False vocal fold surface waves during *Sygyt* singing: A hypothesis

Chen-Gia Tsai[1]*, Yio-Wha Shau*[2]*, and Tzu-Yu Hsiao*[3]

[1]Graduate School of Folk Culture and Arts, Taipei National University of the Arts, Taipei, Taiwan; [2]Institute of Applied Mechanics, National Taiwan University, Taipei, Taiwan; [3]Department of Otolaryngology, National Taiwan University Hospital, Taipei, Taiwan

`tsai_chen_gia@yahoo.com.tw`

Abstract

Overtone singing is a vocal technique found in Central Asian cultures, by which one singer produces a high pitch of *nF*0 along with a low drone pitch of *F*0. The pitch of *nF*0 arises from a very sharp formant. Current physical modeling of overtone singing asserts that the harmonic at *nF*0 is emphasized by a resonance of the vocal tract. However, this approach could not explain the extraordinarily small bandwidth of this formant.

This paper offers a hypothesis that surface waves (Rayleigh waves) of the false vocal folds might actively amplify the harmonic at *nF*0 in a specific technique of overtone singing: *Sygyt*. We propose a loop for harmonic amplification, which is composed of (1) the vocal tract with resonance *nF*0, (2) surface waves of the false vocal folds, and (3) a varicose jet separating from the false folds. This model receives indirect support from an experimental study on a novel human vocalization, which is characterized by a prominent component at 4 kHz. During this pure tonal vocalization, false fold surface vibrations were detected by ultrasound color Doppler imaging. High-frequency false fold surface waves may also occur during *Sygyt* singing.

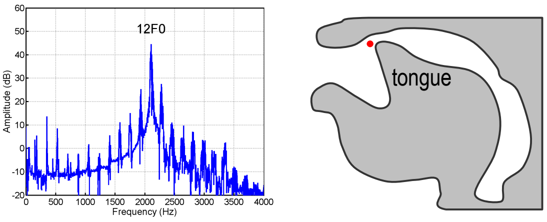
1. Introduction

Overtone singing (or throat singing, biphonic singing) is a vocal technique found in Central Asian cultures such as Tuva and Mongolia, by which one singer produces a high pitch of *nF*0 along with a low drone pitch of *F*0 (*F*0 is the fundamental frequency, *n* = 6, 7, ...13 in typical performances). The voice of overtone singing is characterized by a sharp formant centered at *nF*0, as can be seen in Figs. 1 and 2. Traditional techniques of overtone singing include *Khoomei*, *Sygyt*, *Kargyraa* and others.

There are two approaches of physical modeling of overtone singing: (1) the double-source theory [1], which asserts the existence of a second sound source that is responsible for the melody pitch; and (2) the resonance theory, which asserts that a harmonic is emphasized by an extreme resonance of the vocal tract. The fact that the melody pitches producible by the singer are limited to the harmonic series of the drone was regarded as robust support of the resonance theory [2].

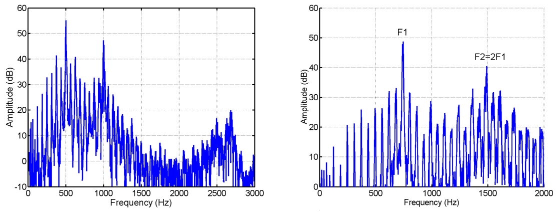
Recent attempts of physical modeling of *Sygyt* were concerned with calculation of the transfer function of the vocal tract using one-dimensional models, successfully predicting the formant frequency [2,3]. From a theoretical standpoint, however, this approach may not be suitable for the tract with a rapidly flaring bell section. A *Sygyt* singer raises the tongue so that the tract shape changes abruptly at the narrowing of the tongue (marked with a red dot in Fig. 1b), where the assumption of planar wave fronts breaks down, and evanescent cross-modes can be excited in this flaring section even at low frequencies [4]. This may leads to errors in transfer function calculation using one-dimensional models. An alternative approach of Matched Asymptotic Expansions for modeling a *Sygyt* singer's vocal tract was proposed in [5].

In a two-resonator theory, a *Sygyt* singer's vocal tract was modeled as a coupled system of a longitudinal resonator that was from the glottis to the narrowing of the tongue, and a Helmholtz resonator that was from the articulation by the tongue to the mouth exit. Experiments showed that for some *Sygyt* voices with a sharp formant two resonances were matched, while a melody pitch can be perceived even in the case of not exactly matched resonances [6]. Although the formant magnitude was shown to be increased by resonance-matching [3], it is unclear whether resonance-matching will reduce the formant bandwidth.



(a) (b)

Figure 1: *(a) Spectrum of a Sygyt voice produced by a singer from Tuva. (b) Vocal tract shape of a Sygyt singer, based on [2]. Because of rapid flaring, the region at the narrowing of the tongue is "compact"; the acoustic field is locally governed by Laplace equation [5].*



(a) (b)

Figure 2: *Two spectrum snapshots of a song produced by a Kargyraa singer from Tuva. The center frequency of the second formant is twice the first one. This "mode-locking", holding at every instant in this song, cannot be explained by tract filtering. An unknown glottal source may produce the outstanding component at F1 and its second harmonic.*

From a psychoacoustic point of view, a small bandwidth of the prominent formant is critical to a clear melody in *Sygyt* singing. A preliminary study using an autocorrelation model for pitch extraction suggested that the pitch strength of $nF0$ increased along with the Q value of this formant, with the formant magnitude playing a secondary role [5]. The spectrum of the *Sygyt* voice shown in Fig. 1a has the 12th harmonic approximately 15 dB stronger than its flanking components. If the amplification of this harmonic cannot be explained in terms of vocal tract impedance, it should be attributed to the source signal.

The insufficiency of the resonance theory is even more notable in another technique of overtone singing: *Kargyraa*. A *Kargyraa* singer uses his false vocal folds to produce a low-pitched drone, manipulating his mouth opening to change the vocal tract resonance. Spectra in Fig. 2 show that the center frequencies of the first and second formants of *Kargyraa* voices always stand in the ratio of 1:2. This strange phenomenon suggests an unknown glottal source that produces the outstanding component at $F1$, and its second harmonic.

The goal of this study is to offer a physical model based on a nonlinear loop that explains the harmonic amplification in *Sygyt*. This model asserts that surface waves (Rayleigh waves) of the adducted false vocal folds can actively amplify a harmonic. We first discuss the interactions between the false vocal fold surface waves (FVFSWs), the glottal flow and acoustic waves. A preliminary experiment that provided indirect evidence of this model is then addressed.

2. Theory

2.1. Rayleigh surface waves

The Rayleigh surface wave is a specific superposition of a transverse wave and a longitudinal wave of an elastic solid (see, e.g. [7]). Its amplitude is significant only near the surface and attenuates exponentially with the depth. The trajectories of material particles are ellipses. At the surface the normal displacement is about 1.5 times the tangential displacement. The velocity of Rayleigh waves, independent on the wavelength, is about 0.9 times the transverse wave velocity. Rayleigh's theory of surface waves has been generalized to viscoelastic solids (see, e.g. [8]).

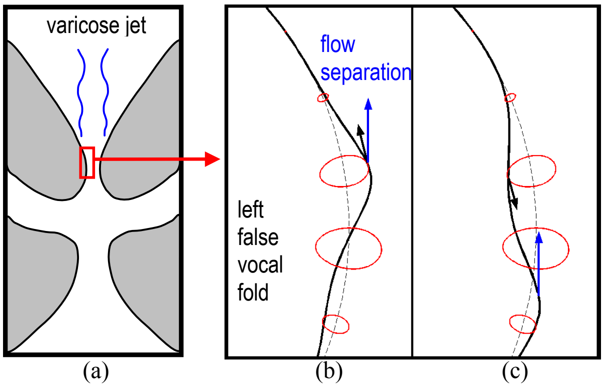
The assumption of Rayleigh surface wave on the false vocal folds is supported, although indirectly, by recent measurements of the medial surface dynamics of the vocal folds [9]. The trajectories of fleshpoints were approximately ellipses, with the length ratio of the two axes varying in the range of 1.5-2.0. This value is in remarkable agreement with Rayleigh's theory of surface waves.

2.2. Physical modeling of *Sygyt*

Here we propose a physical model that describes how FVFSWs absorb the energy of the glottal flow and acoustic waves.

The false folds are significantly adducted during *Sygyt* singing. Hence, the volume flow through them (U_F) is sensitive to FVFSWs. FVFSWs are supposed to be triggered by the acoustic pressure, which is predominated by the resonance of the vocal tract $nF0$. So we assume a FVFSW with the frequency of $nF0$.

Based on the assumption of elliptic movements of fleshpoints on the false folds, snapshots of this wave can be obtained. The ellipses in Figs. 3b and 3c represent the trajectory of fleshpoints. We estimate the energy exchange between the flow and the tissue occurs at one point. In Fig. 3b the work done by the viscous flow at this point is positive. In Fig. 3c the flow separates upstream, performing no work (or positive work, if back-flow appears) at this point. It can easily be seen that over a period the FVFSW absorbs energy from the flow in the vicinity of the flow separation point, which moves back and forth at a crest of the FVFSW, modulating the flow through the false folds at frequency of $nF0$. This induces varicose oscillations of U_F, which produce the harmonic at $nF0$ in the source signal. This harmonic is in turn reinforced by the strong vocal tract resonance at $nF0$.



(a) (b) (c)

Figure 3: *False vocal fold adduction and snapshots of the surface wave on the left false fold. The dashed curve represents the rest position of the surface. See the text.*

The net work done by the sinusoidal acoustic wave with frequency $nF0$ at a point on the false fold over a period can be positive or negative, depending on the phase relationship between the FVFSW and the acoustic pressure. We suppose that within a half wavelength of the FVFSW in the vicinity of the flow separation point, the FVFSW absorbs the acoustic energy of the harmonic at $nF0$. Away from this flow separation point, the FVFSW is expected to decay rapidly because of large viscous losses in the tissue during high-frequency vibrations. We thus conclude that the total work done by the acoustic wave on the FVFSW is positive.

To sum up, a loop for *Sygyt* is established in terms of (1) linear resonator: the vocal tract with resonance at $nF0$, (2) energy source: pressure difference across the false glottis, and (3) nonlinear amplifier: a flow separating from curved walls with mucosal layers receiving acoustic feedback. This self-sustained oscillator differs from the true vocal folds in that the false fold mucosa does not vibrate at any intrinsic resonance, but rather respond to the acoustic pressure.

2.3. Discussion

The present model explains the crucial role of the adduction of the false folds in *Sygyt* technique. Because of this adduction the flow velocity over their mucosal layers is high enough to supply the energy for sustaining FVFSWs. It is interesting to note that FVFSWs have been observed in patients suffering from ventricular dysphonia [10], although their frequencies appeared to be much lower than those during *Sygyt* singing.

From an empirical standpoint, learning *Sygyt* is much more difficult than it is implicated by the resonance theory. In workshops of overtone singing, it has been repeatedly observed that only very few people are able to produce voices with a clear melody pitch. The present model predicts that one cannot sing *Sygyt* well even when manipulating the tract shape perfectly, because his false folds are not correctly adducted, or their mucosal layers do not have a proper shape, thickness, and viscoelastic properties.

The loop described in our model tends to “unify” the double-source theory and the resonance theory of overtone singing. Whereas the true vocal folds and the vocal tract are, as usual, viewed as the independent source and filter, the false fold mucosa plays a key role in introducing acoustic feedback into the loop for harmonic amplification.

The present model for *Sygyt* might also shed new light on the production of high-frequency, whistle-like voice type of birds, dolphins, whales, and groaning dogs. In this regard, our model is an updated version of the double-source theory [1], which already drew parallels between the sounding mechanisms of overtone singing and the whistle-like voice type, which is produced with the false folds adducted.

It is interesting to compare the harmonic-amplification loop with the sounding mechanism of flute-type instruments, which is based on a loop composed of a vibrating jet and acoustic waves filtered by a resonator. In the case of flutes the jet separates from the musician’s lips, traveling along the mouth of the resonator towards a sharp edge. When the instrument produces a tone, the jet oscillates at one of the resonances of the pipe. The acoustic flow field near the flow separation point excites sinuous oscillations of the jet. At the sharp edge, the jet is directed alternately toward the inside and the outside of the resonator. This pulsing injection induces an equivalent pressure difference across the mouth that excites and maintains acoustic waves in the pipe [11]. The jet, like the false fold mucosa, does not vibrate at any intrinsic resonance. It should be noted that the acoustic *flow* induces *sinuous* oscillations of the jet at the mouth hole of a flute, whereas the acoustic *pressure* excites FVFSWs that induce *varicose* oscillations of the glottal flow.

While a varicose jet is essential for whistle-like sound production, the role of wall vibration is not fully understood. It has been suggested that the sounding mechanism of human whistling is a loop composed of the jet and the oral cavity with a prominent resonance. The pressure fluctuations due to the acoustic wave at the flow separation point could induce varicose oscillations of the jet without any wall vibration. This model is in an interesting contrast to our model of *Sygyt*, which assumes vibrations of the compliant walls. To examine the assumption of FVFSWs in our model of *Sygyt*, we measure surface vibrations during whistle-like singing *in vivo*.

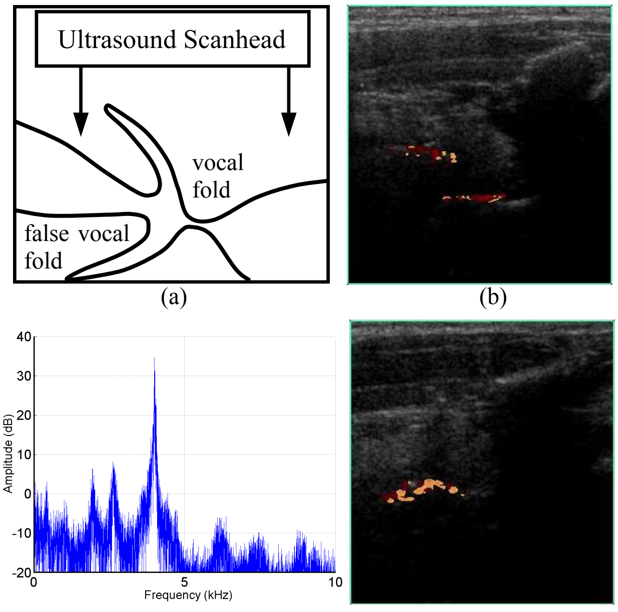
3. Experimental Study

3.1. Whistle-like voice type

The present model of “varicose jet oscillations induced by surface waves of curved walls in the vicinity of the flow separation point” may provide insight into the production of the whistle-like voice type in birds and mammals. It has been suggested that the production mechanism of bird whistled song might be related to a retraction of the syringeal membranes while in oscillation so that they no longer completely close, leading to a great reduction in the harmonic content of the flow. An alternative explanation of whistled song is that it is produced by pure aerodynamic means without any vibrating surfaces [12]. However, recent experimental studies favor the sounding mechanism of vibrating surface [13,14].

After some practice, human can imitate dog’s groaning to produce high-frequency whistle-like voices, which have a prominent component approximately at 4 kHz, as shown in Fig. 4c. We hypothesize that the mechanism underlying this vocalization is a varicose jet induced by FVFSWs.

Medical ultrasound (US) provides an ideal noninvasive method for observing high-frequency surface vibrations with small amplitude, because the vibratory artefact of color Doppler imaging (CDI) detects surface velocity rather than displacement. In previous studies, the CDI was used to measure the frequency and the length of the vocal folds during normal phonation [15,16]. In the present experiment we employ this technique to detect FVFSWs during whistle-like singing.



(a) (b)

(c) (d)

Figure 4: *(a) Schematic of US coronal scan of the glottis. (b) Display of CDI color artefacts during a breathy vocalization. Surface vibrations on the right vocal fold and false fold can be observed. (c) Spectrum of a pure tonal voice. (d) Display of CDI color artefacts during this vocalization. Surface vibrations on the right false vocal fold can be observed.*

3.2. Methods

A commercially available, high resolution US scanner (HDI-5000, ATL, Bothell, WA) with a 5- to 12-MHz linear-array transducer (L12 to 5 38 mm, ATL) was used in this study. The frame rate in B-mode was about 25 Hz. In the color mode, the pulse-repetition rate was 10,000 Hz and the measuring velocity range was set at 0 to 128.3 cm/s with baseline offset, which resulted in a frame rate of about 7 Hz. The US scan head was placed horizontally at the midline of the thyroid cartilage lamina on one side (Fig. 4a). The subject is the first author of this paper, who is a healthy man aged 33 with normal vocal function. For this experiment he had practiced the whistle-like vocalization for a week.

3.3. Results

CDI color artefacts detected surface vibrations of the right false vocal fold during pure tonal singing (Fig. 4d). During warming up of this vocalization, surface vibrations of the right vocal fold and the false fold were observed (Fig. 4b).

The frequency of pure tonal singing was found to range from 3.7 kHz to 4.6 kHz. Out of this range the voice lose the pure tonal characteristic, with breathy noises accumulating at the prominent resonance.

4. Concluding Remarks

The observation of false fold surface vibrations during pure tonal singing provides indirect support of our model for *Sygyt*. As FVFSWs may generate 4 kHz pure tonal voices with the second harmonic 30 dB (or more) weaker than the fundamental, it should be possible that a *Sygyt* singer amplifies a selected harmonic of the voice produced by the true vocal folds through FVFSWs.

The role of acoustic feedback in FVFSW generation is not fully understood. When the acoustic wave filtered by the resonator is strong enough to trigger FVFSWs, a loop for pure tonal vocalization may be established. If not, periodic FVFSWs may not occur. The laryngeal ventricle may be the Helmholtz resonator that is responsible for the prominent resonance at 3.7-4.6 kHz. However, this “resonance” model appears against experimental results about bird’s pure tonal vocalization [13,14]. If the frequency of surface waves is not determined by the tract resonance, it should be determined by the tissue curvature, elastic properties, and the flow speed. In the case of *Sygyt* singing, however, it has not been reported that a singer manipulates the false folds to change the melody pitch. Further research is needed to compare the sounding mechanisms of *Sygyt* singing and the pure tonal vocalization.

One implication of our surface wave model is that the vertical motion of fleshpoints on the true/false vocal folds may be critical to their self-sustained oscillation. The two-mass and three-mass models of the vocal folds [17,18] do not take into account the ellipse-like motion of vocal fold fleshpoints, which is consistent with Rayleigh’s theory of surface waves and has been demonstrated in excised canine larynx experiments [9]. We suggest that the vertical motion of fleshpoints near the flow separation point can absorb the kinetic energy of the glottal flow through viscous shear force.

The effect of surface viscous shear stress exerted by a flow also plays a central role in the system of a pair of fluttering flags in wind. This system shows some notable similarities of the glottis. When the inter-flag distance lies in a definite range the flags flutter in an out-of-phase state and generate a pulsating flow, with striking similarities of the vocal fold vibration in the chest register. Flow visualizations showed significant shear stress on the flags exerted by the flow [19]. This finding suggests that viscous shear stress on the vocal fold mucosa should not be ignored, especially in the vocalizations with a large open quotient.

Next to the viscosity effect, the surface shear stress may be attributed to the carrying-along of the varicose flow. It was observed in a pair of flags that the flag wave propagates along with the flow, while the wave of an isolated flag propagates in the direction opposite to the flow. Note that the surface shear stress dominates the system of a pair of flags but not an isolated flag [19]. It is likely that the surface shear stress is due to the effect that a varicose or sinuous flow carries along the flag wave. This approach may shed new light on the mechanism of the self-sustained oscillation of the vocal folds.

5. References

[1] Chernov, B.; and Maslov, V. 1987. Larynx double sound generator. *Proc. XI Congress of Phonetic Sciences*, Tallinn 6, 40-43.

[2] Adachi, S.; and Yamada, M. 1999. An acoustical study of sound production in biphonic singing, Xöömij. *J. Acoust. Soc. Am.* 105(5), 2920-2932.

[3] Kob, M. 2002. *Physical modeling of the singing voice.* PhD thesis, Aachen University (RWTH).

[4] Pagneux, V.; Amir, N.; and Kergomard, J. 1996. A study of wave propagation in varying cross-section waveguides by modal decomposition. Part I. Theory and validation. *J. Acoust. Soc. Am.* 100, 2034-2048.

[5] Tsai, C.G. 2004. Physics and perception of overtone singing. URL: http://jia.yogimont.net/overtonesinging/

[6] Kob, M.; and Neuschaefer-Rube, C. 2004. Acoustic properties of the vocal tract resonances during *Sygyt* singing. *Proc. of the International Symposium on Musical Acoustics*, Nara, Japan.

[7] Achenbach, J.D. 1984. *Wave propagation in elastic solids*. Elsevier, New York.

[8] Romeo, M. 2001. Rayleigh waves on a viscoelastic solid half-space. *J. Acoust. Soc. Am.* 110 (1), 59-67.

[9] Berry, D.A.; Montequin, D.W.; and Tayama, N. 2001. High-speed digital imaging of the medial surface of the vocal folds. *J. Acoust. Soc. Am.* 110(5), 2539-2547.

[10] Nasri, S.; Jasleen, J.; Gerratt, B.R.; Sercarz, J.A.; Wenokur, R.; and Berke, G.S. 1996. Ventricular dysphonia: a case of false vocal fold mucosal traveling wave. *Am. J. Otolaryngol.* 17(6), 427-431.

[11] Verge, M.P.; Caussé, R.; Fabre, B.; Hirschberg, A.; Wijnands, A.P.J.; and van Steenbergen, A. 1994. Jet oscillations and jet drive in recorder-like instruments. *Acustica* 2, 403-419.

[12] Gaunt, A.S.; Gaunt, S.L.L.; and Casey, R.M. 1982. Syringeal mechanics reassessed: evidence from Streptopelia. *Auk* 99, 474-494.

[13] Brittan-Powell, E.F.; Dooling, R.F.; Larsen, O.N.; and Heaton, J.T. 1997. Mechanisms of vocal production in budgerigars (Melopsittacus undulatus). *J. Acoust. Soc. Am.* 101, 578-589.

[14] Ballintijn, M.R.; and Cate, C.T. 1998. Sound production in the collared dove: a test of the ‘whistle’ hypothesis. *J Experimental Biology* 201, 1637-1649.

[15] Shau, Y.W.; Wang, C.L.; Hsieh, F.J.; and Hsiao, T.Y. 2001. Noninvasive assessment of vocal fold mucosal wave velocity using color Doppler imaging. *Ultrasound Med. Biol.* 27, 1451-1460.

[16] Hsiao, T.Y.; Wang, C.L.; Chen, C.N.; Hsieh, F.J.; and Shau, Y.W. 2002. Elasticity of human vocal folds measured in vivo using color Doppler imaging. *Ultrasound Med. Biol.* 28, 1145-1152.

[17] Ishizaka, K.; and Flanagan, J.L. 1972. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst. Tech. J.* 51(6), 1233-1268.

[18] Story, B.H.; and Titze, I.R. 1995. Voice simulation with a body cover model of the vocal folds. *J. Acoust. Soc. Am.* 97, 1249-1260.

[19] Zhang, J.; Childress, S.; Libchaber, A.; and Shelley, M. 2000. Flexible filaments in a flowing soap film as a model for one-dimensional flags in a two-dimensional wind. *Nature* 408, 835-839.