The Supraglottal Constriction in Tibetan Chants*

——Electroglottographic Evidences

YOSHINAGA Ikuyo, KONG Jiangping

Abstract

The phonation of Tibetan chants was examined using electroglottographic analysis. The supraglottal constriction, which was considered to generate their peculiar phonation qualities, was examined by a comparison with the vocal fold adduction. The results suggest that the frequency of the ventricular fold oscillation was the same as F0, and the closing peak of the ventricular fold adduction occurred approximately 171 degree after the vocal fold adduction. As the ventricular fold adduction immediately followed the glottal release, it generated the glottal pulse with double peaks in the corresponding glottal airflow. Low OQ_{egg} value was observed because the ventricular fold adduction, which occurred during the glottal release, lowered its value.

Keywords:

supraglottal constriction; ventricular folds; open quotiend (OQ_{egg}); speed quotient (SQ_{egg}); phase difference

1. Introduction

Tibetan lamas in their red robes chant sutras. The pure sounds of their low sonorous pitch heal listeners. *Sabda-vidya*, which deals with ancient Indian linguistic and grammatical studies including singing sutras, was one of the five fields of academics study in ancient India and was deeply treasured and successfully handed down by Tibetan Buddhists.

The supraglottal constriction is estimated to occur in Tibetan Buddhist chants using electroglottographic (EGG) and acoustic analyses [1]. These supraglottic phonations have been reported to occur in certain singing modes, for instance, Mongolian *Kargyraa* in 'throat singing' [2], ethnic and pop style singing [3], and Japanese traditional Noh singing [4]. The supraglottal constriction is widely considered to be caused by movements of the ventricular and aryepiglottic folds (Fig.1), the ventricular folds oscillate at the speed of F0, F0/2 or F0/3 in vocal-ventricular mode (VVM), so do the aryepiglottic folds at the frequency of F0/2 in growl voice [1, 2, 3, 4, 5]. The earlier EGG assessments of phonation types revealed that period-doubling EGG waveforms are the characteristics of VVM [6]. A certain mode of Tibetan chants is assessed as harsh voice [1], in which the ventricular folds generally become involved in the phonation of the true vocal folds [7]. These irregular supraglottic phonations have been found not only in singing but also in vocal fry, voice instabilities, and infant vocalizations. These irregular vocalizations are often interpreted as period-doubling bifurcations, and the corresponding acoustical signals often show sudden jumps to subharmonic regimes [8, 9].

In this paper, we deal with a certain singing mode of Tibetan Buddhist chants. The phase difference between the adduction of the vocal folds and that of the ventricular folds and EGG-based parameters of those adductions are examined using electroglottographic analysis to investigate the supraglottal movements and their contributions for the production of peculiar sounds.

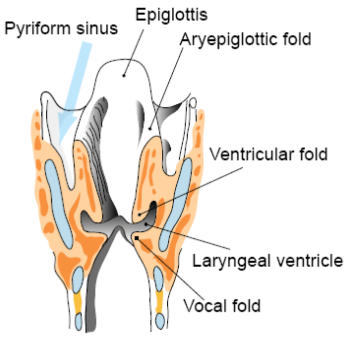

Figure1. Coronal view of the larynx, as seen from behind [3]

2. Material and Method

This section describes the voice materials, calculation method of Electroglottography (EGG)-based parameters, and data processing procedure.

2.1 Voice Material

The phonation of Tibetan Buddhist chants was studied in one Tibetan male monk from *Kumbum* Monastery of *Dge-Lugs-Pa*. The monk was 31 years old, with 18 years of priest experience, when the recording was performed. He was also a teacher at the Monastery with an excellent reputation for his chanting.

The sustained vowel /a/ phonated at 93.7 Hz (≒F2#), whose EGG signal was formed as clear period-doubling waveforms, was studied in this research.

The data acquisition took place at Kumbum Monastery in Qinghai province, China. The EGG signal was obtained by an EGG system (Electroglottograph Model 6103; Kay, USA). The audio signal was recorded by a Sony Electret Condenser Microphone. Those signals were simultaneously recorded and digitized at 16-bit resolution at a sampling frequency of 44.1 kHz.

2.2 Parameter Calculation Method

The EGG signals provide meaningful information only when the vocal folds repeat contact and de-contact during vibration. Therefore, contact-based analysis is the common algorism. A few parameters can be extracted from the EGG waveform that roughly correspond to the open quotient (OQ) and speed quotient (SQ). Because the EGG and airflow waveforms differ from each other qualitatively, OQ_{egg} and SQ_{egg} are employed in this study as the EGG-based parameters. Fig.2 shows that a period of EGG signal can be divided into contact and de-contact phases. Furthermore, the contact phase can be divided into contacting and de-contacting.

Basically, three kinds of EGG calculation methods are proposed, i.e., criterion-level [10], DEGG [11, 12, 13, 14, 15, 14] and the combination of the criterion-level and DEGG methods, called the hybrid method [17, 18]. The DEGG method is employed in this research, in which the glottal closing instance (GCI) and glottal opening instance (GOI) are determined as the maximum and minimum values of the DEGG waveform (Fig.2).

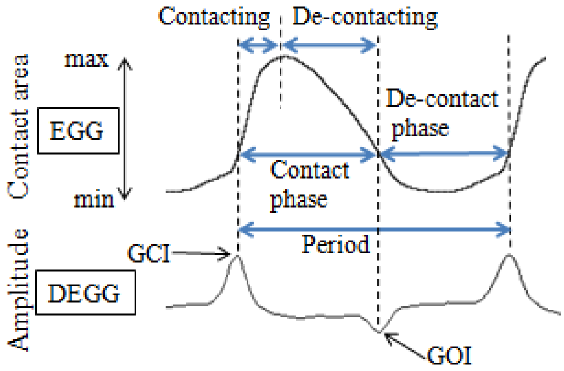

Figure 2. EGG waveform and phases of vocal fold contact.

Three EGG-based parameters are extracted: F0, OQ_{egg} and SQ_{egg}. The definitions of F0 and OQ_{egg} are described as: F0=1/period and OQ_{egg}%=de-contact phase/period*100. Although the SQ_{egg} can be varied in detail across researchers, the definition used in this study is SQ_{egg}%=de-contacting/contacting*100 [19].

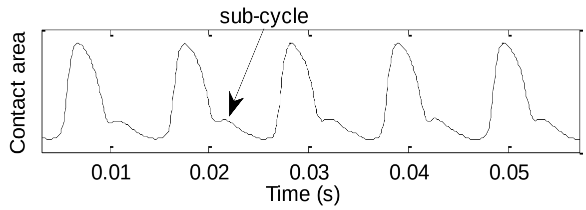

Figure 3. Five vibratory cycles of the EGG signal from vowel /a/ phonated at F2# (92.5 Hz).

The EGG waveform in the data from Tibetan chants demonstrates period-doubling phenomena (Fig.3). These phenomena are quite similar to what was observed in VVM [5, 6]. Furthermore, the ventricular or the aryepiglottic folds are considered to be involved in the supraglottal constriction, however, the aryepiglottic folds oscillate at the speed of $F_0/2$ because the aryepiglottic folds locate higher than the ventricular folds with less developed muscles. Indeed, it is quite reasonable to admit that sub-cycles observed in this study are caused by the oscillation of the ventricular folds.

The time-based parameters of the DEGG waveforms also yield information about periodicity and time patterns of the vibratory events. They are generally referred to as period time (T0), GCI, and GOI. And we call the maximum instance of the vocal fold contact area ‘MI’ in this study (see Fig.4).

Fig.5 shows the DEGG waveform which corresponds to the period-doubling EGG waveform. The parameters of the vocal fold adduction and the ventricular fold adduction are extracted. To avoid confusion between them, '$_2$' is suffixed to the parameters of the latter. Two DEGG-based parameters are extracted: T0=GCI(n+1)-GCI(n) and $T0_2=GCI_2(n)$—GCI(n), 'n' is the number of the glottal cycles. T0 indicates the duration of each glottal cycle, $T0_2$ indicates the duration between the closing peak of the vocal fold and that of the ventricular fold closure. The phase difference between them is given as: $(T0_2/T0)*360°$.

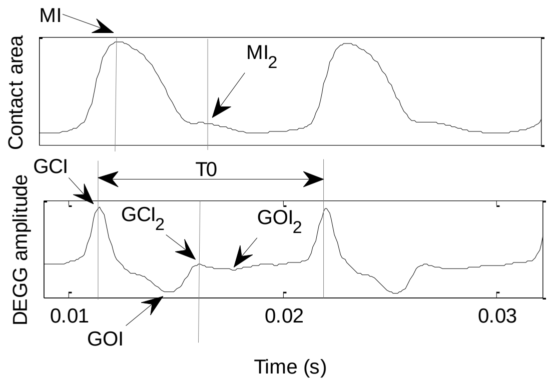

Figure 4. Time-based parameters of the EGG and DEGG waveforms.

According to the definitions of EGG-based parameters described above, three parameters of the vocal fold oscillation are given as: F0=1/T0, $OQ_{egg}\%=(T0-(GOI-GCI))/T0*100$, and $SQ_{egg}\%=(GOI-MI)/(MI-GCI)*100$. And OQ_{egg2} and SQ_{egg2} of the ventricular fold oscillation are given as: $OQ_{egg2}\%=(T0-(GOI_2-GCI_2))/T0*100$ and $SQ_{egg2}\%=(GOI_2-MI_2)/(MI_2-GCI_2)*100$.

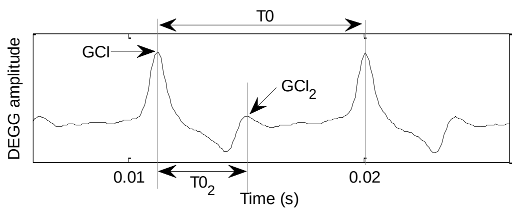

Figure 5. Definitions of T0 and $T0_2$

2.3 Data Processing

The recorded file was downsampled to 11,025 Hz. And the EGG rumble, which was caused by up and down laryngeal movements, was filtered out by a high-pass filter with the cutoff frequency set at 60 Hz because it could affect or mislead the parameter extraction. The parameter values for all of the cycles were extracted using the DEGG method and were saved in an Excel file, the parameter values of 247 data points were extracted from the file concretely. The data processing was performed by a Matlab-based program.

3. Phase Analysis

This section describes the phase difference between the oscillation of the vocal folds and that of the ventricular folds using DEGG analysis.

3. 1 Duration of T0 and $T0_2$

Fig.6 shows the durations of T0 and $T0_2$, they are presented by black and gray dots respectively. The abscissa indicates time, and the ordinate indicates the duration of T0 and $T0_2$ of each cycle.

It seems that T0 is more constant than $T0_2$, the standard deviation (SD) of $T0_2$ is 0.16ms which is 0.09 ms higher than that of T0 (Table II). The maximum value of $T0_2$ is 5.44 ms, and the minimum value is 4.72 ms. The range between them ups to 0.73 ms, on the other hand, the range of T0 is only 0.36 ms which is about a half of $T0_2$. These results suggest that the duration from the vocal fold closure to the ventricular fold closure is not as stable as the duration of each glottal cycle. The mean values of T0 and $T0_2$ is 10.67 ms and 5.07 ms. In other words, the ventricular fold closure occurs approximately 5.07 ms after the vocal fold closure, and the next vocal fold closure follows 5.6 ms after the ventricular fold closure.

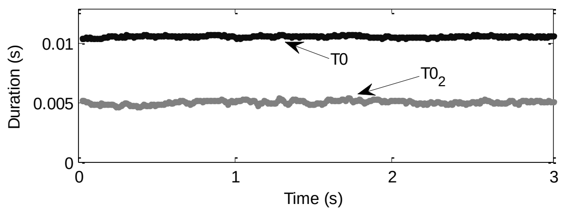

Figure 6. Durations of T0 and $T0_2$.

TABLE III MEAN AND SD VALUES OF T0 AND $T0_2$

	Mean (ms)	SD (ms)
T0	10.67	0.07
$T0_2$	5.07	0.16

3.2 Phase difference between the vocal fold and the supraglottal oscillations

Glottal cycles can be described with 0-360 degree scale. In this study, GCI is described as 0° as well as 360° because it is the end of the glottal cycle as well as the beginning of the next glottal cycle.

Fig.7 shows the phase difference between the adduction of the vocal folds and that of the ventricular folds, it is presented by black dots. The abscissa indicates time, and the ordinate indicates the phase difference. Fig.7 roughly indicates that the supraglottal closure occurs approximately 180° after the vocal fold closure, and it shows that the values are not very stable. Table IV shows the mean and SD values of the phase difference between the adduction of the vocal folds and that of the ventricular folds. The mean value of the phase differences is 171.17° with the SD of 5.26°. The earlier researches on VVM using the high-speed videoendoscopy revealed that the ventricular fold closure occurred during the 480°-560° interval during the vocal folds were open [5]. In other words, the ventricular fold closure occurred during the 120°-200° interval every other glottal cycle because the frequency of the ventricular fold oscillation was F0/2 in the above case. The mean value of the phase differences between the adduction of the vocal folds and that of the ventricular folds in this research is 171.17°, the frequency of the ventricular fold oscillation is equal to that of the vocal fold oscillation. The phase difference of 171.17° is also in the 120°-200° interval, and the ventricular fold adduction also occurs during the glottal release in this case.

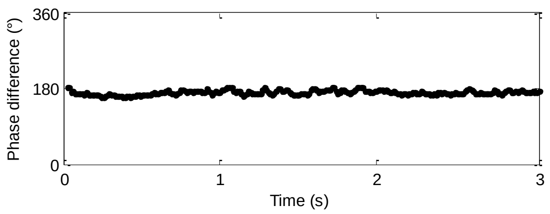

Figure 7. Phase difference between the vocal fold and ventricular fold adductions

TABLE V MEAN AND SD VALUES OF THE PHASE DIFFERENCE BETWEEN THE VOCAL FOLD AND VENTRICULAR FOLD ADDUCTIONS.

	Mean (°)	**SD (°)**
The phase difference between the vocal fold and ventricular fold adductions	171.17	5.26

[20] has reported on the relationship between the glottal air flow and vocal fold contact area. According to the result, opening of upper fold margins corresponds to the first airflow of the pulse, and the pulse terminates closely before lower fold margins close. Indeed, an upside-down image of the EGG waveforms roughly demonstrates the glottal airflow. Fig.8 shows the estimated glottal airflow based on the results from the phase analysis. An upside-down image of the EGG waveforms is shown as the estimated glottal airflow. 0° and 171° indicate the decreasing peaks of the airflow, the former is caused by the vocal fold closure, and the latter is caused by the ventricular fold closure. A pulse can be described as follow: the airflow starts to be released, the volume velocity of the glottal pulse decreases closely before 171°, and the peak volume velocity of the pulse follows it, then the pulse terminates. As a result, the glottal airflow is formed with two peaks.

As mentioned above, the SD values of the phase difference between the adduction of the vocal folds and that of the ventricular folds are not so stable that the volume velocity of the glottal airflow at the first peak, which occurs closely before 171°, is affected by the unstable phase difference. In short, the glottal airflow is transformed by the ventricular fold constriction which seems to play an important role in generating their peculiar voice qualities.

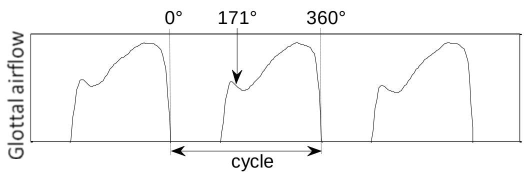

Figure 8. Estimated glottal airflow.

4. Parameter Analysis

This section describes the characteristics of EGG-based parameters of the vocal fold and ventricular fold oscillations to observe their adduction mechanisms.

4.1 OQ_{egg}

As mentioned above, OQ_{egg} indicates OQ_{egg} of the vocal fold oscillation, OQ_{egg2} indicates OQ_{egg} of the ventricular fold oscillation in this study. Fig.9 shows the OQ_{egg} and OQ_{egg2} values, black dots indicate OQ_{egg}, and gray dots indicate OQ_{egg2}. The abscissa indicates time, and the ordinate indicates the OQ_{egg} and OQ_{egg2} values of each cycle.

Table VI shows the mean and SD values of OQ_{egg} and

OQ_{egg2}. The mean OQ_{egg} is 67.64%, the SD value is 1.73%. The mean OQ_{egg2} is 74.59%, the SD value is 0.28%. The mean OQ_{egg2} is 6.95% higher than OQ_{egg}, it suggests that the contact phase of the ventricular fold oscillation is shorter than that of the vocal fold oscillation. Furthermore, the OQ_{egg2} value is more stable than OQ_{egg}, for the SD value of the former and the latter are 0.28% and 1.73%. The stable T0 and unstable $T0_2$ were the result from the earlier section, unstable OQ_{egg} seems to be caused by the unstable $T0_2$. It seems plausible that the ventricular fold constriction, which shortly follows the glottal release, influences the vocal fold movements. Correlation analysis between OQ_{egg} and $T0_2$ reveals that there is a negative correlation of -0.492 ($p<0.01$) between them.

The real OQ_{egg} value can be extracted by subtracting the contact quotient (CQ_{egg}) value of the ventricular fold oscillation from OQ_{egg} because the ventricular fold constriction occurs at the de-contact phase of the vocal fold oscillation. The mean real OQ_{egg} is 42.23% which is characterized as significantly low OQ_{egg}. It is quite natural that the phonation containing the supraglottal constriction has low OQ_{egg}.

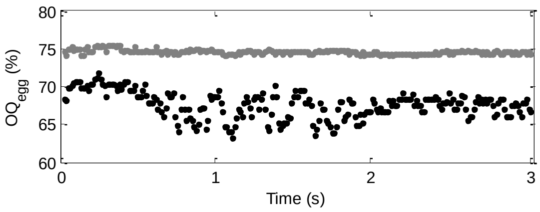

Figure 9. OQ_{egg} values of the glottal and the supraglottal adductions.

TABLE VII MEAN AND SD VALUES OF OQ_{egg}

	Mean OQ_{egg} (%)	SD (%)
The glottal oscillation	67.64	1.73
The supraglottal oscillation	74.59	0.28

4.2 SQ_{egg}

As mentioned above, SQ_{egg} indicates SQ_{egg} of the vocal fold oscillation, SQ_{egg2} indicates SQ_{egg} of the ventricular fold oscillation in this study. Fig.10 shows the SQ_{egg} and SQ_{egg2} values, black dots indicate SQ_{egg}, and gray dots indicate SQ_{egg2}. The abscissa indicates time, and the ordinate indicates the SQ_{egg} and SQ_{egg2} values of each cycle.

Table VIII shows the mean and the SD values of SQ_{egg} and SQ_{egg2}. The mean SQ_{egg} is 150.52% with the SD value of 11.00%, and the mean SQ_{egg2} is 100.73% with the SD value of 2.17%. The mean SQ_{egg} is 49.79% higher than SQ_{egg2}. These results suggest that the strong force of the glottal adduction raises the energy at the high frequency region in the vocal fold oscillation. On the other hand, low SQ_{egg2} resulted from less forced adduction because the ventricular folds are incapable of becoming tense, since they contain very few muscle fibres [21]. Fig.10 shows that the SQ_{egg} values are unstable, the SD value reaches 11% which is 8.83% higher than that for the ventricular fold oscillation. Correlation analysis reveals that there is a negative correlation of -0.629 ($p<0.01$) between the OQ_{egg} and the SQ_{egg}.

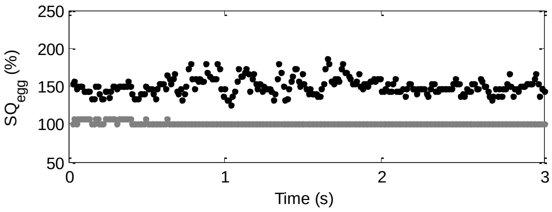

Figure 10. SQ_{egg} values of the glottal and the supraglottal adductions.

TABLE IX MEAN AND SD VALUES OF SQ_{egg}

	Mean SQ_{egg} (%)	SD (%)
The glottal oscillation	150.52	11.00
The supraglottal oscillation	100.73	2.17

EGG-based parameter values of the ventricular fold oscillation are rather stable compared to those of the vocal fold oscillation. It is quite natural that the values of the latter are complicated because the muscles of the vocal folds and those to control the vocal folds are much more developed than those of the ventricular folds. The mean value of real OQ_{egg} is 42.23%, its significantly low OQ_{egg} is one of the characteristics of pressed phonation. The OQ_{egg} of the vocal fold oscillation and the phase difference between the adduction of the vocal folds and that of the ventricular folds are negatively correlated, and the OQ_{egg} and SQ_{egg} of the vocal fold oscillation are negatively correlated. In short, the ventricular fold oscillation occurs unstably in phases, but the values of OQ_{egg2} and SQ_{egg2} are very stable. On the contrary, the duration of each glottal cycle is quite stable, but the values of OQ_{egg} and SQ_{egg} are not very stable because of the unstable occurrence of the ventricular fold oscillation in phases.

5. Conclusion

The ventricular fold constriction was estimated to occur approximately at 171° phase difference after the vocal fold closure. As the ventricular fold constriction immediately followed the glottal release, the estimated glottal airflow was formed with double peaks of the glottal pulse. Though F0 was very stable, the OQ_{egg} and the SQ_{egg} of the vocal fold oscillation were somehow unstable because of the unstable phase difference between the adduction of the vocal folds and that of the ventricular folds. The ventricular fold constriction contributed to lower the OQ_{egg} value. These phonation techniques involving the supraglottal structures made a great role in generating their peculiar phonation qualities.

Further physiological research using tools such as high-speed cameras is needed to clarify the ventricular adduction. Synthesize and perceptual evaluations are also expected as future work.

6. Acknowledgment

Thanks are due to Gazangduojie from Kumbum Monastery for the recording and valuable advices on this research.

This study was supported by China Social Sciences Funds of the Ministry of Education (Grant No: 10JJD740007).

7. References

[1] I. Yoshinaga and J. Kong, "Some phonatory characteristics of Tibetan Buddhist chants," Journal of the Phonetic Society of Japan, vol. 15(2), pp. 83-90, 2011.

[2] P. Lindestad, M. Södersten, B. Merker, and S. Granqvist, "Voice source characteristics in Mongolian 'throat singing' studied with high-speed imaging technique, acoustic spectra, and inverse filtering," Journal of Voice, vol. 15(1), pp. 78–85, 2001.

[3] K-I. Sakakibara, L. Fuks, H. Imagawa, and N. Tayama, "Growl voice in ethnic and pop styles," In Proceedings of the International Symposium on Musical Acoustics. Nara, 2004.

[4] I. Yoshinaga and J. Kong, "Laryngeal vibratory behavior in traditional Noh singing," Tsinghua Science and Technology, vol. 17(1), pp. 94-103, 2012.

[5] L. Fuks, B. Hammarberg, and J. Sundberg, "A selfsustained vocalventricular phonation mode: acoustical, aerodynamic and glottographic evidences," KTH TMHQPSR, vol. 3, pp. 49–59, 1998.

[6] J. H. Esling, "Laryngographic study of phonation type and laryngeal configuration," Journal of the International Phonetic Association, vol. 14(02), pp. 56–73, 1984.

[7] J. Laver, The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press, 1980.

[8] H. Hollien, J. Michel, and E. D. Thomas, "A method for analyzing, vocal jitter in sustained phonation," Journal of Phonetics, vol. 1, pp. 85–91, 1973.

[9] I. R. Titze, R. J. Baken, and H. Herzel, "Evidence of chaos in vocal folds vibration in Vocal fold physiology," in New Frontiers in Basic Science, I. R. Titze, Eds. San Diego: Singular Publishing Group, 1993, pp. 143–188.

[10] M. Rothenberg and J. J. Mahshie, "Monitoring vocal fold abduction through vocal fold contact area," Journal of Speech and Hearing Research, vol. 31, pp. 338–351, 1988.

[11] N. Henrich, "On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation," Journal of the Acoustic Society of America, vol. 115, pp. 1321–1332, 2004.

[12] D. G. Childers, D. M. Hicks, G. P. Moore, and L. Eskenazi, "Electroglottography and vocal fold physiology," Journal of Speech and Hearing Research, vol. 33, pp. 245–254, 1990.

[13] D. G. Childers and A. K. Krishnamurthy, "A critical review of electroglottography," Critical Reviews in Biomedical Engineering, vol. 12, pp. 131–161, 1985.

[14] D. G. Childers and J. N. Larar, "Electroglottography for laryngeal function assessment and speech analysis," IEEE Transactions on Biomedical Engineering, vol. 31, pp. 807– 817, 1984.

[15] D. G. Childers, G. P. Moore, J. M. Naik, J. N. Larar, and A. K. Krishnamurthy, "Assessment of laryngeal function by simultaneous, synchronized measurement of speech, electroglottography and ultra-high speed film," In Proceedings of the Eleventh Symposium on Care of the Professional Voice. New York, 1983.

[16] D. G. Childers, J. M. Naik, J. N. Larar, A. K. Krishnamurthy, and G. P. Moore, "Electroglottography, speech and ultra-high speed cinematography," In Vocal Fold Physiology and Biophysics of Voice, I. R. Titze and R.

Scherer, Eds. Denver: Denver Center for the Performing Arts, 1983, pp. 202–220.

[17] D. M. Howard, "Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers," Journal of Voice, vol. 9(2), pp. 163–172, 1995.

[18] D. M. Howard, G. A. Lindsey, and B. Allen, "Toward the quantifi cation of vocal effi ciency," Journal of Voice, vol. 4, pp. 205–212, 1990.

[19] J. Kong, On Language Phonation [Lunyuyanfasheng]. Beijing: Central Nationalities University Press, 2001.

[20] M. Rothenberg, "Some relations between glottal air flow and vocal fold contact area," In Proceedings of the Conference on the Assessment of Vocal Pathology. National Institutes of Health, 1979.

[21] K-I. Sakakibara, H. Imagawa, S. Niimi, and N. Tayama, "Physiological study of the supraglottal structure," In Proceedings of the International Conference on Voice Physiology and Biomechanics. Marseille, 2004.