

## Growl Voice in Ethnic and Pop Styles

Ken-Ichi Sakakibara<sup>1,2</sup>, Leonardo Fuks<sup>3</sup>, Hiroshi Imagawa<sup>4</sup>, Niro Tayama<sup>5,2</sup>

<sup>1</sup>NTT Communication Science Laboratories, NTT Corporation, Japan

<sup>2</sup>Department of Otolaryngology, The University of Tokyo, Japan

<sup>3</sup>School of Music, Universidade Federal do Rio de Janeiro, Brazil

<sup>4</sup>Department of Speech Physiology, The University of Tokyo, Japan

<sup>5</sup>International Medical Center of Japan, Japan

kis@b1.n1t1.co.jp leofuks@serv.com.ufrj.br

imagawa@m.u-tokyo.ac.jp ntayama@imc1.hosp.go.jp

### Abstract

Among the so-called extended vocal techniques, vocal growl is a rather common effect in some ethnic (e.g. the Xhosa people in South Africa) and pop styles (e.g. Jazz, Louis Armstrong-type) of music. Growl usually consists of simultaneous vibrations of the vocal folds and supra-glottal structures of the larynx, either in harmonic or sub-harmonic co-oscillation.

This paper examines growl mechanism using videofluoroscopy and high-speed imaging, and its acoustical characteristics by spectral analysis and model simulation. In growl, the larynx position is usually high and aryepiglottic folds vibrate. The aryepiglottic constriction is associated to a unique shape of the vocal tract, including the larynx tube, and characterizes growl.

### 1. Introduction

The term *growl* is originally referred to as low-pitched sounds uttered by animals, such as dogs, or similar sounds by humans, and therefore is mainly described by auditory-perceptual impression. Growl is widely observed in singing as well as in shouting and aroused speech.

The growl phonation has been also referred to as the phonation observed in some singing styles, such as the jazz singing style of Louis Armstrong<sup>1</sup> and Cab Calloway, [2, 3]. Many jazz, blues, and gospel singers often use growl in a similar manner. Besides such pop musics from North America, growl styles are widely found in pop music of other areas: in Brazil, samba singers, particularly in carnival lead voices, pop star Elza Soares, and country singing duo Bruno & Marrone; in Japan, Enka (a popular emotive style) singers, such as Harumi Miyako, employ it frequently. Some singers use growl extensively through a song, while others use it as a vocal effect for expressive emphasis.

<sup>1</sup>Satchmo's voice quality revealed roughness in the vocal fold level, however, he often used growl.

In ethnic music, one of the most prominent use of growl is found in *umngqokolo*, which is a vocal tradition of the Xhosa people in South Africa [11]. In Japanese theatre, Noh percussionist's voice, *Kakegoe*, may present growl at the beginning of phonation.

Growl may have perceptual similarities with the rough or harsh voice. In terms of phonetics, growl is sometimes described as the voiced aryepiglottic trill [3]. However, there is no clear evidence of its production mechanism, such as physiological observation of the aryepiglottic vibration.

In throat singing (Tyvan *khöömei* and Mongolian *khöömij*), ventricular and vocal fold vibration was observed for the two different laryngeal voices (drone and *kargyraa*) [4, 9]. In drone, the basic voice in throat singing with a whistle-like high overtone, the ventricular fold vibration is at the same frequency as the vocal fold vibration. In *kargyraa*, which usually sounds one octave (or more) lower than the modal register, the ventricular folds vibrate at  $f_0/2$  when the vocal folds vibrate at  $f_0$ . Moreover, some singers can do triple-periodic *kargyraa* in which the ventricular folds vibrate at  $f_0/3$ . In this paper, the phonation mode with ventricular and vocal fold vibration is called *VVM (vocal-ventricular mode)* [4]. In growl, there is no clear evidence of the ventricular fold vibration.

The growl, drone, *kargyraa*, as well as vocal fry, and some pathological voices may have similar perceptual characteristics related to roughness, creakiness, or harshness. Their acoustics may also have similar features. Therefore, clarifying differences among these phonations requires careful physiological observation.

In this paper, we examine the production mechanism of the growl phonation. Some of the authors (KIS, LF), who can utter several phonation modes, including the VVM, produced the growl phonation by carefully listening to and imitating various samples, as mentioned above. Observation of the laryngeal adjustment using endoscopic high-speed imaging and X-ray videofluoroscopy (partly reported in [1]), confirm the aryepiglottic vibra-

tion in growl. We also discuss the acoustical characteristics and differences between VVM (in particular, kargy-*raa*) and growl.

## 2. Three-tiered sphincter of the larynx

In the human larynx, there is a three-tiered sphincter comprising the vocal folds, the ventricular folds (false vocal folds), and the aryepiglottic sphincter [7] (Fig. 1).

The ventricular folds are incapable of becoming tense, since they contain very few muscle fibres. However, the ventricular folds can be constricted by the action of certain intrinsic laryngeal muscles. In the aryepiglottic region, the constriction is caused by the approximation of the tubercle of the epiglottis (anterior), aryepiglottic folds (lateral), and arytenoids (posterior). In normal phonation, the vibration of the ventricular and aryepiglottic folds is not observed.

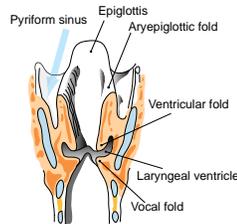


Figure 1: Coronal view of the larynx, as seen from behind.

## 3. X-ray observation

We observed the vertical laryngeal configuration of three different types of phonations (modal, “metallic”, and growl) using X-ray cinematography.

Fig. 2 shows a lateral X-ray view of the phonatory apparatus at rest. A wide pharyngeal space between the epiglottis and the arytenoids is observed. The cricoid cartilage is located at about the level of the fifth cervical vertebrae.

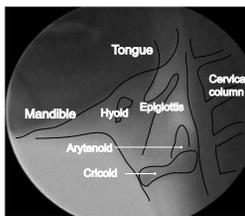


Figure 2: X-ray image of phonatory apparatus at rest, lateral view (subject: LF)

Fig. 3 shows the lateral X-ray views of three different voices: modal (left), “metallic” (center), and growl (right), in /y/ (close front rounded vowel). The metallic voice has a perceptually metallic impression and, in terms of usual phonetic usage, can be interpreted as pharyngealized, a little pressed (not necessarily tense), and raised-larynx. White lines are traced along the edges of

the cricoid, arytenoid, epiglottis, and cervical column. In modal phonation, a wide pharyngeal space is observed. The epiglottis doesn’t depress and its position is almost similar to that when it is at rest. In metallic and growl, the larynx is raised to about the level of the fourth cervical vertebrae. The epiglottis and arytenoid approximate very closely. There is no significant difference of the laryngeal adjustments between metallic and growl.



Figure 3: X-ray images of three different phonations of /y/ about in  $F_3 = 177$  Hz, lateral views. Left: modal. Center: metallic. Right: growl (Subject: LF).

## 4. High-speed images

We observed laryngeal movements in growl directly and indirectly by simultaneous recording of high-speed digital images, EGG (Electroglottography) waveforms, and sound waveforms. The high-speed digital images were captured at 4500 frames/s through an endoscope inserted into the mouth cavity of a singer. Sound and EGG waveforms were sampled at 12 b/s and 18 kHz sf.

In growl phonation, the aryepiglottic region is compressed antero-posteriorly, and the tubercle of the epiglottis and the arytenoid cartilages come into contact (Fig. 4). This antero-posterior compression is in good agreement with the lateral view of growl phonation in Fig. 3. Two-sided chinks generated by the contact of the epiglottic tubercle and arytenoids were observed. Each chink is surrounded by the epiglottis, arytenoid, and aryepiglottic fold. In some cases, both aryepiglottic folds vibrate in almost same phase (Fig. 5), and in other cases, the phases of both seem to be slightly different. Furthermore, in some cases, the vibration of the aryepiglottic folds is unstable and seems to be aperiodic.

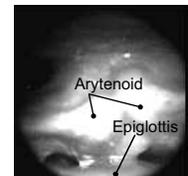


Figure 4: Aryepiglottic region in growl, as seen from above. Upper part is posterior (subject: KIS)

Fig. 5 shows the sound waveform (top), EGG waveform (middle, ordinate corresponds to total contact area of the larynx), and high-speed images. Vertical lines in the sound and EGG are synchronous to the last frames in each column of the high-speed images. The vibrations of the aryepiglottic folds are observed in the high-speed images. In this case, the aryepiglottic fold vibration is likely

to be periodic and the vibration of each side is mostly synchronous.



Figure 5: High-speed images of growl. Top: sound. Middle: EGG. Bottom: images. In images, frame step is 1/4500 ms (subject: KIS).

From the EGG and sound waveform, it is reasonable to conclude that the vocal folds vibrate half-periodically to the aryepiglottic fold vibration. This vibration pattern of the vocal and aryepiglottic folds is same as the VVM with  $f_0/2$ , i.e. kargyaa. The period-double vibration of the aryepiglottic folds generates subharmonics.

Neither the vocal nor ventricular folds were directly observed because the aryepiglottic folds were strongly constricted. Therefore, it is difficult to prove whether the vocal and ventricular fold vibrate or not. However, we conclude that the vocal and aryepiglottic folds vibrate and ventricular folds do not. The basis of this conclusion is as follows.

Smooth transition from modal to growl is frequently achieved by various singers and the subjects, therefore, it is reasonable to claim that, in growl, the vocal folds vi-

brate at almost opposite phases. To take account of the delay of the sound to the EGG, we consider that the maximal excitation of sound and the shape of the EGG waveform were mainly due to the vocal fold vibration. Next, if all three folds had simultaneously vibrated, the phases of their vibration would most likely have been different from each other by aerodynamical constraint. However, it is difficult to ascertain this phenomena from EGG waveform alone. To verify our claim, it is necessary to directly observe the movements of the three folds.

## 5. Acoustical analysis

Fig. 6 shows a spectrogram of the growl voice. Subharmonics appeared in growl. Similar subharmonic oscillation has been observed in kargyaa [4, 6, 9], and in some cases of vocal fry [10]. Perceptual clarification of differences among these phonations is important. Here, however, we focus on acoustical differences between growl and kargyaa.

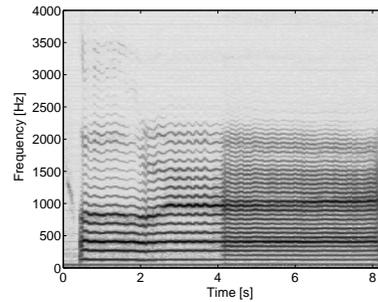


Figure 6: Spectrogram of modal to growl (subject: LF)

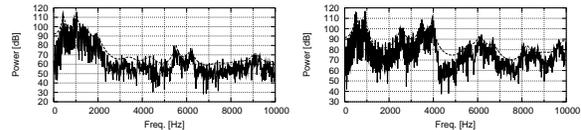


Figure 7: Power spectrum of growl (left) and kargyaa (right) of /o/ (subject: LF)

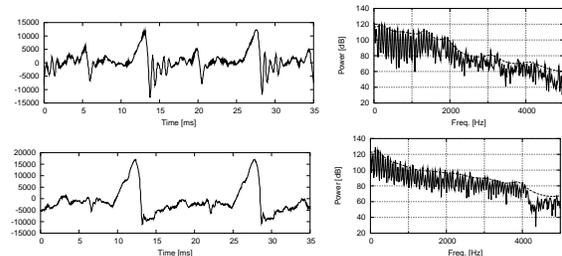


Figure 8: Inverse-filtered source of growl (top) and kargyaa (bottom). Left: sound waveform. Right: power spectrum. Subject: LF.

Fig. 7 shows the power spectra and spectral envelopes of growl and kargyaa. In growl, the range from above 2 kHz has very weak power. Fig. 8 shows the inverse-filtered source and its power spectrum of growl and kar-

gyraa. In growl, a pole is observed at about 1.5 kHz, whereas, in kargyraa, below 4 kHz, the power moderately decreases.

Physiologically, generation of subharmonics is concluded to be caused by the vocal fold vibration in vocal fry, ventricular fold vibration in kargyraa, and the aryepiglottic vibration in growl. In kargyraa, the ventricular fold constriction contributes to the generation of the laryngeal ventricle resonance, which appears as a zero in the laryngeal source. In growl, the aryepiglottic constriction constructs a deeper and larger cavity consisting of the laryngeal ventricle, ventricular fold region, and laryngeal vestibule (Figs. 1, 3, 4). Therefore, the resonance frequency of the cavity must be lower than that of the laryngeal ventricle. Fig. 9 shows the spectra of the synthesized laryngeal source obtained using the two-by-two mass model [8]. For simplicity, the aryepiglottic and ventricular fold vibration and vocal tract are omitted. The pole in the source of growl is at about 1.5 kHz and is lower than in kargyraa.

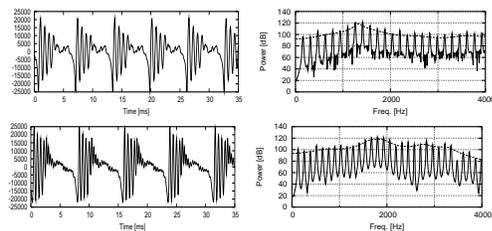


Figure 9: Synthesized sources and spectra of growl (top) and kargyraa (bottom) using the two-by-two mass model.

We also roughly calculated the resonance frequencies of the laryngeal ventricle for kargyraa and the laryngeal cavity for growl by using a Helmholtz resonator. In kargyraa, we assume that the body cylinder (the laryngeal ventricle) has 0.4 cm height and 1.5 cm<sup>2</sup> area and the neck cylinder (the ventricular fold region) 0.8 cm height and several areas. In growl, we assume the body has 2.0 cm height and a 1.02 cm<sup>2</sup> cross-sectional area, and the neck (the aryepiglottic area) 0.4 cm height and several areas (Table 1). If the constricted regions have equal area, the resonance frequency of the source in growl is always lower than that in kargyraa.

Table 1: Resonance frequencies in growl and kargyraa, calculated by a Helmholtz resonator.

Area of neck (cm <sup>2</sup> )	Resonance freq. in growl (Hz)	Resonance freq. in kargyraa (Hz)
0.05	1232	1694
0.10	1674	2342
0.25	2465	3550
0.50	3252	4807
1.00	4227	6429

## 6. Discussions and conclusions

In growl, the larynx position is higher than in the modal case, and the aryepiglottic region is strongly approxi-

mated. The aryepiglottic folds vibrate, as well as vocal folds, and contribute to the subharmonic oscillation. The resonance frequency of the cavity induced by the aryepiglottic constriction is lower than that of the laryngeal ventricle, and this characterizes the growl voice.

The mechanism of the supraglottal constriction is still controversial. The supraglottal constriction is widely considered to be caused by an activity of the aryepiglottic muscle, however, from our physiological observations and previous histological observation of the supraglottal muscles [5], the constrictions of the aryepiglottic and ventricular folds are presumably caused by different mechanisms.

The power of the subharmonics in growl is seemingly lower than in kargyraa, but further analysis is needed to clarify this. Perceptual evaluation of differences among various subharmonic phonations, such as growl, kargyraa, and vocal fry, will be addressed as future work. Analysis of other perceptually similar singing styles, such as Sardinian singing, will also be addressed as future work.

## Acknowledgments

We thank Samuel Araújo, Parham Mokhtari, Seiji Niimi, Makoto Ogawa, Satoshi Takeuchi, and Mamiko Wada for their helpful discussions.

## References

- [1] S. Araújo and L. Fuks. Práticas vocais no samba carioca: um diálogo entre a acústica musical e a etnomusicologia. In N. M. Claudia and T. M. Refnada and T. Elizabeth Ed., *Ao encontro da Palavra Cantada: poesia, música e voz*, pp.278–288, Viveiros de Castro Ltda., 2001.
- [2] J. C. Catford. *Fundamental Problems in Phonetics*, Edinburgh Univ. Press., 1977.
- [3] J. H. Esling. Pharyngeal consonants and the aryepiglottic sphincter, *J. International Phonetics Association*, 26(2):65–88, 1996.
- [4] L. Fuks, B. Hammarberg, and J. Sundberg. A self-sustained vocal-ventricular phonation mode: acoustical, aerodynamic and glotto-graphic evidences, *KTH TMH-QPSR*,3/1998:49–59, 1998.
- [5] M. Kimura, K.-I. Sakakibara, H. Imagawa, R. Chan, S. Niimi, and N. Tayama. Histological investigation of the supra-glottal structures in human for understanding abnormal phonation, *J. Acoust. Soc. Am.*, 112:2446, 2002.
- [6] P.-Å. Lindestad, M. Sodersten, B. Merker, and S. Granqvist. Voice Source Characteristics in Mongolian “Throat Singing” Studied with High-Speed Imaging Technique, Acoustic Spectra, and Inverse Filtering, *J. Voice*, 15(1):78–85, 2001.
- [7] J. J. Pressman. Sphincters of the larynx, *A. M. A. Arch. Otolaryngol.*, 59(2):221–236, 1954.
- [8] K.-I. Sakakibara, H. Imagawa, S. Niimi, and N. Osaka. Synthesis of the laryngeal source of throat singing using a 2×2-mass model, *Proc. ICMC 2002*, 5–8, 2002.
- [9] K.-I. Sakakibara, T. Konishi, K. Kondo, E. Z. Murano, M. Kumada, H. Imagawa, and S. Niimi. Vocal fold and false vocal fold vibrations and synthesis of kh’ò’omei, *Proc. of ICMC 2001*, 135–138, 2001.
- [10] R. L. Whitehead, D. E. Metz, and B. H. Whitehead. Vibratory patterns of the vocal folds during pulse register phonation, *J. Acoust. Soc. Am.* 75(4):1293–1297, (1984).
- [11] H. Zemp, Ed. *Les Voix du Monde — Une anthologie des expressions vocales*. 3 vol. CDs with book, CMX374 1010.12, CNRS/Musée de l’Homme, 1996.