

Voice production in death metal singers

Cornelia Eckers^a, Diana Hütz^a, Malte Kob^{b,c}, Peter Murphy^{b,d}, Diana Houben^a and Bernhard Lehnert^b

^aZuyd University, department of speech and language pathology; ^bChair of Phoniatrics and Pedaudiology, RWTH Aachen University; ^cErich-Thienhaus-Institute, University of Music Detmold; ^dDepartment of Electronic and Computer Engineering, University of Limerick (IRL)

Abstract

Singers in “death metal” style, a deep and harsh, i.e. kind of grunting, form of singing, often use a unique form of voice production called “growling”. The underlying mechanisms of voice production as well as possible potentially harmful influences on the singer’s voice have not been intensively studied. Supraglottal laryngeal constriction appears to play a major role in metal singing. In this study we present an examination method using a questionnaire including the voice handicap index as well as a phonetogram and combined simultaneous electroglottography and acoustic analysis. Furthermore a laryngoscopy with a rigid optic, for singing vowels, and a transnasal flexible laryngoscopy, for growling, was carried out to investigate in particular the production of death metal singing in seven subjects. None of the subjects complained about persistent voice problems and none had signs of secondary organic voice disorders like vocal fold nodules or inflammation. However, most showed minor irregularities on stroboscopic examination. Further findings reveal that at least two forms of supraglottal laryngeal constriction are used to obtain the typical death metal sound which differs in use of vocal folds and ventricular folds as well as in supraglottal adduction effort. The study shall help to answer the question whether death metal singing can be performed in a manner such that the risk of voice damage can be reduced or avoided.

Introduction

Across singing styles different voice production mechanisms are used. Some styles employ overtone singing with e.g. a whistle-like high overtone like “drone” [10]. Others may use undertone singing. This form of singing is produced with low frequencies. Undertone singing includes the application of the ventricular folds inside the larynx [11]. A similar form of singing is often used in metal styles such as “death metal”.

In metal-jargon this type of singing is called “growling”. Growling is originally mentioned as low-pitched sounds uttered by e.g. dogs and other animals. It has been observed in other singing styles as well, such as the jazz singing style of Louis Armstrong and Cab Calloway [3], [4].

With the increasing number of metal-singers the question of stress of the voice production mechanism and of resultant voice disorders has gained more important and more relevance. In some metal-forums and also in some articles voice disorders of metal-singers are mentioned [9]. Because of this knowledge, metal-singers are potential patients for speech therapists.

Considering the physiology, Sakakibara et al. (2004) found the aryepiglottic folds vibrating in addition to the vocal cords (at a rate of 2:1) for two singers producing growl voice [10].

However, there is not enough research in death metal singing respectively growling of death metal singers to give some voice trainings or therapy to patients in metal-singing. Out of this lack of research the following study resulted. Its purpose is getting some more basic knowledge over the physiological production of the growling voice, the acoustic aspects of growling and the symptoms and factors of stress of death metal singing.

Method

Laryngeal movements in growling were observed directly and indirectly by recording of endoscopic digital images, EGG (electroglottography) waveforms and sound waveforms. The normal and the growling voice were tested with a phonetogram to get more information about the capacity of the growling voice and symptoms of voice disorders. Furthermore the stress of voice and possible voice disorders are examined within a structured interview as well as with the voice handicap index. The investigations took place at the clinic of Phoniatrics and Pedaudiology at the University Hospital Aachen.

Subjects/Participants

Seven death metal singers participated in this research. Their experience in death metal singing resp. growling varied from 1½ to 19 years. Because of the voice is reported to be optimum between the age of 20 and 40 years, all participants were in this range (range = 22 – 34) [5]. Prior to performing the research tasks each participant signed a privacy policy and a declaration of consent.

Data collection

The participants were asked to sustain the vowel /a/ in a continuous manner going from modal register, to growling and back to modal register in almost every investigation. It was required that each participant sustained the vowel for a minimum of 3 seconds at a comfortable loudness level. During the acoustic measurements and phonetogram analysis the distance between the subject’s mouth and the microphone was 30 cm. The EGG and acoustic measurements were recorded simultaneously. Two electrodes were positioned right and left from the larynx while the subjects were singing from modal register into growling and back into modal register. The data from the acoustic and EGG measurements were saved with “Audacity”, free software to edit audio tracks.

The investigation of the vocal performance was first done regular in modal register with the Software “Phoneto” [8]. Then the subjects had to produce the growling style. In order to get some valuable results, they had to try singing in the deepest and the highest possible tones in growling, in addition to the given tones in the phonetogram.

The endoscopy was carried out by a doctor of the clinic of Phoniatics and Pedaudiology, University Hospital Aachen (the last author of this paper). The first endoscopic investigation was carried out with a rigid optic. Here the subjects were asked to sustain the vowel /i/ to get some information of the organic healthiness of the larynx and the vocal folds. To get some more knowledge about the physiological aspects of growling, there was also done a transnasal flexible laryngoscopy. In this way the subjects were able to produce the fluent passage from modal register into growling and back into modal register without urge to gag. Here they had to sustain the vowel /a/.

Moreover the subjects had to answer some open and closed questions related to the growling voice, the production, complaints, capacity of voice etc. in the form of a structured interview. All interviews were saved with Audacity.

Finally the participants had to fill in the Voice Handicap Index (VHI). This was supposed to bring some information over possible existing voice disorders of the subjects [7].

Data analysis and results

Laryngoscopy

The digital images due to this investigation were interpreted by an image-by-image-observation of the laryngeal structures with the software “VirtualDub” and filling this in a rating sheet for the qualitative analysis of the vibration of the vocal folds by Böhme and Gross (2001) [2].

The first endoscopic investigation with a rigid optic could be done by five of seven subjects. They were physically healthy. The transnasal flexible laryngoscopy could be done by six subjects. During the laryngoscopy two different forms of production of growling voice were identified. One form of production used the vibrating of the ventricular folds with an antero-posterior constriction of the supralaryngeal tract. In another production of growling the aryepiglottic folds vibrated.

Analysis of the Electroglottographic Signal (EGG)

The results of this investigation were qualitatively analyzed by watching precisely at the blow up EGG-waveforms of each subject.

The EGG waveforms for growl differed markedly from those produced for modal register. Two vibratory regimes were observed during growl production: a slow variation occurring at a sub-multiple of the vocal fold fundamental frequency and the vocal fold fundamental frequency.

The waveform for one of the participants had 2 vocal fold oscillations to one slow oscillation (2:1, Fig.1). Four of the participants had 3 vocal fold oscillations to one slow oscillation. The EGG-waveform of a further participant had 4 vocal fold oscillation to one slow oscillation and another participant produced an aperiodic waveform suggesting chaotic vibration at level of the larynx. Because of this, the waveforms were somewhat different in each case. Some example segments are shown in Fig.1.

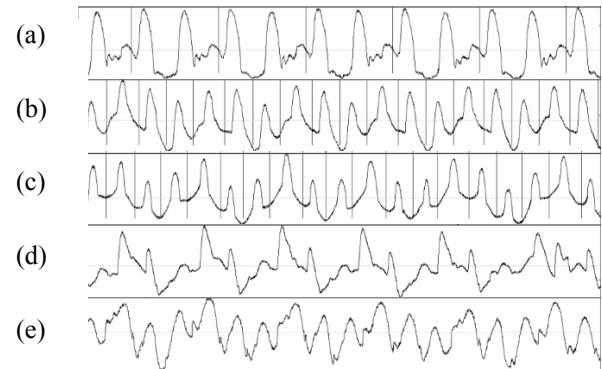


Figure 1: (a) EGG-waveform with a ratio of 2:1. (b-e) Ratio 3:1 in four different participants.

Acoustic Analysis

Analysis with the hoarseness diagram by Michaelis (1999)

The acoustic recordings of /a/ in modal/growling/modal phonation were analysed using the hoarseness diagram [6]. With this, parameters of hoarseness and noise could be calculated. There was a minimum of three audio tracks grabbed to get enough values to analyze. Each audio track was split into three parts: Two parts of modal register (first and third part) and one part of growling (middle). Onset and offset were deleted from each part. After this, reference values for the parameters PPQ (Pitch Perturbation Quotient/jitter), EPQ (Energy Perturbation Quotient/shimmer), GNE (Glottal to Noise Excitation Ratio), irregularity-component

$$IC = 5 + \frac{1}{\sqrt{3}} \left(\frac{\log J3 + 0,374}{0,645} + \frac{\log S15 - 0,757}{0,368} + \frac{\log(1 - MWC) + 1,614}{0,574} \right) \quad (1)$$

and noise-component

$$NC = 1,5 + \frac{GNE3 - 0,695}{2,42} \quad (2)$$

were obtained by analyzing the modal register parts with the hoarseness diagram. The values were compared with the values averaged by analyzing the growling part of each subject with the hoarseness diagram.

The parameters obtained from the growling parts were increased (range: 4,22 to 8,90), compared to the reference values (range: 1,78 to 3,16).

Especially the parameter EPQ was strikingly increased (average: 27.26), compared with reference values (average: 3.93) as seen in the following table:

	EPQ [%]		PPQ [%]	
	growing	reference	reference	reference
S 1	35,14 ± 6,69	13,93± 12,96	15,16 ± 2,76	1,17 ± 1,78
S 2	22,89 ± 8,33	2,73 ± 0,71	12,35 ± 4,47	0,14 ± 0,03
S 3	37,88 ±11,44	3,96 ± 1,89	12,47 ± 8,02	0,16 ± 0,16
S 4	17,64 ±13,51	1,49 ± 0,37	5,36 ± 4,70	0,10 ± 0,05
S 5	6,58 ± 4,14	2,18 ± 0,41	0,24 ± 0,66	0,09 ± 0,04
S 6	39,38 ± 5,28	1,27 ± 0,27	7,39 ± 1,65	0,04 ± 0,00
S 7	31,29 ± 9,49	1,97 ± 0,73	9,84 ± 2,67	0,07 ± 0,03

Table 1: Averaged reference values with standard deviation and values of growing with standard deviation for EPQ and PPQ of each subject (S).

Analysis with Praat

The averaged three audio parts (modal register – growing – modal register) were taken to obtain some values about unvoiced frames and harmonics-to-noise-ratio (HNR).

The results of the analysis of the growing parts of each subject with the software *Praat* were differently. In some audio tracks the growing part had 0% to 17% unvoiced frames and in other ones the unvoiced frames varied from 68% to 77%. Two examples are shown in the following figure:

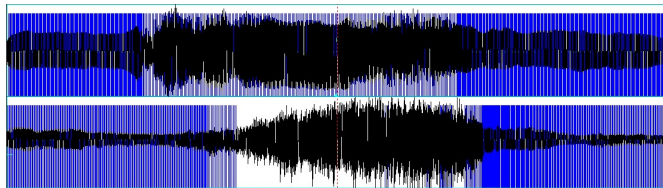


Figure 2: On top: Voice signal with 0% unvoiced frames in the growing part. At the bottom: Voice signal with 77% unvoiced frames in the growing part. Blue vertical lines = voiced frames.

Six of the seven participants displayed HNR distinctly below the standard value of 20dB. The lower the HNR, the more the hoarseness was pronounced.

Phonetogram

The measurements are analyzed in comparison with standard values by Nawka and Wirth (2008) [8]. The phonetogram was done with seven participants in modal register and growing. In modal register there were various results. Only two participants achieved normal values in dynamic range. One showed a light restriction in forte. Four participants exhibited deficits in piano, exemplarily shown in figure 3, and two participants in forte and piano.

The results of the growing-phonetogram were different from the standard values of modal register. Every participant was as loud as, or louder than, his modal register-phonetogram. Furthermore, in growing every participant reached deeper than, or as deep as, their performances in modal register. But they all did not reach the standard values of lowest volume, highest tone, dynamic range and pitch range, as shown in figure 3 as an example.

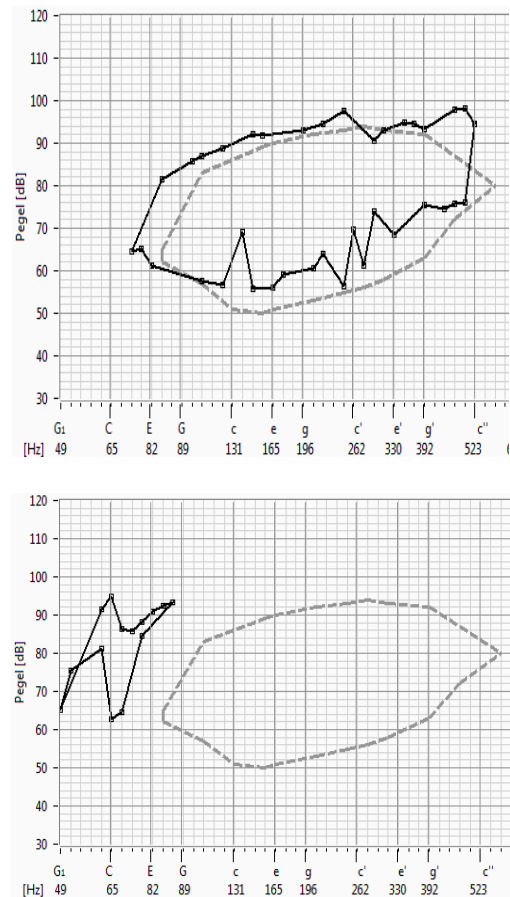


Figure 3: Left: Phonetogram of modal register (black solid line) by one subject with reference male standard register (grey dashed line). Right: Phonetogram of growing voice (black solid line).

Voice Handicap Index (VHI)

The questionnaire of the VHI was given to each subject to fill in their answers. They were compared with standard values [7]. There was only one subject on the lower bound of a light voice disorder. Anybody else did not have a voice disorder by analyzing the VHI.

Interview

Almost each subject stated to feel more pressure in the larynx and/or chest. Further on they said that they need more air to produce the growing. Otherwise the descriptions were various. The capacity of voice amounts to ca. 2 hours. More stress of voice in growing lead to various strength of hoarseness or scratch of throat. The regeneration of the voice normally eventuates over night.

Complaints of voice were scarcely described by the subjects. Only after overloading in growing they stated the described problems.

Discussions and conclusions

There are two major forms of producing growing: one using the ventricular folds and another using the aryepiglottic folds. The possibility that during the vibration of the aryepiglottic folds, the ventricular folds vibrated as well could not be ruled out. This could not be seen during the

laryngoscopic investigations. Contrariwise to this study, Sakakibara et al. (2004) excluded a vibration of the ventricular folds while growling [10].

With the aid of the EGG there could be concluded that in every production of growling the vocal folds are vibrating, too [1]. Related to the ventricular folds or aryepiglottic folds they vibrate 2:1, 3:1 or 4:1. In contrast to this, Sakakibara et al. (2004) found the vocal folds vibrating exclusively in a ratio of 2:1 to the aryepiglottic folds during growling [10].

The acoustic measurements, analyzed with the hoarseness diagram, resulted in high values for Jitter and Shimmer. This resulted in a high irregularity component. The deeper values of GNE resulted in a higher value of the noise component. These values differ from values of the modal register and are comparable to values measured by Michaelis (1999) [6] in subjects with a compensatory ventricular folds phonation or compensatory aryepiglottic phonation.

The analysis with Praat showed different results for unvoiced frames during growling. Some audio tracks of the subjects had many unvoiced frames and others just a few. So it differs from 0% to 77%.

The EGG-waveforms of one participant, who reached 0% unvoiced frames, showed a relation of two vocal fold oscillations to one oscillation of another supralaryngeal structure. Moreover this participant (S5) had, in contrast to the other ones, low values concerning the parameters of the hoarseness diagram. Because of the analysis of his laryngoscopic digital images, it could be imagined, that the phonation of the vocal folds, together with the ventricular folds and without the aryepiglottic parts, in the mentioned ratio leads to results close to standard values.

A simultaneous investigation of laryngoscopy, EGG and acoustic measurements was not placed in this study but may result in better and more comparable values to e.g. prove the assumption described above.

Of the phonetograms arises a decrease of the fundamental frequency in growling and restricted values for lowest volume, highest tone, dynamic range and pitch range. Seen in retrospect, the phonetogram did not seem to be the best instrument to capture the capacity of the growling voice. Several tones could not be recorded by the phonetogram.

The Voice Handicap Index did not show some distinctive voice disorders. Even with the interview this could not be concluded. But there were mentioned clear factors of stress in producing growling for a longer period of time. More than ca. 2 hours of growl production lead to various occurrences of hoarseness or scratch of throat. The VHI was afterwards not the best method to investigate voice disorders. A screening or logopedic investigation might have been more informative.

This study produces more knowledge about the physiological and acoustic aspects of the growling voice. Some more research with multichannel EGG-measurements, ideally simultaneously to acoustic and endoscopic investigations and in a larger sample size, could extend the findings of this study.

Acknowledgements

The authors would like to thank each subject for his participation and the first two authors would like to thank their families and friends for their support.

Dr. Peter Murphy was funded during this period through a DAAD research grant (October, 2008), followed by an EU Erasmus teaching/staff mobility grant (November-December, 2008).

References

- [1] Blomgren, M., Chen, Y., Ng, M.-L., Gilbert, H.-R. (1998). *Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers*. J. Acoust. Soc. Am. 103 (5), Pt. 1
- [2] Böhme, G. (2003) 4. Auflage. *Sprach-, Sprech-, Stimm- und Schluckstörungen: Bd. 1 Klinik*. München, Jena: Urban & Fischer Verlag
- [3] Catford, J. C. (1977). *Fundamental Problems in Phonetics*, Edinburgh Univ. Press.
- [4] Esling, J. H. (1996). Pharyngeal consonants and the aryepiglottic sphincter, *J. International Phonetics Association*, 26(2):65–88
- [5] Hammer, S. (2005) 2. Auflage. *Stimmtherapie mit Erwachsenen. Was Stimmtherapeuten wissen müssen*. Berlin: Springer
- [6] Michaelis, D. (1999). *Das Göttinger Heiserkeits-Diagramm -Entwicklung und Prüfung eines akustischen Verfahrens zur objektiven Stimmgütebeurteilung pathologischer Stimmen*. Dissertation zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultäten der Georg-August-Universität zu Göttingen.
- [7] Nawka, T., Wiesmann, U., Gonnermann, U. (2003). *Validierung des Voice Handicap Index (VHI) in der deutschen Fassung. HNO 51: 921.929*
- [8] Nawka, T., Wirth G. (2008) 5. Auflage. *Stimmstörungen: Für Ärzte, Logopäden, Sprachheilpädagogen und Sprechwissenschaftler*. Köln: Deutscher Ärzte-Verlag
- [9] Nederlands Dagblad (2007). 'Grunten' sloopt de stem [online]. Available: http://www.nd.nl/Document.aspx?document=nd_artikel&id=95336 [2008, 12. Februar]
- [10] Sakakibara, K., Fuks, L., Imagawa, H., Tayama, N. (2004). *Growl Voice in Ethnic and Pop Styles*. Nara, Japan: Proc. ISMA
- [11] Seidner, W., Wendler, J. (2004) 4. Auflage. *Die Sängerstimme – phoniatische Grundlagen der Gesangsausbildung*. Berling: Henschel