Some Phonatory Characteristics of Tibetan Buddhist Chants*

YOSHINAGA Ikuyo, KONG Jiangping

要旨

歌声を形成する言語情報（歌詞）と非言語情報（音高・音長情報を含む旋律及び声質）のうち非言語情報である声質に焦点を当て、電気声門図及び音声信号を基にチベット声明の発声における特徴の解明を試みた。その結果、声明の音源パラメータは低声門開放率及び高声門開閉速度率の特徴をもち、それらに対応する音響特徴として H1-H2 及び H1-A3 が共に低い数値を示した。これらは声明におけるりきみ発声の特徴を表している。次に、電気声門図波形及びスペクトル解析により、単なるりきみだけでなく、声門上構造物の振動によると推測されるザラザラ感（harsh）のある発声も確認された。ここではその周波数は声帯振動と同じ F0 であった。声明における音域は通常発話時の半分の音域に値する 2 半音であり、音高も発話時より 2 半音低めであった。このような特徴からチベット声明は喉詰型発声の伝統を汲むことがわかった。

Keywords

pressed voice, harsh voice, supraglottal constriction, electroglottography (EGG), open quotient (OQ_{egg}), speed quotient (SQ_{egg})

1. INTRODUCTION

Tibetan lamas chant sutras with devout devotion and passion. ‘Outsiders’ listening to these chants often cannot help but be impressed by the unique sounds of their low sonorous pitch. *Shomyo* (sabda-vidya, in Sanskrit), as it has been called, was one of the five fields of academic study in ancient India and was deeply treasured and successfully handed down by Tibetan Buddhists.

In ‘throat singing’, Mongolian *Kargyraa* is the common label for low, bass-pitched singing, and a similar style is found in Tibetan Buddhist chants (Lindestad et al. 2001, Sakakibara 2003). A study conducted in the 1960s used sonograms to hypothesize that the ‘odd harmonics’ found in the chants of Tibetan lamas were produced by double oscillators or asymmetrically vibrating vocal folds (Smith et al. 1967). High-speed video endoscopy and electroglottography (EGG) of non-Tibetan ‘throat singing’ revealed that the ventricular folds oscillated at half of the frequency of vocal folds in a typical phonation mode, which was judged to be perceptually identical to that used in Tibetan Buddhist chants (Fuks et al. 1998). In Mongolian ‘throat singing’, the ventricular fold vibrations were observed via high-speed imaging techniques and kymography (Lindestad et al. 2001). The ventricular folds oscillate at a frequency of F0, F0/2, or F0/3 in vocal-ventricular mode (VVM) (Fuks et al. 1998, Lindestad, et al. 2001, Sakakibara, et al. 2004). These supraglottic phonations have been found not only in singing techniques 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 (Hollien et al. 1973, Titze et al. 1993).

The phonatory characteristics of voice qualities are very important in defining singing techniques. However, perceptual assessments of voice qualities remain ambiguous. Objective assessments, such as acoustical analysis, synthesis, and physiological observation, are needed (Sakakibara 2003).

This paper describes the electroglottographic and acoustic analyses conducted to reveal the voice production mechanisms of certain singing modes of Tibetan Buddhist chants and describes the phonatory characteristics of these voice qualities.

2. METHODS AND MATERIALS

This section describes EGG, the primary experimental method of this research, the calculation method of EGG-based parameters, voice materials, and the data processing procedure.

2.1 Electroglottography

Voice quality is a key issue in describing various singing styles. Perceptual assessment and a variety of instrumental (acoustical and physiological) methods are applied in the definition of voice qualities (Raymond and Martin 2000). EGG, which measures electrical conductance changes between a pair of electrodes placed on the neck, is a noninvasive technique for the observation of vocal fold vibratory patterns. One of the

*Published: Journal of the Phonetic Society of Japan, vol. 15(2), pp. 83-90, 2011.

authors has described five Chinese phonation types using EGG parameters. Table 10 shows that the open quotient (OQ_{egg}) and the speed quotient (SQ_{egg}) are key factors in distinguishing these five phonation types: vocal fry, breathy voice, pressed voice, modal voice, and high-pitched voice. For example, vocal fry is characterized as high OQ_{egg} and SQ_{egg}. These parameters are very important to voice quality assessments.

Table 10 Distinctive features of the source parameter in five phonation types compared with the modal voice. "–" indicates lower and "+" indicates higher than the modal voice (Kong 2001).

	Fry	Breathy	Pressed	Modal	High
F0	–	–	–	±	+
OQ_{egg}	+	+	–	±	–
SQ_{egg}	+	–	+	±	–

Fig. 1 Simplified illustration of the vocal folds, EGG waveform and parameter, and spectral tilt related to the phonation type.

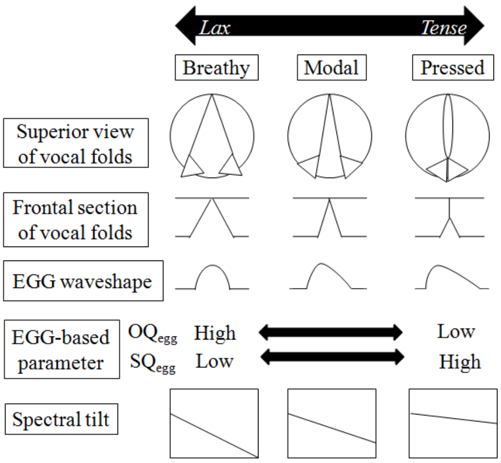


Fig. 1 is a simplified illustration that shows the relationships of voice qualities to physiological, electroglottographic, and spectral properties. The relationship between the EGG waveform and its corresponding frontal section of vocal folds is described according to Titze (1990). The relationship between the superior view of vocal folds and its spectral tilt is determined with reference to Stevens (1977). Low OQegg represents pressed voice with the larger contact area of the vocal folds that can be seen at the frontal section of the vocal folds in Fig. 1. In contrast, high OQegg represents breathy voice because more airflow is released with the longer de-contacting duration. A voice with lower SQegg has weaker energy because of the reduced speed when the vocal folds come into contact. Acoustically, this is reflected by the steeper spectral tilt. In contrast, higher SQegg has more forceful and quicker glottal closure and is accompanied by a more gradual spectral tilt. Thus, the characteristics of these EGG parameter values are reflected in the acoustic features.

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.

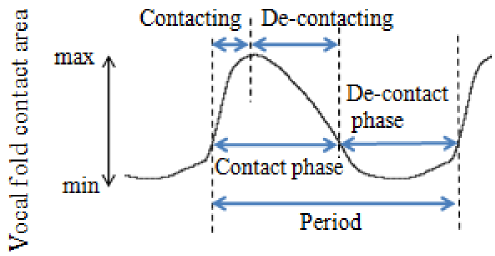


Fig. 2 EGG waveform and phases of vocal fold contact.

Three EGG-based parameters are extracted: F0, OQ_{egg}, and SQ_{egg}. The definitions of F0 and OQ_{egg} are described as follows: 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 research is SQ_{egg}%=de-contacting/contacting*100 (Kong 2001).

There have been discussions on the definition of the glottal closing instance (GCI) and glottal opening instance (GOI) (Baken and Orlikoff 2000, Henrich 2004, Herbst 2004, Howard et al. 1990, Howard 1995). Three kinds of EGG calculation methods are proposed, i.e., criterion-level (Rothenberg 1988), derivative of the EGG signal (DEGG) (Henrich 2004, Childers, Hicks, Moore and Eskenazil 1990, Childers, Moore, Naik, Larar and Krishnamurthy 1983, Childers, Naik, Larar, Krishnamurthy and Moor 1983, Childers and Krishnamurthy 1985, Childers and Larar 1984) and the combination of the criterion-level and DEGG methods, called the hybrid method (Howard et al. 1990, Howard 1995). The DEGG is considered the ideal method to reflect the GCI and GOI, but it is not reliable in the case of imprecise or multiple GCIs and GOIs (Henrich 2004). The EGG waveform in the data from Tibetan chants demonstrates period-doubling phenomena (Fig. 3). In this case, the DEGG signals show double GCIs or GOIs, and the precise setting of the criterion level is necessary so that each instance can be detected. Therefore, the criterion-level method of the 35% threshold is employed in this study, in which the threshold level is determined between the maximum and minimum values of the EGG waveform (Fig. 4).

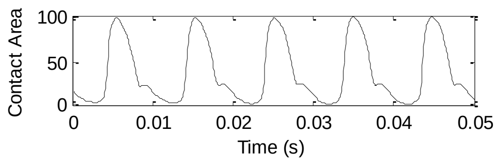


Fig. 3 Five vibratory cycles of the EGG signal from vowel /a/ phonated at G2 (98 Hz).

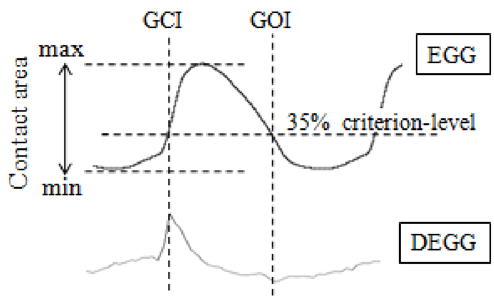


Fig. 4 Definition of GCI and GOI with a 35% criterion-level and derivative EGG (DEGG) waveform.

2.3 Voice Material

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

The voice materials consist of two types: 1) a sutra, *Gadanlajima*, and 2) sustained vowels /a, e, i, o, u/. Gadanlajima is a representative sutra in the Kumbum Monastery. The subject chanted and read the sutra at his comfortable pitch.

The subject sustained five vowels /a, e, i, o, u/ using two styles, chanting and speaking, in material 2. Each semitone was produced from the lowest to highest for the subject's range while attempting to maintain the same volume in chanting and speaking. Thus, the factors that might influence the values of source parameters were eliminated.

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.

2.4 Data Processing

To prepare the recorded files for acoustical analysis, the files were down-sampled to 11.025 kHz. Next, 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 files were divided into smaller pieces in preparation for the batch processing to obtain the value of the EGG parameters. The parameter values for all of the cycles were extracted using the criterion-level method of 35% and were saved in an Excel file. Because a large amount of data processing was needed and the lengths of the recorded files were inconsistent, parameter values at 30 data points were also extracted from each piece of the recorded file and saved in an Excel file. Before extracting the values of the EGG-based parameters, the wavelet transform was applied to each file to reduce the high-frequency noise of EGG signals, which might cause miscalculations in detecting the highest peaks and contacting and de-contacting peaks (Kong and Liew 1998). The data processing was performed by Matlab-based VoiceLab, which was developed by the Linguistic Lab of Peking University.

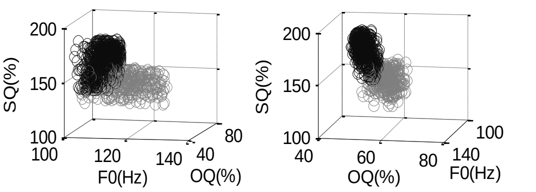
The lengths of the recorded data for chanting and speaking were approximately four minutes for each in material 1. The parameter values of 700 data points were extracted from each data file. Data that indicated abnormally low or high values for parameters were deleted because they may not be from vowels but from voiced consonants. In the case of material 2, parameter values at 30 data points were extracted from each sustained vowel.

3. PARAMETER ANALYSIS

In this section, the EGG parameters are compared with comparisons between chanting and speaking to recognize the inherent features of glottal source in chanting.

3.1 Parameter Distribution of Gadanlajima

Fig. 11 shows the parameter distribution of Gadanlajima for chanting and speaking. The x-, y- and z-axes in Fig. 11a represent F0, OQ_{egg} and SQ_{egg}, and those of Fig. 11b represent OQ_{egg}, F0 and SQ_{egg}, respectively. The parameters for chanting are shown by 700 black circles, and those for speaking are shown by 700 gray circles. Table 11 shows the mean and range of F0, OQ_{egg} and SQ_{egg}. In Fig. 11a, the data for chanting are located at a lower F0 region than that for speaking. The mean F0 of chanting is 102.3 Hz, which is 2 semitones lower than speaking. The F0 range of chanting is a little over 2 semitones (14.9 Hz), which is only half of the range of speaking. Fig. 11b shows the distribution of OQ_{egg}, demonstrating that the values are lower for chanting than for speaking. The mean OQ_{egg} of chanting is 52.8%, which is lower than that of speaking by 4.4%. The range of OQ_{egg} is 7% for chanting and 10.1% for speaking. Because the F0 range of speaking is wider than chanting, it is quite natural that the OQ_{egg} range of speaking is also wider. The SQ_{egg} for chanting is significantly higher than that of speaking. The mean SQ_{egg} value of chanting is 232.1%, which is 88% higher than that of speaking. Its range is 49.4%, which is 14.5% wider than that of speaking.



(a) (b)

Fig. 5 The distribution of parameters for chanting (black) and speaking (gray).

Table 11 The mean and range of F0, OQ_{egg} and SQ_{egg}.

		F0 (Hz)	OQ_{egg} (%)	SQ_{egg} (%)
Chanting	Mean	102.3	52.8	232.1
	Range	14.9	7.0	49.4
Speaking	Mean	116.1	57.2	144.1
	Range	28.0	10.1	34.9

To summarize, the chanting of Gadanlajima is characterized by low F0, low OQ_{egg} and high SQ_{egg} compared to speaking (see Table 12).

Table 12 Parameter characteristics of Gadanlajima.

	F0	OQ_{egg}	SQ_{egg}
Chanting	–	–	+
Speaking	+	+	–

3.2 Parameter Distribution in Sustained Vowels

It is common for OQ_{egg} and SQ_{egg} to co-vary with F0. Because there is a pitch range difference between chanting and speaking for Gadanlajima, the sustained vowels with the same pitch height are examined for chanting and speaking in this section. The results from sustained vowels in his entire pitch range show that the distribution of OQ_{egg} and SQ_{egg} in his high-pitch region do not show a significant difference between chanting and speaking. This is because 2~4 semitones near the lowest pitch region are used for actual chanting and speaking (cf. Table 11). Therefore, a significant difference is observed in the parameter values obtained from the low-pitch region. Thus, the pitch range compared here is limited from F2# (92.5 Hz) to B2 (123.5 Hz). Parameter values of 520 data points are extracted from both chanting and speaking. Fig. 6 shows the distribution of OQ_{egg} and SQ_{egg} of chanting (black) and speaking (gray). The x-axis and y-axis of Fig. 6 represent OQ_{egg} and SQ_{egg}. Table 13 shows the mean value and range of parameters. The OQ_{egg} in chanting is 3.5% lower than that in speaking. The OQ_{egg} range of chanting is 10.1% and that of speaking is 12.2%. The latter is slightly wider. The distribution of SQ_{egg} is separated between chanting and speaking. The mean SQ_{egg} of chanting is 19.9% higher, and the SQ_{egg} range is 15.4% narrower than speaking.

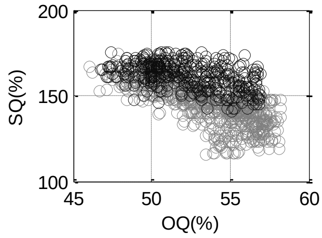


Fig. 6 The distribution of OQ_{egg} and SQ_{egg} in chanting (black) and speaking (gray) phonated at the range of F2#~ B2.

Table 13 The mean and range of OQ_{egg} and SQ_{egg} phonated at the range of F2#~ B2.

	OQ_{egg} (%)		SQ_{egg} (%)	
	Mean	Range	Mean	Range
Chanting	50.3	10.1	164.5	35.1
Speaking	53.8	12.2	144.6	50.5

Thus, the low OQ_{egg} and high SQ_{egg} are the common features in the chanting (see Table 14), which agrees with the result from Gadanlajima. The lower OQ_{egg} in chanting indicates the longer duration of vocal fold contact, suggesting that more pressed phonation is employed. The higher SQ_{egg} means that vocal fold contact is more rapid, which results in the higher energy in the higher frequency region.

Table 14 Parameter characteristics of sustained vowels.

	OQ_{egg}	SQ_{egg}
Chanting	Low	High
Speaking	High	Low

4. SPECTRAL MEASURE ANALYSIS

This section describes the acoustic manifestations that correspond to the results from the time domain analysis. The measurements include the H1-H2 and H1-A3.

4.1 H1-H2 Measurement

Spectral measure analysis is often used as an acoustic method to assess voice qualities. The difference in amplitude between the first and second harmonics (H1-H2) is a common measure to judge the tightness of vocal fold closure. The lower the H1-H2 is, the smaller the open quotient of vocal fold vibration becomes to produce pressed voice. For instance, breathy voice has high H1-H2, and creaky voice has low H1-H2. The lower OQ_{egg} is obtained for chanting from the EGG analysis, which suggests that more pressed phonation is used. The H1-H2 is expected to be low in chanting because it reflects the open quotient of glottal vibration (Bickley 1982).

Fig. 7 is the result of the H1-H2 values of sustained vowel /a/ at F2#~B2. The reason for adopting only vowel /a/ in this section is that the first formant of vowel /a/ has a higher frequency that hardly influences the values of H1 or H2. Fig. 7 shows that the H1-H2 of chanting is lower than speaking, as expected. The H1-H2 value of chanting is 7.6 dB and that of speaking is 11.4 dB, which is 3.8 dB higher than the former (Table 15). This result suggests that chanting has more pressed voice quality than speaking, which agrees with the result of low OQ_{egg} in chanting from EGG analysis.

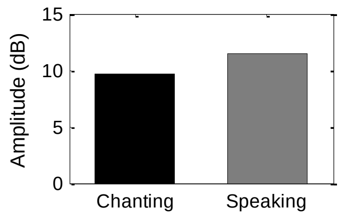


Fig. 7 H1-H2 value of vowel /a/ phonated at the range of F2#~B2.

Table 15 H1-H2 value of vowel /a/ phonated at the range of F2#~B2.

	Chanting	Speaking
Amplitude (dB)	7.6	11.4

4.2 H1-A3 Measurement

The difference in amplitude between the first harmonic and third formant frequency (H1-A3) is one of the common measures to judge the spectral tilt. The lower the H1-A3 is, the smaller the spectral tilt becomes. For instance, breathy voice has large H1-A3, which is indicated as a steep spectral tilt, unlike creaky voice, which has low H1-A3, as indicated by a small spectral tilt. The low H1-A3 is expected in chanting because the high SQ_{egg} is reflected by the low H1-A3 (Stevens and Hanson 1995).

Fig. 8 shows the H1-A3 values of sustained vowel /a/ at F2#~B2. The H1-A3 value of chanting is lower than that of speaking, as expected. The H1-A3 value of chanting is 21.8 dB and that of speaking is 28.4 dB, which is 6.6 dB higher than the former (Table 16). This suggests that chanting has a smaller spectral tilt than speaking. This finding agrees with the result of high SQ_{egg} in chanting from the EGG analysis; namely, chanting has stronger energy in the high-frequency region.

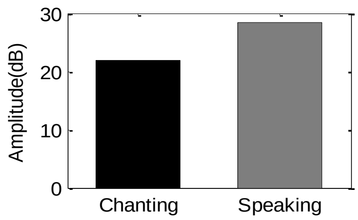


Fig. 8 H1-A3 value of vowel /a/ phonated at the range of F2#~B2.

Table 16 H1-A3 value of vowel /a/ phonated at the range of F2#~B2.

	Chanting	Speaking
Amplitude (dB)	21.8	28.4

Low H1-H2 and low H1-A3 in chanting from the spectral analysis (Table 17) agree with the results of low OQ_{egg} and high SQ_{egg} from the EGG analysis. Thus, chanting is described as a more pressed phonation with stronger energy in the high-frequency region. The results from the EGG parameter analysis in the time domain are reflected in the results from the acoustic parameter analysis in the frequency domain.

Table 17 Spectral characteristics of chanting and speaking.

	H1-H2	H1-A3
Chanting	Low	Low
Speaking	High	High

5. PHONATION ANALYSIS

In this section, the phonation mechanism of chanting is investigated by observing the EGG and DEGG waveforms and spectrograms.

5.1 Phonation Mode

EGG and acoustic parameter analyses were conducted in the earlier sections. The results indicate that pressed phonation with strong energy in the high-frequency region is one of the characteristics of chanting. To go a step further, the shape of EGG and DEGG waveforms are observed. The period-doubling pattern is a typical EGG waveform of chanting, as shown in Fig. 3. The subcycles in the period-doubling waveform seem to be derived from the phase difference between vocal fold and supraglottal oscillations. This feature of the EGG waveform is quite similar to what was observed in VVM (Fuks. et al. 1998).

The subcycles are more clearly seen in the DEGG waveform. Fig. 9 shows the EGG waveform of sustained vowel /a/ phonated at G2 (98 Hz) and the corresponding DEGG waveform. The EGG waveform is characterized as period-doubling, and the DEGG waveform is characterized as double GCIs. A question arises as to whether the GCI with lower amplitude yields because of supraglottal adduction. Although further physiological experiments are needed, it is not unreasonable to assume that the GCI with lower amplitude is caused by the supraglottal adduction. The glottal opening is immediately followed by the supraglottal adduction, and the supraglottal adduction is followed by the actual glottal release (de-contact phase).

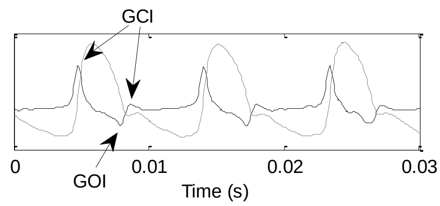


Fig. 9 EGG and DEGG waveform during sustained vowel /a/ phonated at G2.

5.2 Harshness

The Tibetan Buddhist chants are perceived as sounds with low and tense voice containing audible airflow noise. Regarding the harshness setting, which is one of the subcategories in the phonatory setting, Laver (1980) suggests,

‘This is the setting where the ventricular folds become involved in the phonation of the true vocal folds by squeezing closed the ventricle of Morgagni and pressing down on the true vocal folds, …… In order to bring the ventricular folds to this position, a high degree of muscular tension is needed, and the effect is normally to make phonation auditorily very harsh.’

Irregularity in pitch and spectral noise are the characteristics of harsh voice (Laver 1980). Fig. 10 compares a spectrogram of sustained vowel /a/ phonated at G2# (103.8 Hz) in chanting to that of speaking. The spectrogram of chanting shows noise lying over a wide frequency region. In contrast, significant noise is not found in the spectrogram of speaking.

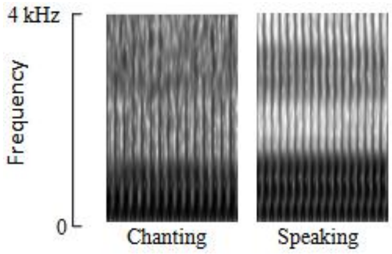


Fig. 10 Wide-band spectrograms of sustained vowel /a/ phonated at G2# in chanting and speaking.

Harsh and growl phonations are typical extended vocal techniques involving supraglottal oscillations. As mentioned above, the frequency of the ventricular fold oscillation is found as F0, F0/2, or F0/3 in previous literature. In the case of growl voice, the frequency of the aryepiglottic fold oscillation is found as F0/2 (Sakakibara et al. 2004). If we assume supraglottal oscillation in the chanting voice, its frequency is thought to be equal to that of vocal fold vibration, and supraglottal and glottal closures occur alternately. Then, the noise is due to excessive constriction and friction at the level of the supraglottal structure. The characteristics of these irregular vocalizations are often interpreted as period-doubling bifurcations and subharmonics. The subharmonics are usually recognized in the narrow-band spectrograms; however, they are not found in this study because the frequency of the supraglottal oscillation is equal to F0.

Period doubling in the EGG waveform, double GCIs in the DEGG waveform, and the widespread noise in the spectrograms are considered to result from the supraglottal activities that immediately follow the vocal fold opening.

6. CONCLUSION

Supraglottal constriction and adduction were estimated to occur in Tibetan Buddhist chants. Pressed phonation with a small spectral tilt resulting from the EGG and acoustic analyses, period doubling in the EGG waveform, double GCIs in the DEGG waveform, and the noise in the spectrograms were supportive evidence for these occurrences. Tibetan Buddhist chants of Kumbum Monastery maintain throat-singing traditions.

Further physiological research using tools, such as high-speed cameras, is needed to clarify the laryngeal vibratory mechanism of GCIs. Other styles of Tibetan Buddhist chants should also be investigated in future work.

7. ACKNOWLEDGEMENTS

This study was supported by China Social Sciences Funds (Grant No: 10&ZD125).

8. REFERENCES

[1] Baken, R. J. and R. F. Orlikoff (2000) *Clinical Measurement of speaking and voice*. New York: Delmar learning.

[2] Bickley, C. (1982) “Acoustic analysis and perception of breathy vowels.” *Speech Communication Group Working Papers*, 73-93. Cambridge, Mass: MIT, Research Laboratory of Electronics.

[3] Childers, D. G., D. M. Hicks, G. P. Moore and L. Eskenazi (1990) “Electroglottography and vocal fold physiology.” *Journal of Speech and Hearing Research* 33, 245–254.

[4] Childers, D. G. and A. K. Krishnamurthy (1985) “A critical review of electroglottography.” *Critical Reviews in Biomedical Engineering* 12, 131–161.

[5] Childers, D. G. and J. N. Larar (1984) “Electroglottography for laryngeal function assessment and speech analysis.” *IEEE Transactions on Biomedical Engineering* 31, 807–817.

[6] Childers, D. G., G. P. Moore, J. M. Naik, J. N. Larar and A. K. Krishnamurthy (1983) “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)*.

[7] Childers, D. G., J. M. Naik, J. N. Larar, A. K. Krishnamurthy and G. P. Moore (1983) “Electroglottography, speech and ultra-high speed cinematography.” In I. R. Titze and R. Scherer (eds.) *Vocal Fold Physiology and Biophysics of Voice*, 202–220. Denver: Denver Center for the Performing Arts.

[8] Fuks, L., B. Hammarberg and J. Sundberg (1998) “A self-sustained vocalventricular phonation mode: acoustical, aerodynamic and glottographic evidences.” *KTH TMH-QPSR* 3, 49–59.

[9] Henrich, N. (2004) “On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation.” *Journal of the Acoustic Society of America* 115, 1321-1332.

[10] Herbst, C. (2004) “Evaluation of various methods to calculate the EGG contact quotient.” Diploma Thesis, Department of Speech, Music and Hearing, KTH, Stockholm.

[11] Hollien, H., J. Michel and E. Doherty (1973) “A method for analyzing, vocal jitter in sustained phonation.” *Journal of Phonetics* 1, 85–91.

[12] Howard, D. M. (1995) “Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers.” *Journal of Voice* 9 (2), 163-172.

[13] Howard, D. M., G. A. Lindsey and B. Allen (1990) “Toward the quantification of vocal efficiency.” *Journal of Voice* 4, 205–212.

[14] Kong, J. (2001) *Lunyuyanfasheng* [On Language Phonation]. Beijing: Central Nationalities University Press.

[15] Kong, J and W-C. A. Liew (1998) “Wavelet analysis of speech phonations of Chinese speakers.” In *Proceedings of the 1st International Conference on Chinese Spoken Language Processing (Singapore)*.

[16] Laver, J. (1980) *The Phonetic description of voice quality*. Cambridge: Cambridge University Press.

[17] Lindestad, P., M. Södersten, B. Merker and S. Granqvist (2001) “Voice source characteristics in Mongolian ‘throat singing’ studied with high-speed imaging technique, acoustic spectra, and inverse filtering.” *Journal of Voice* 15(1), 78–85.

[18] Raymond, D. K. and J. B. Martin (2000) *Voice quality measurement*. California: Singular Publishing Group.

[19] Rothenberg, M. and J. J. Mahshie (1988) “Monitoring vocal fold abduction through vocal fold contact area.” *Journal of Speech and Hearing Research* 31, 338-351.

[20] Sakakibara, K-I. (2003) “Production mechanism of voice quality in singing.” *Journal of the Phonetic Society of Japan* 7 (3), 27-39.

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

[22] Smith, H., Stevens, K. N. and R. Tomlinson (1967) “On an unusual mode of chanting by certain Tibetan lamas.” *Journal of the Acoustic Society of America* 41, 1262-64.

[23] Stevens, K. N. (1977) “Physics of larynx behavior and larynx modes.” *Phonetica* 34, 264-279.

[24] Stevens, K. N. and H. Hanson (1995) “Classification of glottal vibration from acoustic measurements.” In O. Fujimura and H. Hirano (eds.) *Vocal fold physiology: Voice quality control*, 147–170. San Diego: Singular Publishing Group.

[25] Titze, I. R. (1990) “Interpretation of the electroglottographic signal.” *Journal of Voice* 4, 1-9.

[26] Titze, I. R., R. J. Baken and H. Herzel (1993) *Evidence of chaos in vocal folds vibration in Vocal fold physiology*. San Diego: Singular Publishing Group.