대한음성언어의학회지
제 8 권 제 1 호 1997

Quantitative Analysis of Glottal Cycles According to Frequency and Intensity Variations in Normal Speakers

Young-Ik Son, M.D., Kyungah Lee, M.S., CCC-SLP
Jun-Sun Ryu, M.D., Chung-Hwan Baek, M.D.

Department of Otorhinolaryngology-Head and Neck Surgery, College of Medicine, Sungkyunkwan University, Samsung Medical Center, Seoul, Korea

=국문초록=

발성의 강도와 주파수 변화에 따른 성대 움직임의 정량적 분석

성균관대학교 의과대학 이비인후과학교실, 삼성서울병원
손영익 · 이경아 · 류준선 · 백정환

비디오스트로보스코피를 이용한 성대주기(glottal cycle)의 관찰은 질환의 감별 및 치료 전후의 결과를 비교하는데 중요한 정보를 제공하지만, 관찰자의 주관적 견해나 판단능력에 따라 결과의 판독에 많은 영향을 미칠 수 있다. 이에 비디오스트로보스코피 소견의 객관화를 위한 시도가 최근 보고되고 있으나 아직은 그 성과나 결과 등이 참고치로 활용하기에는 부족한 실정이다.

이에 저자들은 정상 성인을 대상으로 발성시의 주파수와 강도의 변화에 따른 glottal cycle의 변화를 정량화함으로써 추후 연구나 임상적용등의 기본자료로서 활용하고자 하였다.

정상성인 남녀 각 5명을 대상으로 평상시 대화수준, 고음, 큰 소리의 세가지 조건에서, 지속적인 /이/ 발성을 후두비디오스트로보스코프를 이용하여 녹화한 뒤, image analysis 프로그램(KSIP, Kay Elemetrics Corp., NJ, USA)을 이용하여 glottal area waveform(GAW)과 amplitude/symmetry waveform(A/SW)을 구하였다. 네개의 연속적인 glottal cycle의 GAW를 이용하여 peak glottal area, opening & closing phase, open quotient, speed quotient, baseline offset 등의 평균값을 각각의 조건에서 비교하여 보았으며, A/SW를 이용하여 성대의 종축을 기준으로 3등분한 뒤 전, 후, 중간 부위에서 성대운동의 진폭과 그 움직임의 양상을 서로 비교하여 보았다.

발성의 조건에 따라 성대주기와 성대운동의 진폭과 그 양상이 달라짐을 객관적으로 판단할 수 있었으나, 이를 참고치나 기준치로 활용하기 위하여는 더 많은 수의 피검자를 대상으로, 발성의 조건을 체계적으로 변화시킨 연구가 뒷받침 되어야 하리라고 사료되었다.

중심 단어 : 비디오스트로보스코피 · 정량화 · 주파수 · 강도.

Introduction

Physicians' and scientists' effort to visualize the vocal fold movement, which is essential for understanding human phonatory mechanism, has been existent for a long time and introduced a variety of instruments such as electroglottography and high speed photography beginning with a laryngeal mirror[2]. Among them, stroboscopy enormously broadened and deepened our knowledge about vocal fold vibratory patterns. Since late 1980s, stoboscopy has been popularized in many Korean clinics and laboratories, and now is enjoying much use and discussions about its findings.

There are certain parameters clinicians typically look for when they examine laryngeal mechanism of their patients through stroboscopy. Those include fundamental frequency, closure configuration, horizontal and vertical phase difference, amplitude of the horizontal excursion of each vocal fold, mucosal wave, periodicity, existence of nonvibrating portions[1]. Although visual inspection adds much information to clinical decision-making process, examination and interpretation in this way may be very subjective and dependent on the clinician's visual acuity, and clinician's bias may play a significant role.

Recent studies are in a stage pursuing more objective and quantitative data. Kay Elemetrics corporation(Lincoln Park, NJ, USA) has developed a software called KSIP(Kay Stroboscopy Image Processing) which complements the clinician's subjective visual analysis of stroboscopic examination by calculating graphic representations of vocal fold dynamics derived from the stroboscopy recordings. A portion of this software was based on the research paper by Peak Woo[5].

This study aimed to suggest a possible way to quantify normal vocal fold vibration and to determine the effects of intensity and frequency changes on glottal vibratory patterns, and to define phase difference along the horizontal length of vocal folds in normal speakers.

Materials and Methods

Five females and five males, aged between 23 and 32 years old, participated in this study. All of them were employers of the otorhinolarygologic clinic where this study was done, and judged as having nonpathological larynges and normally-sounding voices by a physician and a speech-language pathologist.

Kay's Rhino-Laryngeal Stroboscope system(RLS, Model 9100) and Kay Stroboscopy Image Processing (KSIP) software were used.

Stroboscopic examination was administered with following instructions. Every patient was given an instruction of phonation "Say /ee/ for more than 8 seconds, in 3 different conditions.", first, in a comfortable loudness and pitch level, next, in high-pitched modal phonation, finally, in louder phonation. For high-pitched phonation, subjects phonated at a pitch level which was more than 100Hz higher than that of comfortable pitch level. For louder phonation, subjects phonated at 20dB louder level than that of comfortable loudness level. Particular emphasis was given to ensure the visibility of full length of the larynx during the recording, and to maintain a modal register.

Twenty sequential video frames in one vibratory cycle were captured(Fig. 1). The glottal area was measured by counting the pixel number within the automated trace of the glottal area in each video frame, and the glottal area waveform(GAW), which is a representation of glottal width changes during phonation, was obtained by plotting these values(Fig. 2). When necessary, editing work-ups by the investigator's manual touchup procedures, which was primarily removal of the image of mucous strand, were performed for a more accurate GAW calculation(Fig. 3). Six parameters-peak glottal area, number of frames of opening and closing phases, baseline offset, open quotient and speed quotient were obtained from the GAW analysis.

After defining midline of the glottis, the distance of each vocal fold lateral excursion from the midline was automatically traced and plotted(Fig. 4). Amplitude/symmetry waveform(A/SW), which represents the lateral excursion of the margin of the vocal folds was obtained from these plotting at anterior, middle, and posterior one third portion along the vocal fold length.

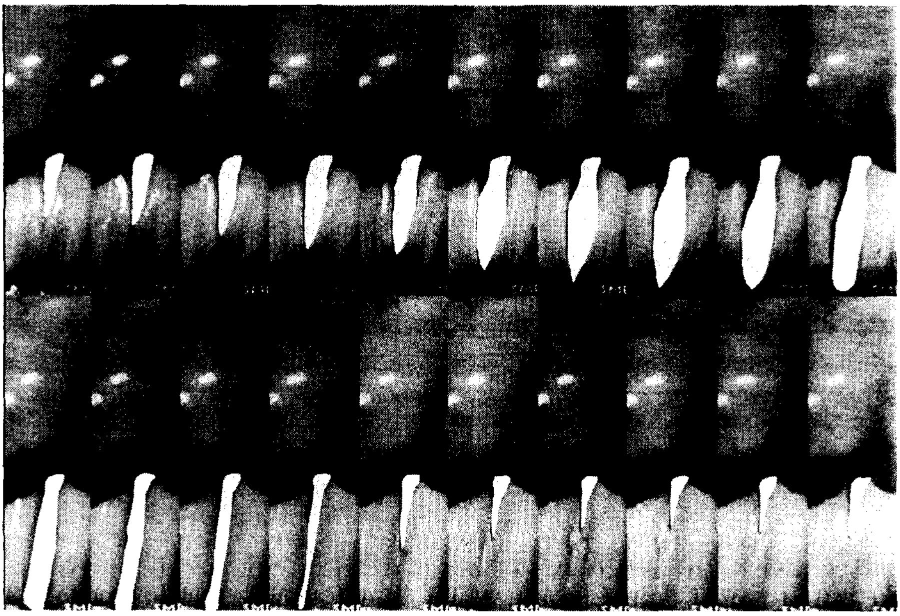

Fig. 1. Twenty sequential video frames in one vibratory cycle are captured, and the glottal area is automatically traced and converted to white-color filled space in each video frame.

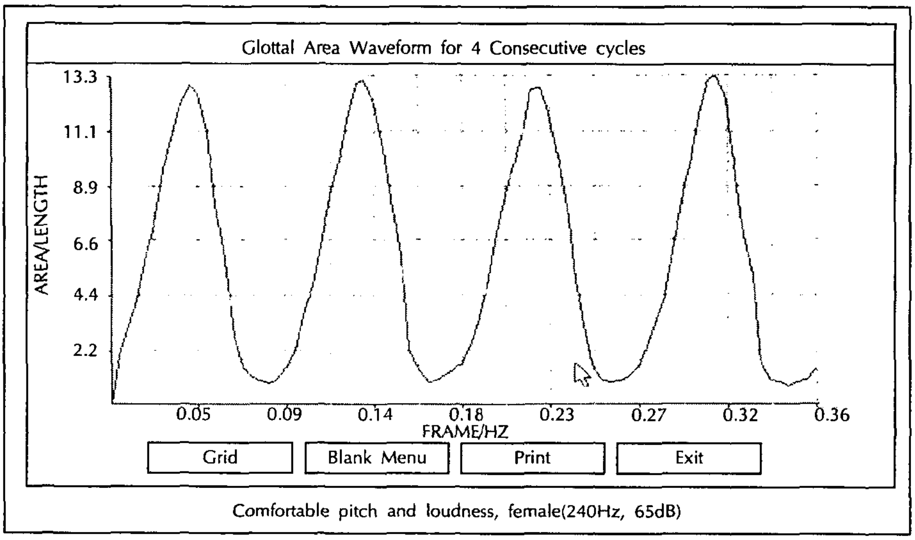

Fig. 2. The glottal area was measured by counting the pixel number within the automated trace of the glottal area in each video frame, and the glottal area waveform(GAW) was obtained by plotting these values. GAW for 4 consecutive glottal cycles in one female speaker is shown, which phonated at comfortable pitch and loudness level.

During comfortable phonation, starting frames of open and closed phase at the anterior, middle and posterior points were obtained and plotted accordingly from A/SW analysis(Fig. 5).

For analysis of GAW and A/SW, 4 consecutive vibratory cycles, rather than 1 or 2 cycles, were analyzed to get a sample of more representative glottal vibrating cycle of each subject. Due to the different endoscopic

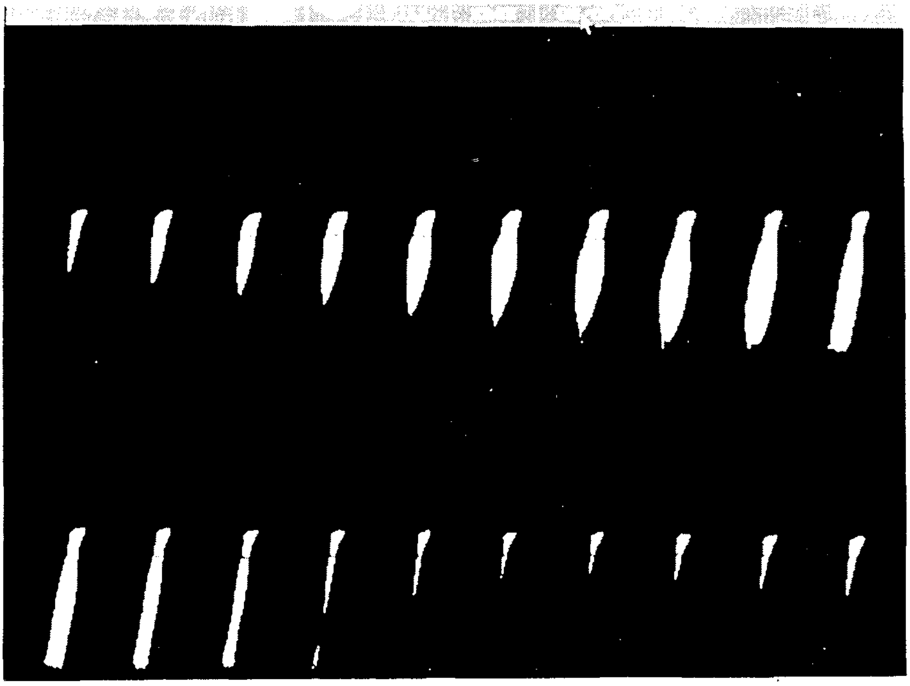

Fig. 3. Automatically traced glottal area can be modified by manual touchup procedure to detect an accurate edge, and to remove the adverse effect of secretional strand.

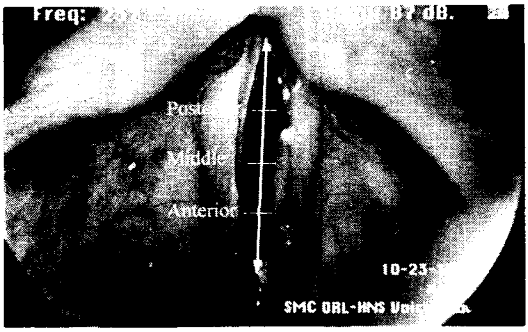

Fig. 4. After defining midline of the glottis, the distance of each vocal fold lateral excursion from the midline was automatically traced at anterior, middle, and posterior one third portion along the vocal fold length.

positions, angles and distances to vocal fold in different speakers, relative glottal area was calculated by dividing GAW by the observed glottal length.

Because the number of subjects was too small, statistical processing was not performed on these values, and values from these analyses were only compared to comfortable level phonation.

Results

1. Glottal area waveform(GAW) analysis(Table 1)

Peak glottal area, which is the largest glottal area reached during one glottal cycle, was significantly reduced in high-pitched phonation in both male and female speakers. The difference between in comfortable phonation and louder phonation was only slight, its value increased in male speakers and slightly decreased in female speakers. The mean values of peak glottal area of male speakers were smaller than those of female speakers except in louder phonation.

The number of frames of opening phase, which is the number of frames in which the vocal folds are moving laterally, decreased both in high-pitched and louder phonations for both male and female speakers. Females had larger mean values than males across all 3 phonation modes.

The number of frames of closing phase, which is the number of frames in which the vocal folds are

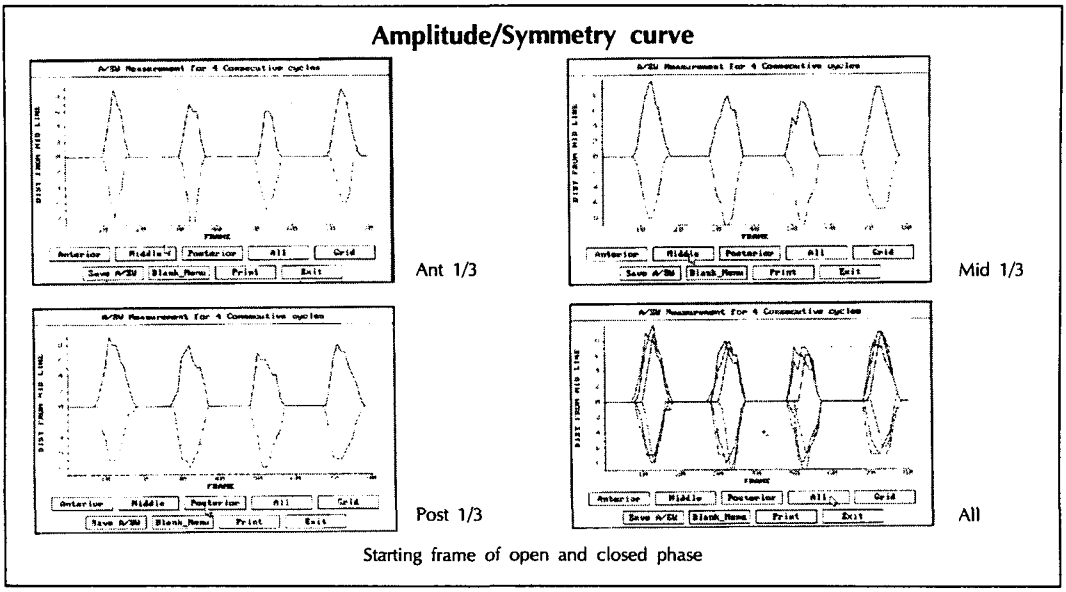

Fig. 5. Amplitude/symmetry waveform(A/SW), which plotted from automatical tracing of vocal fold lateral excursion can be displayed as separate graphs according to anterior, middle, posterior 1/3 point, or altogether in a single graph.

Table 1. Averaged values of parameters obtained from glottal area waveform(GAW) analysis

	Female			Male		
	High	Comfort	Loud	High	Comfort	Loud
Open quotient	0.80	0.83	0.70	0.76	0.69	0.51
Speed quotient	0.78	1.21	1.60	0.61	1.14	1.12
Opening phase	6.84	8.87	8.63	5.75	7.35	5.38
Closing phase	9.17	7.73	5.45	9.50	6.30	4.73
Peak glottal area	11.54	14.60	13.29	9.25	13.58	16.47
Baseline offset	1.43	0.78	0.05	1.05	0.04	0.02

moving medially, exhibited its highest value in high-pitched phonation and lowest value in louder phonation for both male and female speakers. There were no significant sex differences in these values.

The number of frames of closed phase, which is the number of frames in which the midportions of the vocal folds are in contact, had its highest value in louder phonation for speakers of both sexes. Its increase was more dramatic in male speakers than in female speakers.

Open quotient, which is the open period divided by the total period, exhibited some decrease in louder phonation in both sexes.

Speed quotient, which is the glottal opening time divided by glottal closing time, significantly decreased in high-pitched phonation compared to comfortable phonation for both sexes.

Baseline offset, which is the lowest value of glottal area reached, had its highest value in high-pitched phonations for speakers of both sexes.

2. Amplitude/symmetry waveform(A/SW) analysis (Table 2)

Female speakers exhibited a remarkable difference in starting frame of open phase. That is, their posterior 1/3 point was opened most quickly, middle 1/3 point the next, and anterior 1/3 point the slowest. Male speakers did not have this discrepancy of open phase.

For closed phase, male speakers showed much faster starting of closed phase than female speakers. No re-

Table 2. Starting frame of open phase and closed phase obtained from amplitude/symmetry waveform(A/SW) analysis

	Female		Male	
	Open phase	Closed phase	Open phase	Closed phase
Anterior 1/3	7.63	17.23	4.50	13.80
Middle 1/3	6.32	17.32	4.60	13.85
Posterior 1/3	4.69	18.05	4.85	13.90

markable difference was observed along the 3 points in either group of speakers.

Discussion

Since the advent of 1990s, videostroboscopy has been enjoying increasing popularity in medical diagnosis of laryngeal pathology and clinical investigation of vocal fold dynamics. However, examination and interpretation of stroboscopic findings may be subjective and dependent on the clinician's visual acuity, and clinician's bias may play a significant role, and there are no normative data available yet.

In 1996, Peak Woo[5] analyzed videostrobolaryngoscopic images of vocal fold behavior of 65 male and female speakers and provided some quantitative measures that could supplement and corroborate perceptual observations. Some of his findings, but not all of them, were replicated in our study.

In our study, female speakers exhibited larger values for peak glottal area than male speakers except in louder phonation. Woo presented higher values of peak glottal area in male speakers than in female speakers for all modes. This difference can be interpreted as follows. We included cartilaginous portions of the vocal folds when selecting anterior and posterior commissures. During phonation, the cartilaginous portion of the vocal fold of male speakers typically do not vibrate. However, in cases of female speakers whose glottal cycle is characterized by persistent posterior glottal chink throughout the vibrating cycle, the computer automatically calculated glottal area including those chinks.

In louder phonation, open quotient obviously decreased in both sexes, which does well correspond to the findings of previous studies. Open quotient decreased owing to the decrease of both closing and opening phase. In female speakers, decrease of closing phase was more prominent compared to that of comfortable phonation. Male speakers had a slight decrease in speed quotient, which is a ratio between opening phase and closing phase. This is contrary to what Murty, et al.(1990) observed in their study using electroglottography[3]. However, the difference between the current study and the previous study in terms of numeric value is only slight and does not seem to have clinical relevance.

In high-pitched phonation, speed quotient decreased dramatically in both sexes, which was due to the increase of closing phase and the decrease of opening phase. According to the study of Murty et al. (1991)[4] such changes were statistically insignificant. We assure further study with larger samples will clarify the meaning of these changes.

In female speakers, starting frames of open phase from A/SW analysis showed a unique finding, that is, their posterior 1/3 point was opened most quickly, middle 1/3 point the next, and anterior 1/3 point the slowest. Any clear explanation regarding this difference can not be given at this moment.

Although KSIP is considered to be an effective and fairly reliable instrument among so far developed, it still has some limitations. One of those limitations is related to visibility of the total length of vocal folds. In some subjects, it is hard or even impossible to visualize the vocal fold in its entirety due to the subject's anatomical variations. So the marking of anterior and posterior commissure could not be performed accurately. Another limitation is, some pathologic conditions such as vocal fold paralysis or mass lesion may prohibit clinicians from this kind of analysis, because patients with those medical conditions typically are unable to sustain phonation as long as the instrument requires. Touchup procedures, which are necessary to improve the automatic edge detection in many cases where color contrast between the glottis and the vocal folds is poor or dissatisfactory, also permits the clinicians subjectivity to interfere with a correct evaluation.

This study provides an objective basis and the possibility of quantification of the stroboscopic findings such as image shape, amplitude, area, and their changes according to frequency and intensity variations. However, quantitative data or normative data will be actually usable and more reliable, only if the study is performed in larger samples and in systematic conditions such as step-wise variations of intensity and frequency.

Summary

To set up an objective basis for the evaluation of the stroboscopic findings, video-strobolaryngoscopic images of vocal fold vibration in 5 female and 5 male normal speakers were analyzed using an image analysis computer program called KSIP(Kay Storoboscopy Image Processing, Kay Elemetrics Corp., NJ, USA).

Four consecutive vibratory cycles were compared in comfortable, louder, high-pitched /ee/ phonation for every subject. Findings mostly replicated earlier studies including glottal chinks which were observed in most female speakers throughout the cycles and clear distinction between female and male speakers in their vibratory patterns as well as intensity and frequency-related differences. However, there were some findings incompatible with those from previous studies which may be attributable to technical problems.

This study may provide an objective basis of the stroboscopic findings such as image shape, amplitude, area, and their changes according to frequency and intensity variations. We anticipate that further study with larger samples can provide an objective criteria for normal vibratory characteristics of the laryngostroboscopic findings.

References

1) Casiano RR, Zaveri V, Lundy DS : *Efficacy of videostroboscopy in the diagnosis of voice disorders. Otolaryngology-Head and Neck Surgery. 1992 ; 107(1) : 726-727*
2) Hirano M, Bless D : *Videostroboscopic Examination of the larynx, San Diego, Singular Publishing Group. 1993 : 23-34*
3) Murty GE, Carding PN, Kelly PJ : *The effect of intensity on combined glottography. Clin Otolaryngol. 1990 ; 16 : 399-400*
4) Murty GE, Carding PN, Kelly PJ : *The effect of frequency on combined glottography. Clin Otolaryngol. 1991 ; 16 : 298-301*
5) Woo P : *Quantification of Videostroboscopic Findings-Measurement of the Normal Glottal Cycle. The Laryngoscope 106(3) supplement 79, 1996*