|---|---|---|---|---|
| GA | Quad | -0.1854 | 0.0088 | -0.00002 | |
| GA | Cub | -0.0079 | 0.0045 | 0.000012 | -6E-08 |
| Width | Quad | -0.2339 | 0.0134 | -0.00004 | |
| Width | Cub | -1.1844 | 0.0369 | -0.0002 | 3.4E-07 |
| Length | Quad | 0.1247 | 0.0059 | -0.00001 | |
| Length | Cub | 0.3364 | 0.0007 | 0.000024 | -8E-08 |



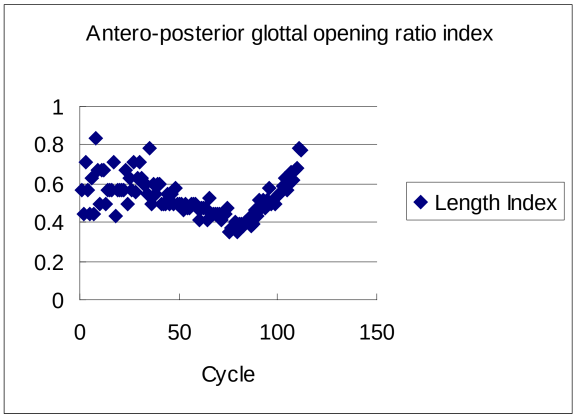
# Appendix 9

## Scatterplots and coefficients of lines of regression for Subject 9 (IC)

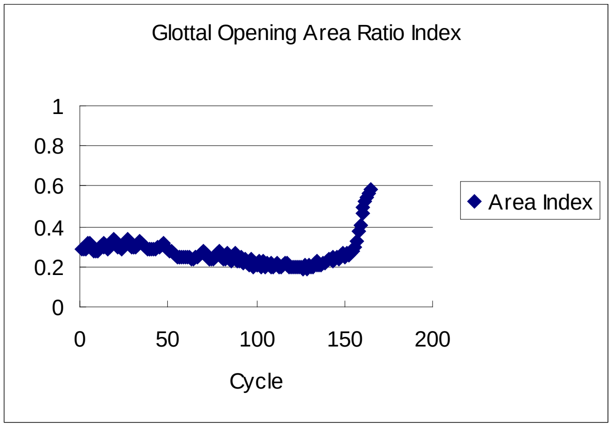


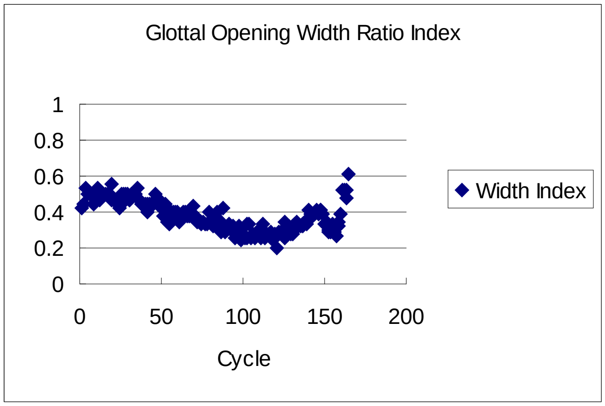


| Ratio Index | Method | constant a | b | c | |
|---|---|---|---|---|---|
| GA | Quad | 0.2371 | -0.004 | 0.000065 | |
| GA | Cub | 0.1077 | 0.0094 | -0.0002 | 0.0000017 |
| Width | Quad | 1.1406 | -0.0156 | 0.000069 | |
| Width | Cub | 0.6673 | 0.0335 | -0.001 | 0.0000064 |
| Length | Quad | 0.6883 | -0.0069 | 0.000054 | |
| Length | Cub | 0.5216 | 0.0104 | -0.0003 | 0.0000023 |

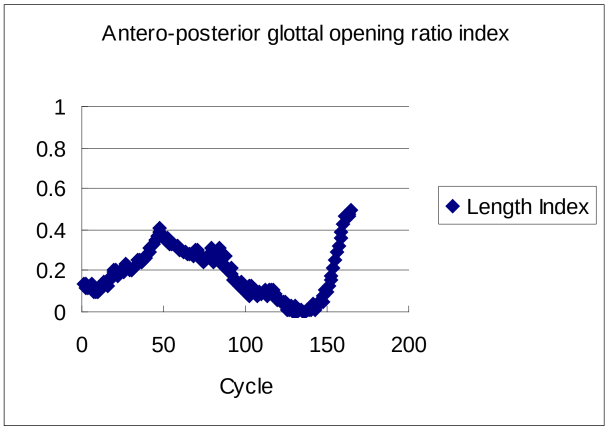


## Appendix 10
## Scatterplots and coefficients of lines of regression for Subject 10 (KT)

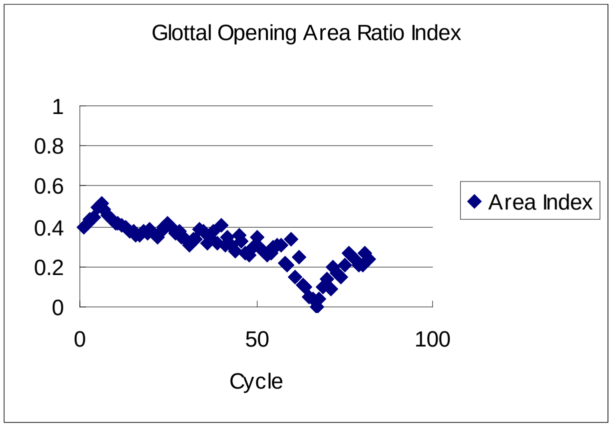


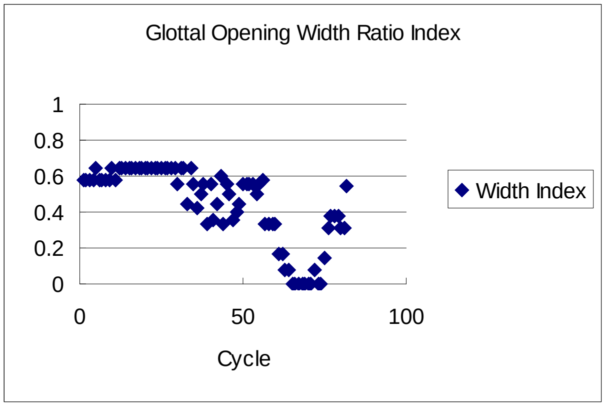


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.3707 | -0.0034 | 0.00002 | |
| GA | Cub | 0.26 | 0.0045 | -0.0001 | 4.8E-07 |
| Width | Quad | 0.5649 | -0.0045 | 0.000021 | |
| Width | Cub | 0.4737 | 0.002 | -0.00008 | 3.9E-07 |
| Length | Quad | 0.2208 | -0.0001 | 0 | |
| Length | Cub | -0.0528 | 0.0194 | -0.0003 | 0.0000012 |

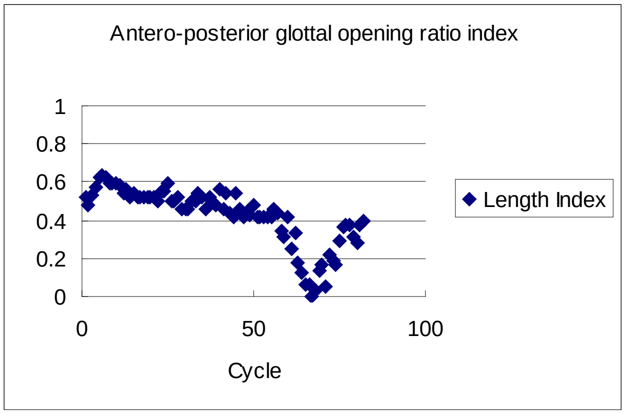


## Appendix 11
## Scatterplots and coefficients of lines of regression for Subject 11 (NC)

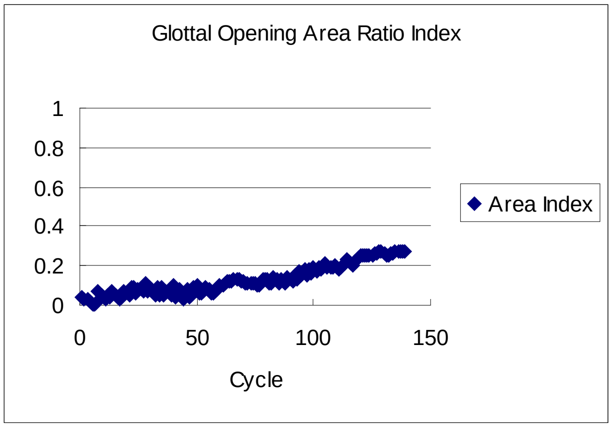


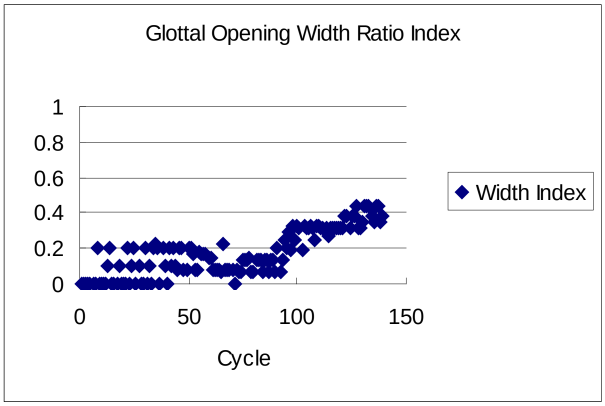


| Ratio Index | Method | constant a | b | c | |
|---|---|---|---|---|---|
| GA | Quad | 0.4696 | -0.0039 | -7E-08 | |
| GA | Cub | 0.4113 | 0.0042 | -0.0002 | 0.000002 |
| Width | Quad | 0.6583 | -0.002 | -0.00006 | |
| Width | Cub | 0.4673 | 0.0248 | -0.0009 | 0.0000064 |
| Length | Quad | 0.5942 | -0.0023 | -0.00003 | |
| Length | Cub | 0.4893 | 0.0124 | -0.0005 | 0.0000035 |

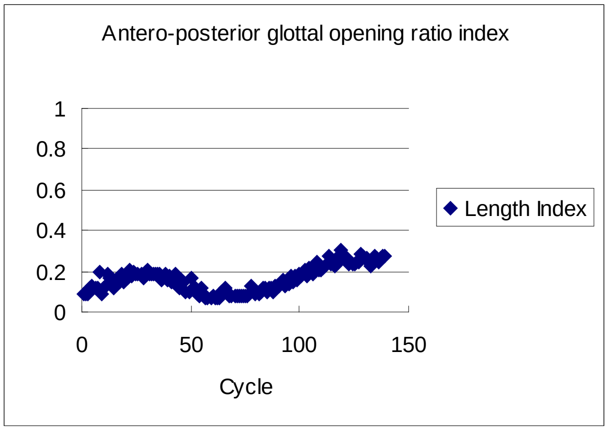


## Appendix 12
## Scatterplots and coefficients of lines of regression for Subject 12 (OC)

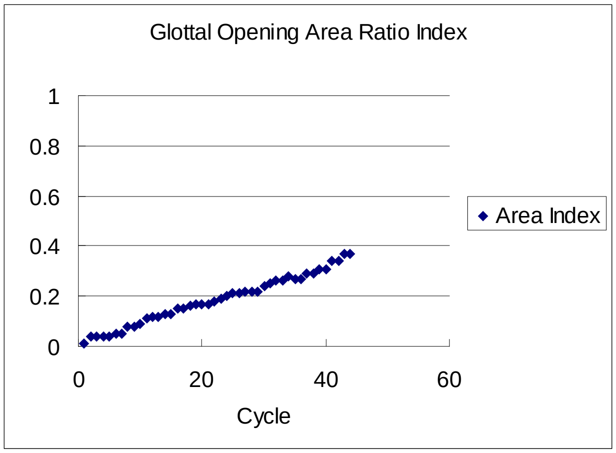


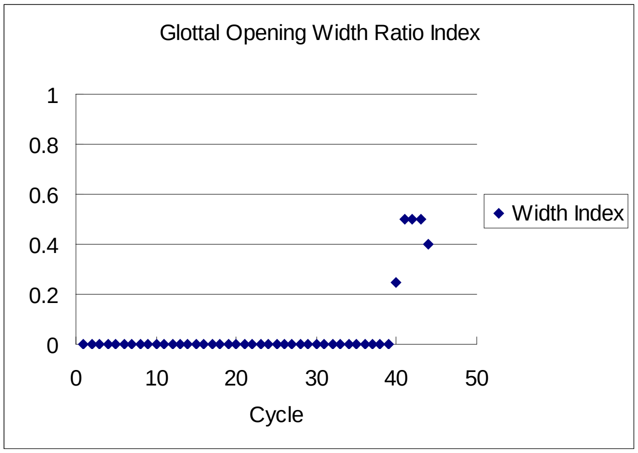


| Ratio Index | Method | constant a | b | c | |
|---|---|---|---|---|---|
| GA | Quad | 0.0418 | 0.0002 | 0.000011 | |
| GA | Cub | 0.0345 | 0.0008 | 4.3E-07 | 5.2E-08 |
| Width | Quad | 0.0537 | -0.0002 | 0.000021 | |
| Width | Cub | 0.0114 | 0.0033 | -0.00004 | 0.0000003 |
| Length | Quad | 0.1873 | -0.0026 | 0.000024 | |
| Length | Cub | 0.1553 | 0.000076 | -0.00002 | 2.3E-07 |

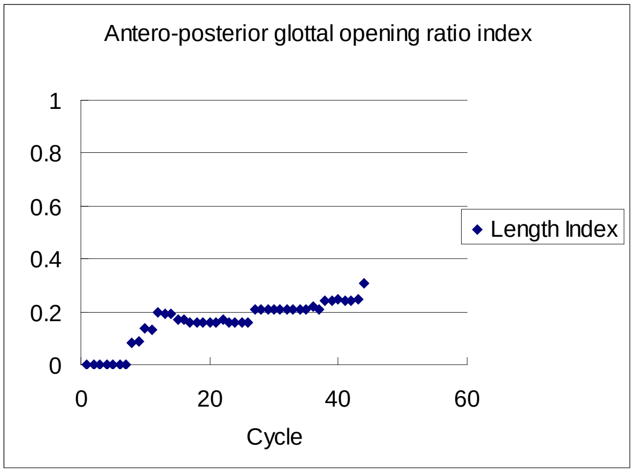


# Appendix 13
## Scatterplots and coefficients of lines of regression for Subject 13 (RL)*male

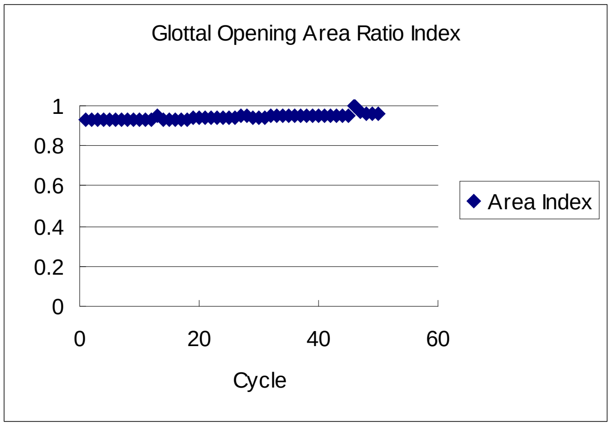


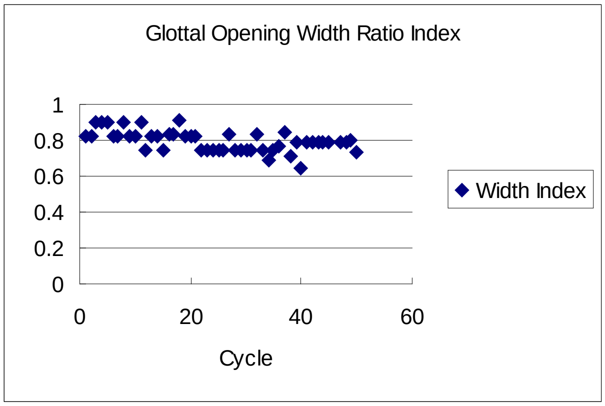


| Ratio Index | Method | constant a | b | c | |
|---|---|---|---|---|---|
| GA | Quad | 0.0138 | 0.0076 | 0.0000038 | |
| GA | Cub | -0.0034 | 0.0119 | -0.0002 | 0.0000036 |
| Width | Quad | 0.0985 | -0.018 | 0.0005 | |
| Width | Cub | -0.0841 | 0.0281 | -0.002 | 0.000038 |
| Length | Quad | -0.0136 | 0.0117 | -0.0001 | |
| Length | Cub | -0.069 | 0.0257 | -0.0009 | 0.000011 |

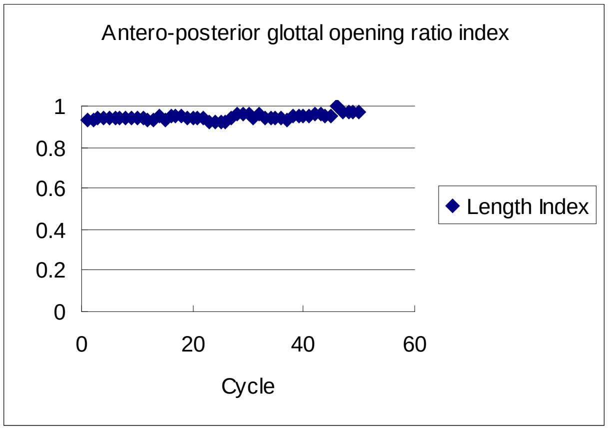


# Appendix 14
## Scatterplots and coefficients of lines of regression for Subject 14 (RC)

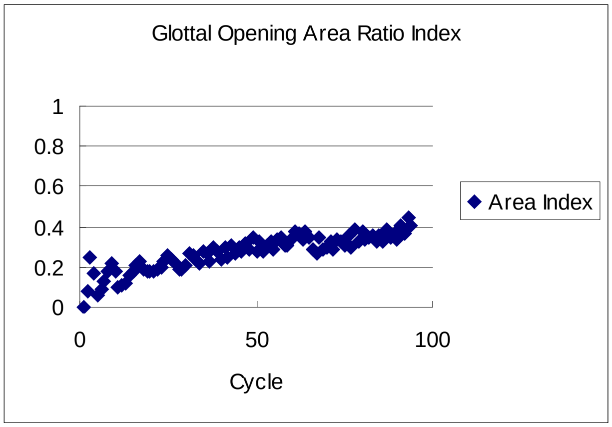


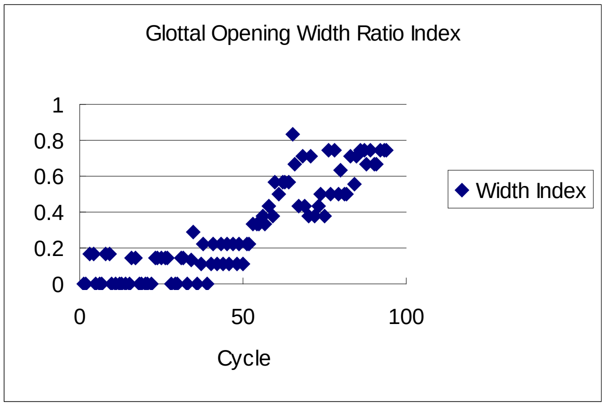


| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.9279 | 0.0004 | 0.0000073 | |
| GA | Cub | 0.9275 | 0.0005 | 0.0000029 | 5.8E-08 |
| Width | Quad | 0.8937 | -0.008 | 0.0001 | |
| Width | Cub | 0.8627 | -0.001 | -0.0002 | 0.0000044 |
| Length | Quad | 0.9421 | -0.001 | 0.000024 | |
| Length | Cub | 0.9363 | 0.0006 | -0.00004 | 8.3E-07 |

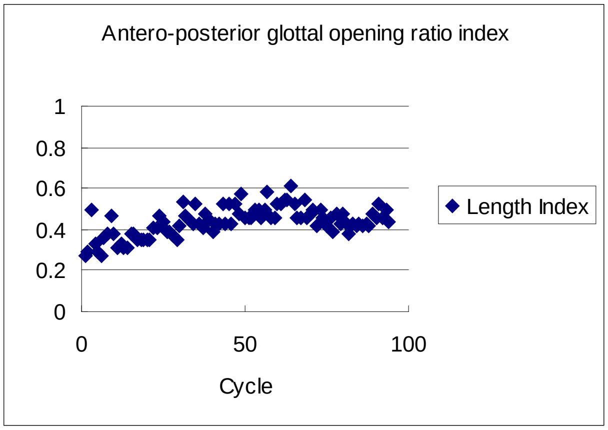


# Appendix 15
## Scatterplots and coefficients of lines of regression for Subject 15 (WF)





| Ratio Index | Method | constant | a | b | c |
|---|---|---|---|---|---|
| GA | Quad | 0.0988 | 0.0053 | -0.00003 | |
| GA | Cub | 0.0806 | 0.0076 | -0.00008 | 4.1E-07 |
| Width | Quad | 0.0069 | 0.0016 | 0.000075 | |
| Width | Cub | 0.1142 | -0.012 | 0.0004 | -0.000002 |
| Length | Quad | 0.2922 | 0.0062 | -0.00005 | |
| Length | Cub | 0.3026 | 0.005 | -0.00002 | -2E-07 |



# Acknowledgments

I would like to express my gratitude to Dr. Edwin Yiu for his enduring support and guidance throughout the course of this study. I would also like to thank Dr. Kong Jianping from Beijing University for his development and advice on the use of the HSVP program. I would like to thank Dr. Karen Chan, Ms. Elaine Kwong and Ms. Cecilia Fung for their great help in subject recruitment and data collection. I would also like to dedicate my gratefulness to Ms. Pang Ka Lee for acting as inter-rater. Thanks are expressed to all the volunteered subjects and staff of the Division of Speech and Hearing Sciences for their kind assistance, and to my friends and family for their encouragement and support.