Comparison of voice production types of ‘western’ overtone singing and South Siberian throat singing

Sven Grawunder
University of Halle, Germany
E-mail: grawunder@sprechwiss.uni-halle.de

ABSTRACT

The investigator uses a non-invasive methodology based on a synchronous recording of voice, laryngographic signal and subglottal resonance on 2 ‘European’ overtone singers and 2 ‘Siberian’ semi-professional throat singers. For the upper registers of throat singing an extremely short contacting phase was obtained and SQ seems to be appropriate to distinguish between both phonation types. Lower registers share doublet and triplet waveform patterns, but differ significantly in jitter and shimmer rates ($j < 0.1$; $sh < 0.15$ for ThS). Other findings are a narrowed bandwidth of a dominant F2 and an ostinato formant between F3 and F4, which are remarkable for all throat singing styles.

1. INTRODUCTION

The terms overtone singing (OtS) and throat singing (ThS) are still often taken as synonyms but differ after a closer look extremely in their ‘voice technique’. Singers of both types can differentiate at least 5 voice production modes. So are for overtone singing modal register (with 2 different articulatory settings) and sometimes vocal fry (‘Strohbass’ register) described [13]. For throat singing the classification matches for the most part with traditional terminologies, which otherwise indicate a broader stylistic variety [11]. Therefore it needs to be assumed that there are certain similarities among styles of the same timbre and range. So for this study only articulatory similar OtS and ThS styles of the same assumed phonation type and register have been selected. Each of these ThS styles inhere 2 ‘substyles’ (tuvan names: höömei, syghyt, kargyraa type 1, kargyraa type 2). Endoscopic examinations [[7],[8] [9]] discovered that a ventricular-vocal fold register, an aryepiglottic-vocal fold mode but also an aryepiglottic-ventricular fold register [9] in throat singing are productively used. The paper presents an investigation that serves as pilot study for a corpus of 20 singers from southern Siberia (including Mongolia and the Republics of the Russ. Fed. Tuva, Gorno-Altai, Hakassia). The intention is to find significant characteristics that allow a distinction between ‚western’ (modal) overtone singing and South Siberian throat singing and that serve as model for further studies in other fields based on a non-invasive methodology.

2. METHODS

2.1 Material

The singers were asked to produce sustained passages of different voice production modes and of different vowel qualities (/o/, /u/, /a/, /e/, /i/) and harmonic structure (scaling up and down) on a comfortable pitch. One of the ‘western’ singers had also experience and training in throat-singing. He was asked also to produce voice production modes and singing styles of both categories. All other recordings on ThS have been carried out during fieldwork in southern Siberia in the years 2000-2002.

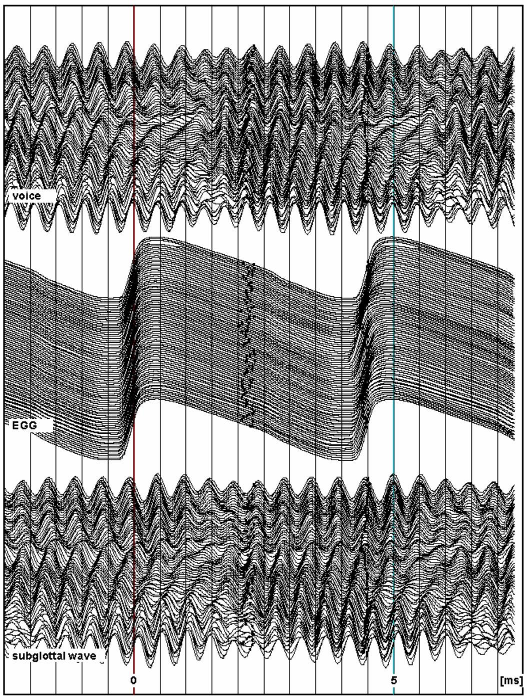

Figure 1: signals from voice (Vx), laryngograph (Lx) and subglottal wave (Sx) – consecutive cycles are ordered along the ordinate axis – EGG serves as basis for synchronizing all on the point of glottal closure (delay time 1.4 ms)

2.2 Recordings and Acoustical Analysis

Recordings were made by connecting a DAT-Recorder (Sony TCD-100) and DCC-Recorder (Philips DCC170) (both devices allow a sampling rate of 44.1 kHz) or alternatively by a 4 channel tape recorder (Tascam

Portastudio 414MKII). For the electroglottographic measurement a portable electro-laryngograph (Kay Laryngograph Ltd.) was used. And a Lavaliere microphone (H&H EM 100 N) served as simple contact 'accelerometer', being placed on the skin of the singer's jugular fossa it could be hold by the singers themselves while singing [15].

Acoustic analysis including fundamental frequency perturbation or periodicity analysis of all three signals, formant and bandwidth analysis has been done by Praat [5] and WaveSurfer [23], cascading of waveforms by Glottal Segmentation [10] and harmonicity (NHR) analysis by the software MDVP & Multi-Speech (Kay Elemetrics Ltd.).

4. RESULTS

4.1 geometry variations of Lx

Observing the different Lx-waveform shapes several deviations from modal (normal) can be stated. So Mod_2 shows a saw type signal with steep closing (contacting) phases and peak skewing. And in $AEF\text{-}VF_1$ mode in some samples a skirt bolging or ramping can be observed [2], while for $AEF\text{-}VF_2$ an incisive 'knee' can be found [4].

4.2 contact signal measurements

The analyses of electro-glottographic values shows already for such a small set of samples some articulate differences in the investigated Closed Quotient (CQ), Closing Quotient (CiQ), Open Quotient (OQ) and Speed Quotient (SQ) (**Table 1**). So SQ seems to be an appropriate dimension to differentiate modal (Mod1 and Mod2) and 'aryepiglottic sphinctered modal' (AEF1 und AEF2). CiQ values support the observed steepness and bandwidth measurements. The for $AEF\text{-}VF_1$ obtained OQ seems unexpectedly high.

	Mod_1	Mod_2	$AEF\text{-}VF_1$	$AEF\text{-}VF_2$
OQ	0,264 (0,094)	0,281 (0,060)	0,304 (0,055)	0,105 (0,074)
CiQ	0,089 (0,017)	0,128 (0,018)	0,051 (0,013)	0,074 (0,037)
CQ	0,125 (0,052)	0,260 (0,080)	0,192 (0,052)	0,167 (0,017)
SQ	0,189 (0,062)	0,349 (0,094)	0,123 (0,052)	0,122 (0,021)
Lx-jitter	0,060 (0,020)	0,082 (0,025)	0,046 (0,011)	0,061 (0,025)
Lx-shimmer	0,047 (0,022)	0,083 (0,015)	0,018 (0,003)	0,013 (0,002)
Vx-jitter	0,037 (0,047)	0,022 (0,009)	0,030 (0,022)	0,040 (0,015)
Vx-shimmer	0,037 (0,020)	0,021 (0,004)	0,042 (0,005)	0,046 (0,007)

Table 1: mean values of EGG-measurements (STD in parentheses) and frequency perturbation of reinforced harmonics and throat singing of two singers

On the other hand the production modes of the lower registers indicate a different interpretation of the vocal fold contact area signal. Already the endoscopic findings lead to the assumption that also ventricular fold contact and aryepiglottic fold contact is represented in the waveforms (**Figure 2**). For low register styles with a pitch range from 45 to 70 Hz typical double rhythm and triple rhythm patters can be found [3] [17]. Though contact cycle measurements haven't been made yet.

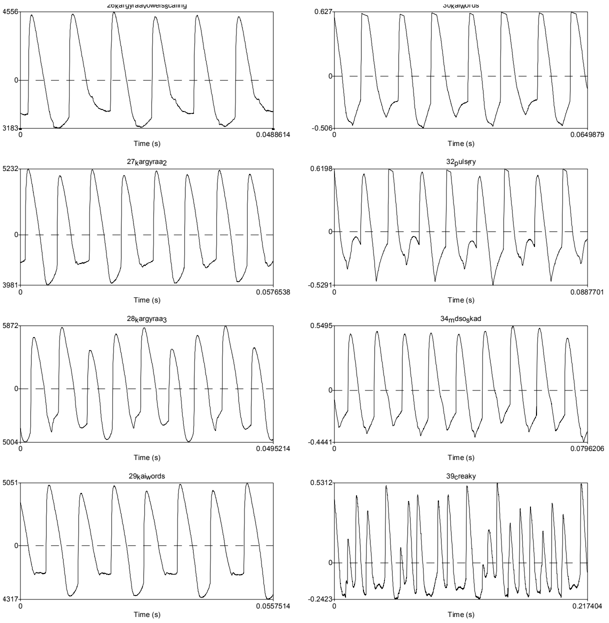

Figure 2: overview on the different findings variety of laryngographic waveforms for the low register (40-75 Hz)

Jitter (periodic jitter), shimmer and NHR measurements were taken on intervalls of quasi sustained passages with an approximate duration of 2,5-3s (Table 1 and 2). The obtained values show for the upper register a rather small straggling, but also without a clear tendency. Lower register values scatter strongly beside vx-jitter. Here 'kargyraa' achieves unexpectedly a rate fewer than 10%. [21]

	pulse (fry)	creaky	VeF-VF Kargyraa	$AEF\text{-}VF_2$ Kargyraa
Lx-jitter	0,149 (0,167)	0,193 (0,238)	0,025 (0,010)	0,012 (0,004)
Lx-shimmer	0,122 (0,099)	0,223 (0,220)	0,039 (0,018)	0,041 (0,043)
Lx-cycle mode	doublets, triplets	single, doublets, triplets	doublets	doublets
Vx-jitter	0,220 (0,122)	0,137 (0,043)	0,062 (0,045)	0,025 (0,020)
Vx-shimmer	0,209 (0,067)	0,342 (0,289)	0,119 (0,043)	0,133 (0,086)

Table 2: average perturbation values for 'low register' phonation modes (Note the high variance)

As for the measured formants and bandwidths of the two often as identical presumed types with a very clear flute-like sound of reinforced harmonics appears also a significant difference in formant enhancement. Especially the bandwidth of the ‘melodic formant’ F2 narrows in ThS exceedingly (**Table 3**) [cf. 13, 8].

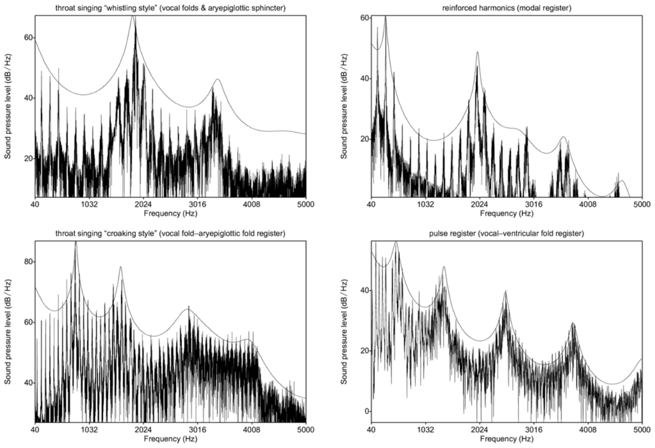

Figure 3: overlaying LPC and FFT spectra of ThS (left) and OtS (right) in upper and lower register

An other significant attribute over all throat singing styles is a F3 (and sometimes F4) ostinato (**Figure 3**). It is clearly associated with a twanging timbre of ThS and serves for the tuvan ThS singers also as a quality characteristic [11]. And it could also stand for the ability of voice penetrance.

Noise-to-harmonic ratio (NHR) seems to be aligned with the degree of constriction (from Mod_1 to AEF-VF_2) in the upper register of ThS (**Table 3**). As for the lower register the ratio appears significantly higher than MDVP treshhold of 0,190.

	Mod_1	Mod_2	AEF-VF_1	AEF-VF_2
F_D	F1	F1	F2, F3, F4	F2, F3
B2	233,8 (262,7) 137,5	167,3 (206,8) 84,3	199,6 (195,6) 121,7	20,95 (12,20) 17,75
NHR	0,073 (0,021)	0,108 (0,018)	0,209 (0,063)	0,473 (0,159)
	Creaky	Pulse (fry)	VeF-VF	AEF-VeF
F_D	F1	F1	F1, F2	F1
B2	99,0 (107,6) 69,7	179,7 (186,2) 106,8	174,3 (174,2) 102,5	42,7 (102,8) 39,9
NHR	0,406 (0,085)	0,729 (0,044)	0,980 (0,127)	0,458 (0,065)

Table 3: bandwidth values (mean, sd, median) and noise-to-harmonic values (mean) of 2 singers

The specific subglottal cavity excitation can be observed in a plot of cascaded waveforms (**Figure 4**). The Sx-wave shows on one hand a synchronization effect in regard to vocal tract shape changes (especially F2) and a dominant F1 in the chest resonator. On the other hand for low register ThS styles also a douple or triple excitation in Sx would be expected – an effect which the material does not support yet.

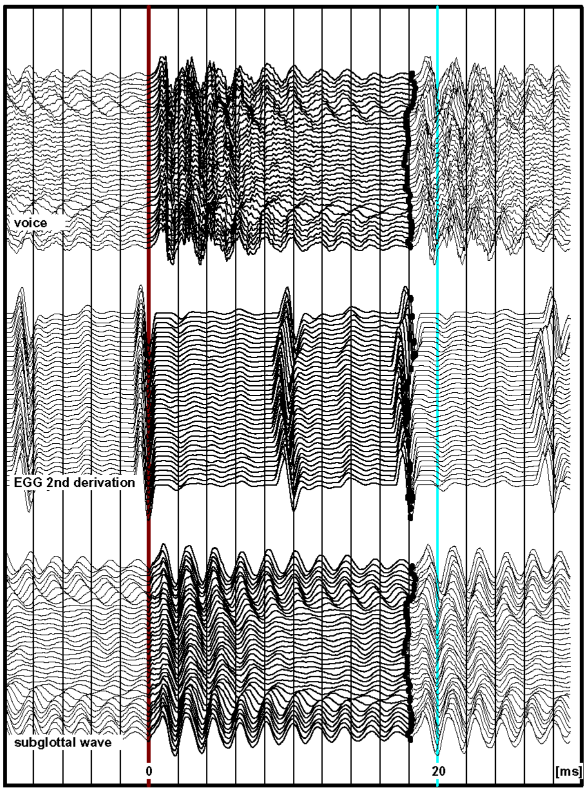

Figure 4: subglottal (delay time 1.4 ms) wave of an /aua/ passage of AEF-VeF phonation mode (style kargyraa)

5. CONCLUSIONS

Bandwidth measurements validate the auditive impression of a sharper enhancement in ThS. The F3/F4 ostinato might be identified with the first small resonator of the epilarynx [19] or better additus laryngis, since all ThS modes share this articulatory adjustment.

On the basis of jitter, shimmer and NHR measures lower register of ThS are appearing more homogeneously which of course could be explained inter alia by a stronger productive use and focus in training of this genre, the usage of pulse register is not very widespread among OtS interpreters.

The dispersion in OQ and CQ values reflects the uncertainty in marking those phases of the contact signal. Due to consisting problems in imaging those vibratory patterns it is unlikely to be solved in the near future. Especially the production modes of all low registers indicate a different interpretation of the vocal fold contact area signal. Overlaying vibratory patterns of different vibrating masses seem to be present in Lx but without noticeable waveform change in Sx. Here a time-based

ISBN 1-876346-48-5 © 2003 UAB

subharmonic analysis could be a better attempt [17]. The examination of the cascaded oscillogram can be linked to spectrographical characteristics in the voice signal, but it offers also a more direct view on adjunctive waveform behaviour. The subglottal pressure wave shows especially time base synchronization effects during the adjustment of the vocal tract and subglottal resonator and displays the excitation of subglottal cavities. Especially observing the articulatory influence seems very promising for further inverstigations [8, 14].

REFERENCES

[1] E.L. Allen and H. Hollien, “A Laminagraphic Study of Pulse (Vocal Fry) Register Phonation,” *Folia phoniatr.*, 25, pp. 241-250, 1973

[2] R.J. Baken and R.F. Orlikoff, *Clinical Measurement of Speech and Voice*, 2nded. San Diego, CA: Singular, 2000.

[3] M. Blomgren, Y. Chen, L. Ng. Manwa and H.R. Gilbert „Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers,” *J. Acoust. Soc. Am.*, 111(2), pp. 1045-1055, 2002.

[4] G. Bloothooft, G. de Krom, S. Jansen and A. van der Heijden, “Electroglottogram recordings during overtone singing”, (Abstract), Voice Conference, Philapdelphia , 1994,

[5] P. Boersma, *PRAAT- Doing Phonetics by Computer*, Software, Vers. 4.0.47, www.praat.org 1998-2003.

[6] Y. Chen, M.P. Robb, H.R. Gilbert “Electroglottographic evaluation of gender and vowel effects during modal and vocal fry phonation,” *J. Speech Hearing Research*, 45, pp. 821-829, 2002.

[7] L.B. Dmitriev, B.P. Chernov, V.T. Maslov, “Functioning of the Voice Mechanism in Double-Source Touvinian Singing,” *Folia Phoniatrica*, 35, pp. 193-197, 1983.

[8] M. Edgerton, D. Bless, S. Thibeault, M. Fagerholm, B. Story “The Acoustic Analysis of Reinforced Harmonics,” Submitted for review to *Journal of New Music Research,* 1999, revised 2002.

[9] L. Fuks, B. Hammarberg, J. Sundberg, “A selfsustained vocal-ventricular phonation mode: acoustistal, aerodynamic glottographic evidences,” *THM-QPSR*, 3/1998, pp. 49-59, 1998.

[10] V. Gall and R. Berg, *Glottal Segmentation – Die Feinstruktur von Stimme und Sprache*, Frankfurt/M.: Woetzel, 1998.

[11] S. Grawunder “Der südsibirische Kehlgesang als Gegenstand phonetischer Untersuchungen,” in Hallesche Schriften zur Sprechwissenschaft und Phonetik, Band 10, E. Stock, Ed., pp.

[12] W. Henke, “Signals from external accelerometer during phonation: Attributes and their internal physic,” *MIT Quarterly Progress Report*, 114, pp. 224-231, 1974

[13] F. Klingholtz, “Overtone Singing: Productive Mechanisms and Acoustic Data,” *J. Acous. Soc. Am.*, 7, pp. 118-122, 1993

[14] G. Matern, C. Neuschaefer-Rube, W. Saus, M. Kob, “Ansatzrohrgestaltung beim Obertonsingen,” in *Aktuelle phoniatrisch-pädaudiologische Aspekte 2000/2001*, Band 8, M. Gross and E. Kruse, Eds., pp. 16-18. Heidelberg: Median, 2001.

[15] K. Neumann, V. Gall, H. Schutte, D.G. Miller „Subglottale Druckverläufe – Vergleich intratrachealer Druckmessungen und simultaner externer Kontaktmikrofonregistrierung,“ in *Aktuelle phoniatrisch-pädaudiologische Aspekte 2000/2001*, Band 8, M. Gross and E. Kruse, Eds., pp. 19-23. Heidelberg: Median, 2001.

[16] T. Levin and M. Edgerton “The Throat Singers of Tuva,” *Scientific America,* 218: 3, pp. 70-77, 1999.

[17] J.G. Svec, H.K. Schutte, and D.G. Miller, "A Subharmonic Vibratory Pattern in Normal Vocal Folds", *Journal of Speech and Hearing Research*, 39, pp. 135-143, 1996.

[18] H.K. Schutte, D.G. Miller, and J.G. Švec, „Measurement of formant frequencies and bandwidths in singing,” *Journal of Voice,* 9, pp. 290-296, 1995.

[19] I.Titze and B. Story, “Acoustic interactions of the voice source with the lower vocal tract,” *J. Acous. Soc. Am.*, 101 (4), pp. 2234-2242. 1997.

[20] H. Schutte and D.G. Miller, “Resonanzspiele der Gesangsstimme in ihren Beziehungen zu supra- und subglottalen Druckverläufen: Konsequenzen für die Stimmbildungstheorie,“ *Folia Phoniatrica*, 40, pp. 65-73, 1988.

[21] M.N. Viera, F.R. McInnes and M.A. Jack “On the influence of laryngeal pathologies on acoustic and electroglottographic jitter measures,” *J. Acoust. Soc. Am.*, 103(5), Pt. 1, pp. 2649-2658, 2002.

[22] H. Smith, K.N. Stevens, and R.S. Tomlinson, “On an unusual mode of chanting by certain Tibetan lamas,” *J. Acoust. Soc. Am.*, 41, pp. 1262-1264, 1967.

[23] WaveSurfer is being developed at the Centre for Speech Technology (CTT) at KTH in Stockholm, Sweden, and is provided as open source, under a BSD style license. www.speech.kth.se/wavesurfer/