ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/252334550

Vocal Fold and False Vocal Fold Vibrations in Throat Singing and Synthesis of Khöömei

Article · January 2001

CITATIONS
36

READS
1,106

7 authors, including:



Ken-Ichi Sakakibara
Health Sciences University of Hokkaido
94 PUBLICATIONS **823** CITATIONS
SEE PROFILE



Emi Z Murano
Johns Hopkins University
70 PUBLICATIONS **961** CITATIONS
SEE PROFILE



Seiji Niimi
International University of Health and Welfare
213 PUBLICATIONS **2,326** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Emotional speech View project

LIver MRI View project

All content following this page was uploaded by Emi Z Murano on 10 June 2014.

The user has requested enhancement of the downloaded file.

Vocal fold and false vocal fold vibrations in throat singing and synthesis of khöömei

Ken-Ichi Sakakibara[*1], Hiroshi Imagawa[*2],
Tomoko Konishi, Kazumasa Kondo,
Emi Zuiki Murano[*2], Masanobu Kumada[*3], and Seiji Niimi[*4]

[*1] NTT Communication Science Laboratories, [*2] The University of Tokyo,
[*3] National Rehabilitation Center for the Disabled, [*4] International University of Health and Welfare

Abstract

We observed laryngeal movements in throat singing using physiological methods: the simultaneous recording of singing sounds, EGG, and high-speed digital images. We observed vocal fold and false vocal fold vibration and estimated the vibration patterns. We also estimated the laryngeal voices by using an inverse filtering method and simulated the vibration pattern using a new physical model: 2×2-mass model. From these observations, we propose a laryngeal voice model for throat singing and synthesis system of throat singing.

1 Introduction

Throat singing is a traditional singing style of people who live around the Altai mountains. Khöömei in Tyva and Khöömij in Mongolia are representative styles of throat singing. Throat singing is sometimes called biphonic singing, multiphonic singing, overtone singing, or harmonic singing because two or more distinct pitches (musical lines) are produced simultaneously in one tone. One is a low sustained fundamental pitch, called a drone, and the second one is a whistle-like harmonic that resonates high (in the range from 1 kHz to 3 kHz) above the drone.

Many variations of singing styles in throat singing are classified according to singers and regions. However, it is possible to objectively classify these variations in the terms of a source-filter model in speech production.

The laryngeal voices of throat singing can be classified into (i) a pressed voice and (ii) a kargyraa voice based on listener's impression, acoustical characteristics, and the singer's personal observation on voice production. The pressed voice is the basic laryngeal voice in throat singing and used as drone. The kargyraa voice is a very low pitched voice that ranges out of the modal register.

The production of the high pitched overtone is mainly due to the pipe resonance of the cavity from the larynx to the point of articulation in the vocal tract [1]. In Tyvan khöömei, sygit is a style where singers articulate by touching the tongue to the palate and khöömei is one where they articulate by pursing the lips.

We have physiologically observed two different laryngeal voices and estimated the patterns of the vocal fold and false vocal fold vibrations [6]. We have also simulated the vibration patterns by a physical modeling of the larynx: 2×2-mass model. Based on the physiological observations and the simulation, we propose a new laryngeal voice model and synthesis system for throat singing.

2 Physiological observations

2.1 Methods

We observed laryngeal movements in throat singing directly and indirectly by simultaneous recording of high-speed digital images, EGG (Electroglottography) waveforms, and sound waveforms (Fig. 1). The high-speed digital images were captured through a fiberscope inserted into the nose cavity of a singer at 4501 frames/s. Sound and EGG waveforms were sampled at 12 b/s and 18 kHz sf [4]. Two singers, who are normal, participated as subjects. One studied khöömei in Tyva and the other studied khöömij in Mongolia.

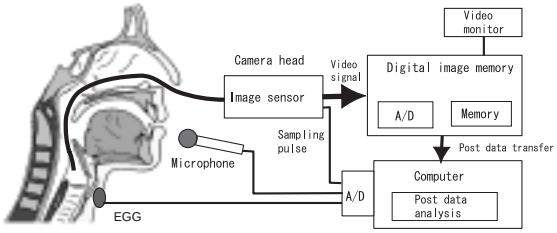


Fig.1: High-speed digital image system.

2.2 Results

Common laryngeal movements are observed among two singers for each of the two laryngeal voices.

contact: K.-I. Sakakibara, kis@brl.ntt.co.jp, NTT Communication Science Labs, 3-1, Morinosato Wakamiya, Atsugi-shi, 243-0198, Japan

Pressed voice

In pressed-voice production, the following features of the laryngeal movements were observed. (1) Overall constriction of the supra-structures of the glottis was observed, thus it was difficult to directly observe vibrations of vocal folds (VFs). (2) Vibration of the supra-structures of the glottis, whose edges are presumably false vocal folds (FVFs), was observed in digital high-speed images. (3) The period of FVFs vibrations was almost equal to the period of the EGG waveform. (4) The slope of the EGG curve changed in the beginning of the closed phase of the FVFs, the impedance of the EGG reached the maximal value when the FVFs were open, and reached the minimal value when they were closed (Fig. 2). The graph at the bottom of Fig. 2 depicts the locus of the edge of FVFs. The upper line (the lower line) is the locus of the left (right, respectively) edges of FVFs.

Kargyraa voice

In kargyraa-voice production, the following features of the laryngeal movement were observed. (1) Overall constriction at the supra-structures of the glottis was observed. (2) The constriction was looser than that in the case of the pressed voice. (3) Vibration of the supra-structures of the glottis, whose edges are presumably FVFs. (4) The phases of FVF vibrations are observed to alternate between almost completely closed and open. (5) Vibration of the VFs was observed during the open period of the FVFs. (6) The double period of vibration of the FVFs were equal to the period of the sound waveform. (7) When the FVFs almost completely closed, the power of sound became weaker. (8) In the EGG waveform, two different shapes alternated, and the period of the EGG waveform was equal to that of the sound waveform (Fig. 3).

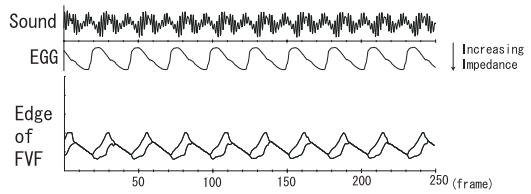


Fig. 2: Pressed voice
(from above, sound, EGG, edges of FVF).

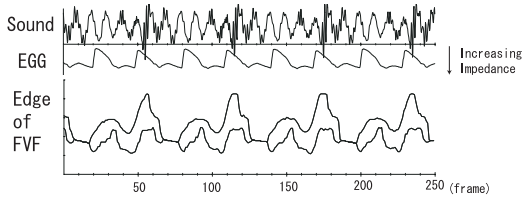


Fig. 3: Kargyraa voice
(from above, sound, EGG, edges of FVF).

2.3 Discussion

Two common features were observed among the mechanisms of the two different laryngeal voice productions: (1) Overall constriction of the supra-structures of the glottis and (2) vibration of the supra-structures of the glottis, which presumably are FVFs. These features are not observed in vowel production in ordinary speech. The differences among the two different laryngeal voice productions are (1) narrowness of the constriction and (2) the manner of FVF vibration.

The EGG waveforms for the pressed voice and karygraa voice represent the contact area of the supra-structures of the glottis as well as that of the VFs. However, taking into account the high-speed digital images and sound waveforms, the EGG waveforms can be assumed to mainly represent the contact area of VFs. Thus, we can conclude that VF vibrations and FVF vibrations have the opposite phase in the pressed-voice case . In the kargyraa voice, the FVFs can be assumed to close once for every two periods of closure of the VFs, and this closing blocks airflow and contributes to the generation of the subharmonic tone of kargyraa.

In a previous study, the open quotient (OQ) in throat singing was estimated to be smaller from the acoustical feature [2]. However, for both the pressed and kargyraa voice, our physiological observation suggests that the OQ is difficult to estimate because of the contribution of the supra-structures of the glottis. Therefore the OQ was not estimated.

In the synthesis of the throat singing sounds, as pointed out in [1], glottal source modeling is needed for reproduction of the timber. Our physiological observations suggests that the glottal source model of throat singing should include the FVF vibrations as well as the VF vibrations [7].

3 Laryngeal voice model of throat singing

In this paper, we define the glottal airflow as the airflow through glottis to the area between FVFs and the laryngeal airflow as the airflow through the area between FVFs to the pharynx.

Glottal airflow estimation

From recorded sounds, we estimated laryngeal airflow using the inverse filtering technique. In the pressed voice, the estimated laryngeal airflow curve had a small notch just after the curve reached a peak, and the closing of the VFs was apparently not complete

(Fig. 4). In the kargyraa voice, the estimated laryngeal airflow curve has two peaks in each period. From our physiological observation, the VFs vibrate twice in each period of the FVF vibration, and the estimated laryngeal airflow curve showed that in one of the two vibrations of VFs, the closing of VFs were not completed (Fig. 5).

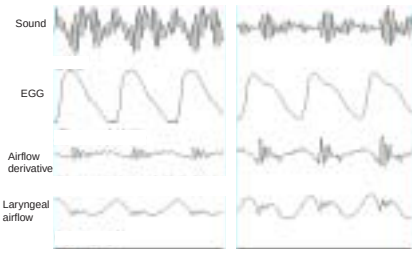


Fig. 4: Inverse filtered laryngeal airflow of pressed voices for two singers.

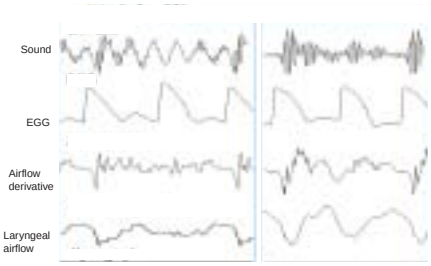


Fig. 5: Inverse filtered laryngeal airflow of kargyraa voices for two singers.

All the power spectra of the estimated glottal airflows showed an increase of power in the range from 1 to 3 kHz, which is where the second formant frequency which corresponds the whistle-like overtone appears in throat singing (Fig. 6–8).

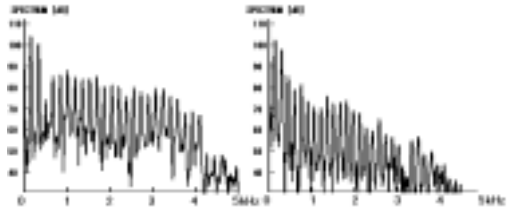


Fig. 6: Inverse filtered airflow spectrum of normal voice for two singers.

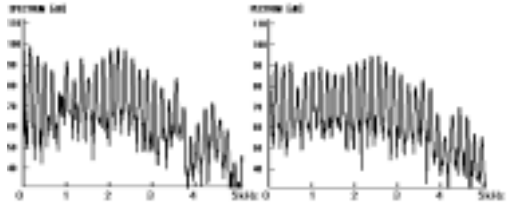


Fig. 7: Inverse filtered airflow spectrum of pressed voice for two singers.

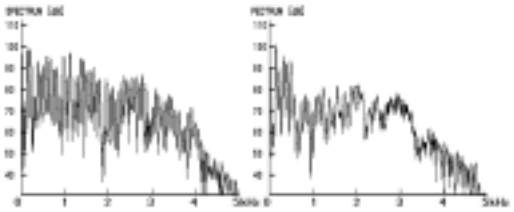


Fig. 8: Inverse filtered airflow spectrum of karygraa voice for two singers.

A 2×2-mass model

For a physical simulation of the VF and FVF vibrations, we propose a 2×2-mass model as a self-oscillating model of VF and FVF vibrations (Fig. 9). This model was devised by introducing a two-mass model for the FVFs to the ordinary two-mass model for the VFs. The mechanical transmission of vibrations between the VFs and FVFs were not considered. The laryngeal ventricle is a cylinder whose sectional area is uniformally 5 $cm^2$and height is 16 cm and not deformed. In the simulation the 2×2-mass model oscillated stably. The simulation of laryngeal movements using the 2×2-mass model agreed with the above assumptions for the two laryngeal movement patterns of throat singing for both the pressed and kargyraa voices (Fig. 10). The 2×2-mass model can simulate ordinary glottal source in the same way as the two-mass model by setting suitable model parameters [3].

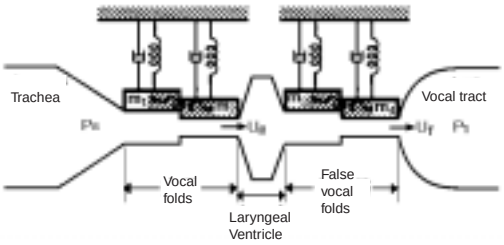


Fig. 9: 2×2-mass model for the VFs and FVFs.

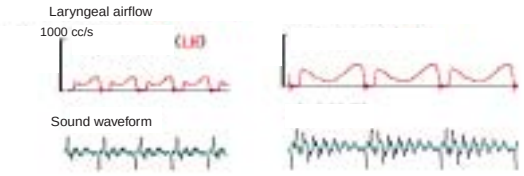


Fig. 10: Laryngeal airflow obtained by using 2×2-mass model
(left: pressed voice, right: kargyraa voice).

Laryngeal voice model

From the physiological observations and estimated laryngeal voices, we assume (1) in pressed-voice production, VFs and FVFs vibrate in almost opposite phase; (2) in karygraa-voice production, two closed

phases of the VFs appeared in one period of a glottal volume flow waveform, and VFs were incompletely closed at one of the two closed phases. Under these assumptions, we propose a laryngeal voice model for throat singing and synthesized throat singing sounds.

Our proposed laryngeal voice model is obtained as follows: We generate almost sine-shaped glottal airflow, because the glottal flow of the throat singing must be symmetric from Fig. 4 (Step 1). The glottal airflow is modulated by the vibration of the FVFs (Step 2). Turbulent noise is added according to the open width of the FVFs (Step 3). The output is convoluted with the transfer function of the laryngeal ventricle (Step 4)[3].

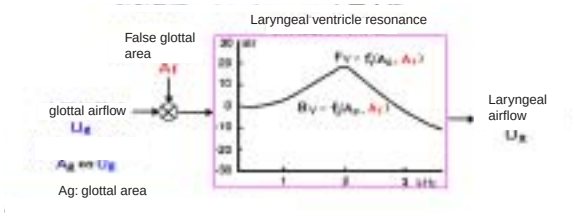


Fig. 11: Block diagram for laryngeal voice model.

4 Synthesis of throat singing

Based on a Klatt synthesizer [5], we propose synthesis model for throat singing, which has the proposed laryngeal voice model as source and time-varying formants obtained from recorded throat singing sounds as resonating filters (Fig. 12). Compared with an ordinary glottal airflow model, some improvements of the timbre were observed.

Conclusion

We observed the laryngeal movements in throat singing. The VF and FVF vibrations were observed. The FVF vibrations contribute to production of both the two laryngeal voices of throat singing. We also estimated the laryngeal voice source and simulated the laryngeal movements by using a 2×2-mass model. Based on these observations, we proposed a laryngeal source model and synthesis model for throat singing. These models can also simulate the normal voice. Consequently, all the power spectrum of the simulated glottal airflows showed the increase of the power on the range less than 3 kHz where the second formant frequency which corresponds the whistle-like overtone in throat singing. Our study indicates the glottal source also contributes the whistle-like overtone production as well as the articulation of the tongue and lips.

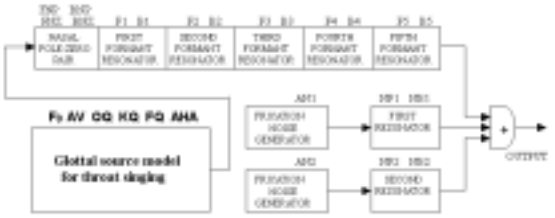


Fig. 12: Block diagram of khöömei synthesizer.

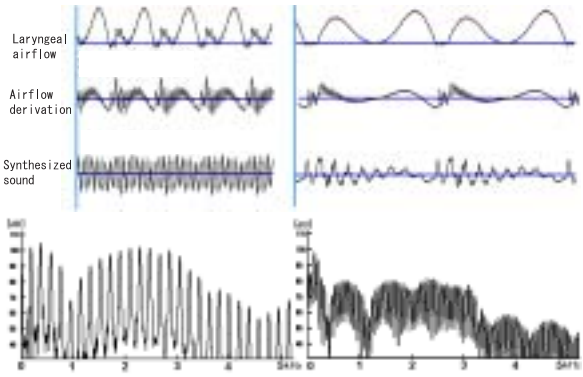


Fig. 13: Synthesized laryngeal airflows, synthesized sounds by khöömei synthesis system, and power spectra of sythesized souds (left: pressed voice, right: kargyraa voice).

Acknowledgments

We wish to thank Seiji Adachi, Zoya Kyrgys, Koichi Makigami, Naotoshi Osaka, Yoshinao Shiraki, and Masahiko Todoriki for their help and useful discussion.

Bibliography

[1] S. Adachi and M. Yamada. An acoustical study of sound production in biphonic singing xöömij. *J. Acoust. Soc. Am.*, 105(5):2920–2932, 1999.

[2] G. Bloothooft, E. Bringmann, M. van Cappellen, J. B. van Luipen, and K. P. Thomassen. Acoustics and perception of overtone singing. *J. Acoust. Soc. Am*, 92(4):1827–1836, 1992.

[3] H. Imagawa, K.-I. Sakakibara, T. Konishi, E. Z. Murano, and S. Niimi. Throat singing synthesis by a laryngeal voice model based on vocal fold and false vocal fold vibrations. *Tech. Rep. IECE*, SP2000-140:71–78, Feb. 2001. in Japanese.

[4] S. Kiritani, H. Imagawa, and H. Hirose. Vocal cord vibration in the production of consonants-observation by means of high-speed digital imaging using a fiberscope. *J. Acoust. Soc. Jpn. (E)*, 17:1–8, 1996.

[5] D. H. Klatt. Software for a cascade/parallel formant synthesizer. *J. Acoust. Soc. Am.*, 67(3):971–995, 1980.

[6] T. C. Levin and M. E. Edgerton. The throat singers of tuva. *Scientific America*, (Sep.1999):80–87, 1999.

[7] K.-I. Sakakibara, S. Adachi, T. Konishi, K. Kondo, E. Z. Murano, M. Kumada, M. Todoriki, H. Imagawa, and S. Niimi. Observation of vocal fold vibrations in tyvan and mongolian throat singing. *Tech. Rep. Musical Acoust., Acoust. Soc. Jpn*, 19-4:41–48, Sep. 2000. in Japanese.

View publication stats