ORIGINAL ARTICLE

Taylor & Francis healthsciences

Vocal fold vibration and voice source aperiodicity in 'dist' tones: a study of a timbral ornament in rock singing

D. Zangger Borch[1], J. Sundberg[2], P.-Å. Lindestad[3] and M. Thalén[1]

From the [1]University College of Music Education in Stockholm (SMI), Sweden, [2]KTH Voice Research Centre, Department of Speech Music Hearing, KTH, Sweden, and [3]Department of Logopedics and Phoniatrics, Karolinska Institutet, Huddinge University Hospital, Sweden

Received 17 October 2003. Accepted 17 May 2004.

Logoped Phoniatr Vocol

The acoustic characteristics of so-called 'dist' tones, commonly used in singing rock music, are analyzed in a case study. In an initial experiment a professional rock singer produced examples of 'dist' tones. The tones were found to contain aperiodicity, SPL at 0.3 m varied between 90 and 96 dB, and subglottal pressure varied in the range of 20–43 cm H_2O, a doubling yielding, on average, an SPL increase of 2.3 dB. In a second experiment, the associated vocal fold vibration patterns were recorded by digital high-speed imaging of the same singer. Inverse filtering of the simultaneously recorded audio signal showed that the aperiodicity was caused by a low frequency modulation of the flow glottogram pulse amplitude. This modulation was produced by an aperiodic or periodic vibration of the supraglottic mucosa. This vibration reduced the pulse amplitude by obstructing the airway for some of the pulses produced by the apparently periodically vibrating vocal folds. The supraglottic mucosa vibration can be assumed to be driven by the high airflow produced by the elevated subglottal pressure.

Johan Sundberg, KTH Voice Research Centre, Department of Speech Music Hearing, KTH, SE-100 44 Stockholm, Sweden. Tel.: +468 790 7873. Fax: +468 790 7854. E-mail: pjohan@speech.kth.se

INTRODUCTION

Vocal technique in singing differs greatly depending on the musical style (1, 2). Much research has been spent on the classical style, while non-classical styles, which would be more commonly used, have been analyzed more rarely. Yet, they merit attention, particularly as the voice use in many of the non-classical styles is sometimes considered detrimental to vocal health. To find out the truth of this view, facts about what characterizes the voice use in different non-classical styles of singing is needed.

The present investigation focuses on the so-called 'dist' singing, a particular timbral ornament frequently used in rock music styles, such as heavy rock and heavy metal as performed by, for example, Whitesnake and AC/DC (3, 43). A typical example of a 'dist' tone can be found between 7.8 s and 8.2 s from the beginning of Whitesnake's recording of 'Crying in the rain' (Coverdale, 1994). Although the term 'dist' may appear as somewhat colloquial, we will use it here, since it is a widely accepted term used among musicians who perform in these styles. The question we ask is what are the acoustic and physiological characteristics of this type of voice use? The answer was found by applying a previously tested method that combined high-speed imaging with inverse filtering, thus illustrating the relationship between vocal fold vibration and transglottal airflow (5).

METHOD

Two experiments were carried out. In experiment 1, a professional singer soloist (co-author DZB) performed examples from his concert repertoire while audio, flow and oral pressure were recorded. The audio was picked up by a TCM 110 TIEPIN microphone, and the flow by a Glottal Enterprises flow mask (Fig. 1a). The lyrics of the songs were replaced by the syllable [pae], thus allowing estimation of subglottal pressure as the oral pressure during the p-occlusions (see, e.g., (6)). Henceforth, this oral pressure will be referred to as the subglottal pressure. The flow and oral pressure signals were stored on separate channels of a TEAC PCM signal recorder.

 ISSN 1401-5439

DOI: 10.1080/14015430410016073

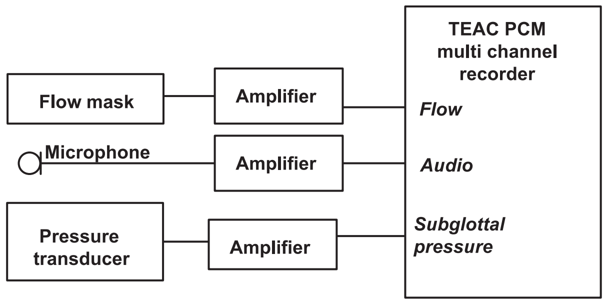

Fig. 1a. Setup at experiment 1, where a flow mask was used to record airflow, a pressure transducer for oral pressure and a microphone for the audio signal.

In experiment 2, the aim was to examine the vibratory and the phonatory mechanisms synchronously in the same singer's production of 'dist' tones. The setup is shown in Fig. 1b. In essence the experiment combined an audio and a high-speed imaging recording of the same singer when sustaining examples of 'dist' tones. The high-speed signal was recorded from a camera attached to a flexible endoscope and a 300 W xenon light source. The signal was stored on a PC provided with an image processor attached to a video recorder. The camera system records up to 4 s at a rate of 1904 images/s. The audio signal was captured by means of a head-mounted TCM 110 TIEPIN microphone and stored on a TEAC PCM signal recorder and also directly into sound files on a PC.

A custom made program (7) was used for the analysis. The program includes kymography as an option showing how the image of a thin slice of the glottis varies over time (8). Thus, in a kymogram time runs from left to right, while up and down correspond to left and right in the glottis. The method offers a valuable complement to the frame-by-frame high-speed imaging as any slice across the glottis can be selected and examined off-line (9). The program also allows Fourier analysis of the vibrations in an image

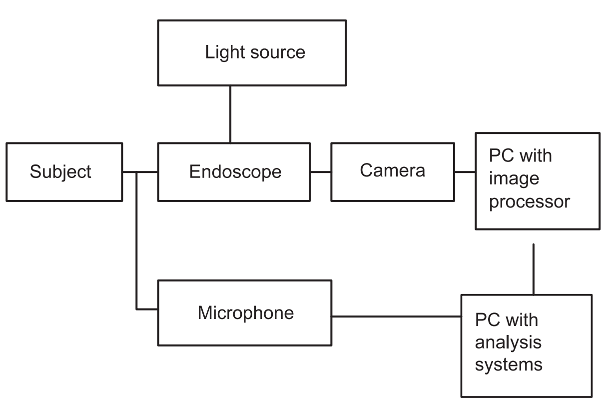

Fig. 1b. Setup at high-speed recording.

sequence, such that the vibration frequency of the various parts of the image can be determined (7).

The audio signal was analyzed by means of a custom made inverse filtering program (DECAP, Granqvist). It displays the waveform and the spectrum before and after the inverse filtering. The frequencies and bandwidths of the filters are set by hand. The Decap program was used also for the analysis of the flow recordings from the first experiment.

RESULTS

A rendering of the song 'Crying in the rain', composed by David Coverdale (4), was selected for analysis from experiment 1. The song has an ambitus of one octave, A#3–A#4 (233 Hz–466 Hz). Fig. 2 shows SPL versus log of subglottal pressure for the tones in this song. By and large, SPL increased with log of subglottal pressure and also with F0, as expected. On average, a doubling of subglottal pressure yielded a 2.3 dB increase of SPL. Similar values have been found in country singers (10), while considerably greater values are typically observed in classically trained singers (11–13).

The 'dist' tones seemed characterized by aperiodicity and/or subharmonics. This is illustrated in the spectrogram shown in Fig. 3. There is considerable sound energy between the harmonic partials of the tone appearing at $3.4 \leq t \leq 4.1$. Subharmonics can be seen during short instances at $t \approx 2.95$, $t \approx 3.2$ and $t \approx 4.2$ s.

As 'dist' examples were recorded on two occasions, the glottograms from both occasions were compared. Such a comparison is shown in Fig. 4. The glottograms obtained from the flow recordings showed a smoother waveform than those obtained from the high-speed session. This was due to the limited frequency range of the flow transducer. A main common characteristic was varying pulse amplitude.

A detailed analysis of flow glottogram properties from the two recording sessions revealed that not only the pulse amplitude varied in an apparently random fashion, but also the closed quotient, that is, the portion of the pitch period during which the glottis is closed (see Fig. 5). The variations of the closed quotient were more pronounced in the high-speed session, while the pulse amplitude variation was similar. This shows that the 'dist' tones produced at the two recording sessions were similar in the sense that both showed a variation of the closed quotient.

An interesting question now is what vibrational characteristics cause this variation in pulse amplitude. Fig. 6 shows two kymograms from the high-speed recording of a 'dist' tone, one from the posterior part of the larynx (Fig. 6a) and one from the middle part

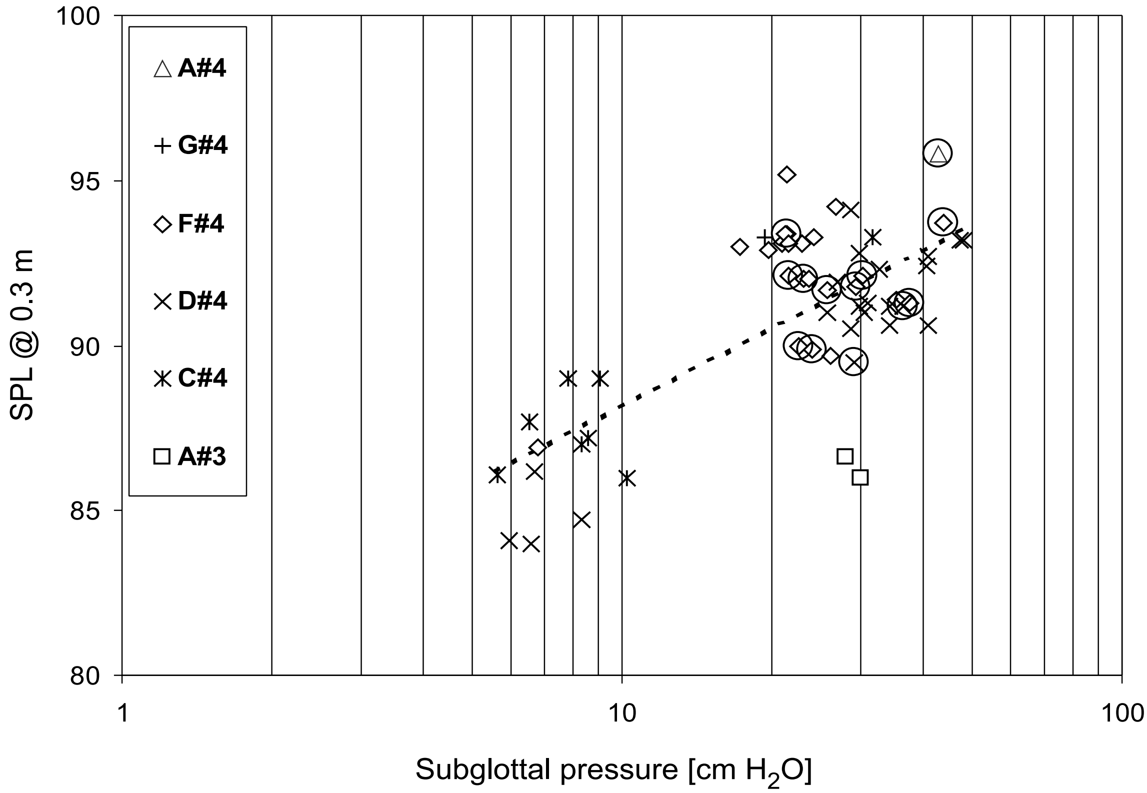

Fig. 2. Subglottal pressure and SPL at 0.3 m recorded in experiment 1. Circled symbols refer to 'dist' tones. The dashed line shows the best linear fit of the data points, characterized by a slope of 2.3 dB per doubling of subglottal pressure.

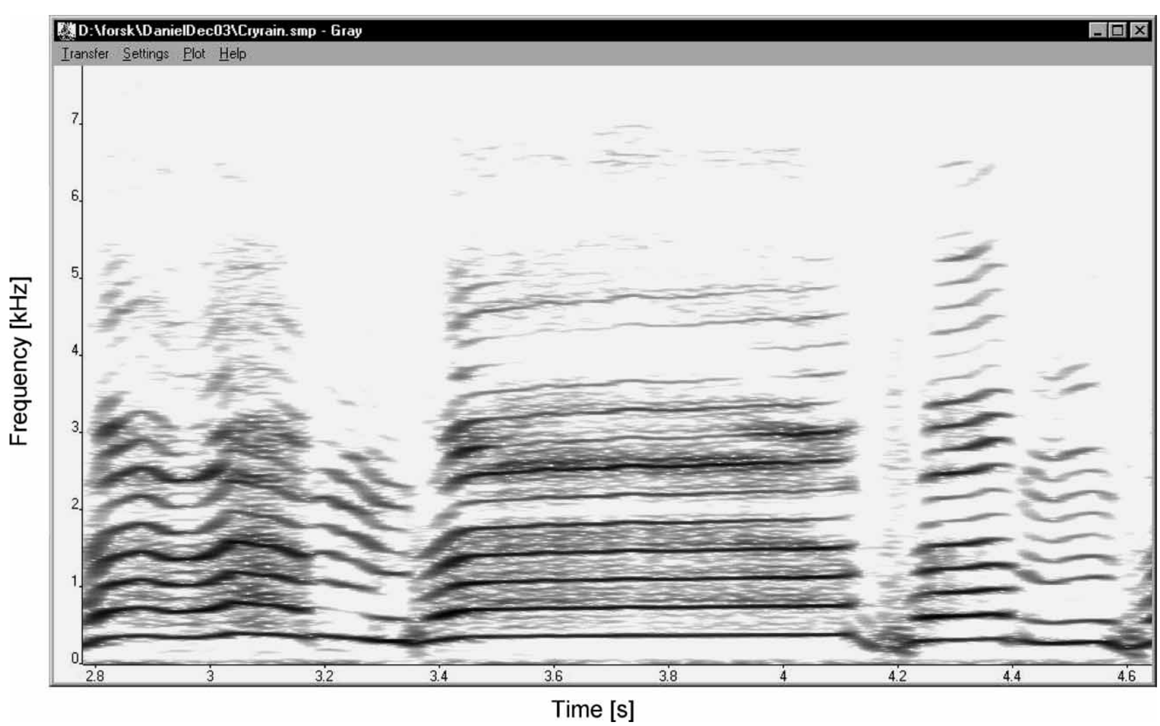

Fig. 3. Spectrogram of a section from the tune 'Crying in the rain' by David Coverdale as performed on the vowel [ae] by the rock singer subject.

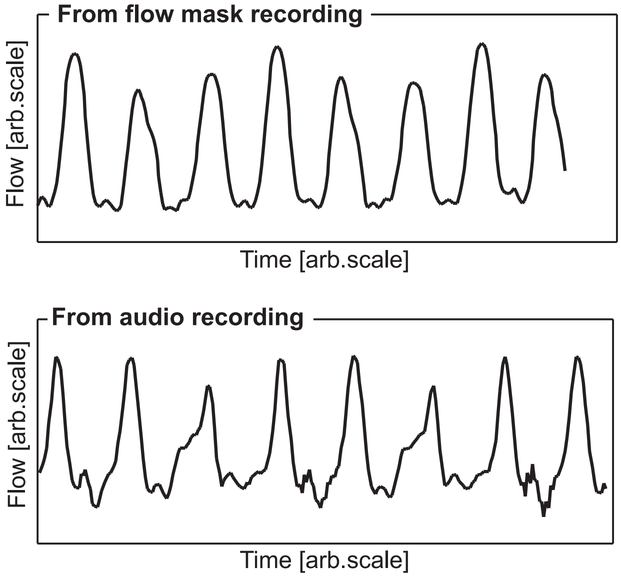

Fig. 4. Examples of flow glottograms of ‘dist’ tones derived by inverse filtering of the flow signal recorded in experiment 1 (upper curve, F0 = 450 Hz) and by inverse filtering the audio signal recorded in experiment 2 (lower curve, F0 = 500 Hz). The time axes represent about 16 ms.

(Fig. 6b). The figure reveals that there are two vibrations: one at the glottis, evident from Fig. 6a; and one, at a lower frequency, in the supraglottal mucosa. Although not shown in the kymogram, a frame-by-frame examination of the movie revealed that the slower vibration included the ventricular folds, and to some extent also the ary-epiglottic folds and the anterior part of the mucosa covering the arytenoid structures. The supraglottic mucosa is seen as the bright, tooth-like pattern, particularly clearly in Fig. 6b. The apparently periodic opening and closing of the glottis can be more clearly observed as an alternation between black and slightly brighter vertical stripes in Fig. 6a. The inverse filtered audio signal (Fig. 6c) shows pulse amplitude variation according to a regular 2+1+2+1+2+1 pattern. Each low amplitude pulse is preceded and followed by a pair of pulses with greater and similar amplitudes. A detailed examination reveals that the low amplitude pulses were synchronous with the approximation of the supraglottic mucosa (see dashed line in Fig. 6b and c). This suggests that the reduction of the pulse amplitude was the result of a narrowing of the supraglottal airway caused by the approximation of the supraglottal mucosa.

A Fourier analysis of the variation of brightness in the kymogram showed two periodicities, one at 160 Hz and the other at 480 Hz, approximately, the latter corresponding to F0 (Fig. 7). These frequencies are harmonically related, since $3 \times 160 = 480$. The audio recording of this ‘dist’ tone also contained these partials, the timbre sounding somewhat like a duodecim interval. This finding indicates that the supraglottic mucosa vibrated at 160 Hz while the vocal fold vibration frequency was 480 Hz. This suggests that in this case there was some coupling between the vocal fold vibration and the vibration of the supraglottic mucosa.

DISCUSSION

In the ‘dist’ tone analyzed in detail here, the supraglottic mucosa was vibrating at a third of the vocal fold vibration frequency. Such a harmonic relationship between vibration frequencies did not seem to characterize all ‘dist’ tones recorded during the first recording session. Rather, a random component was often superimposed on a periodic F0 in many cases.

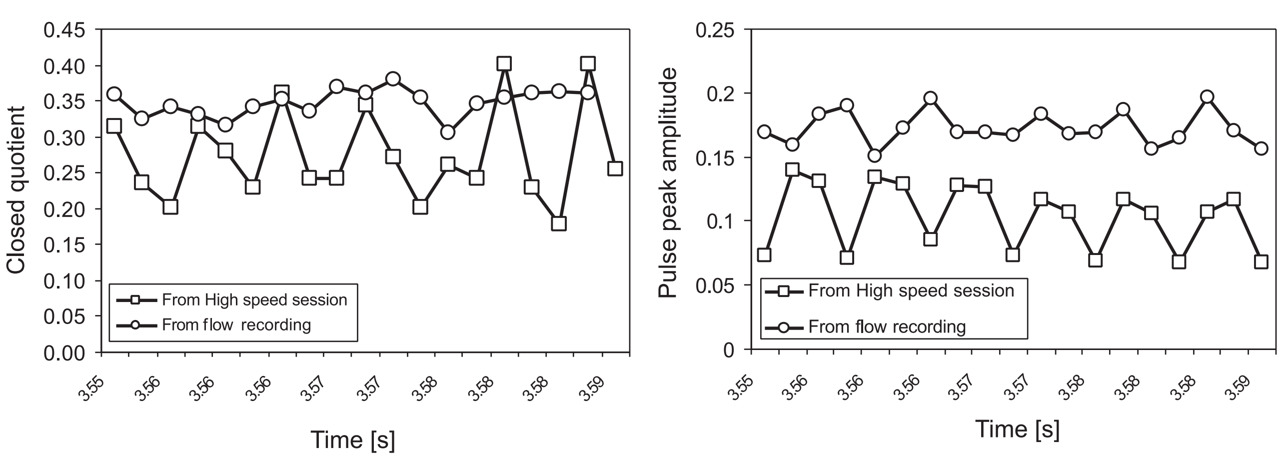

Fig. 5. Closed quotient and pulse amplitude variation in ‘dist’ tones from experiments 1 and 2 (high-speed session and flow recording, respectively).

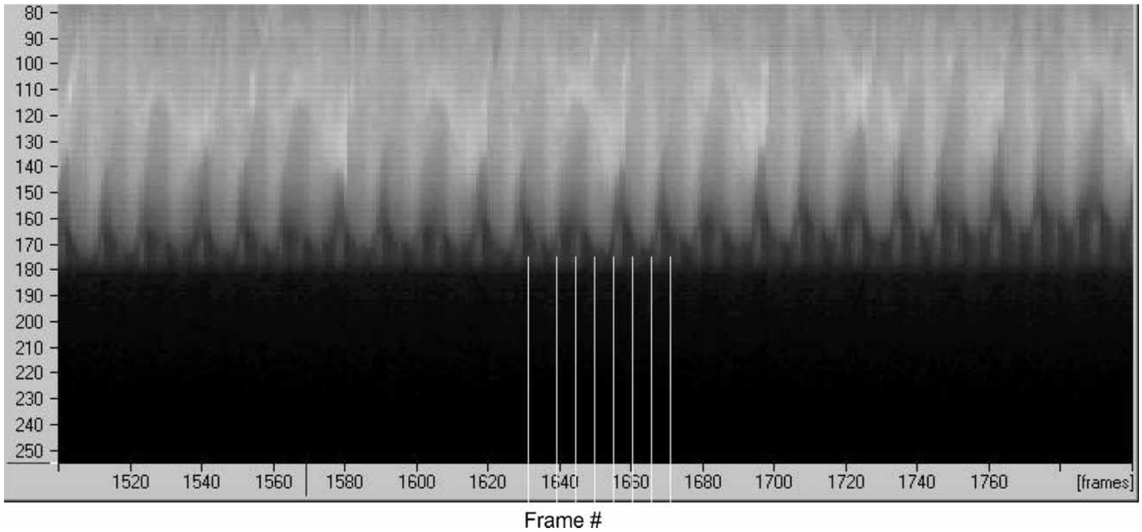

Fig. 6a. Kymogram of the posterior part of the glottis during the production of a 'dist' tone, showing the vocal fold vibrations during frames 1500–1800. The white lines show adjacent moments of glottal opening. Only one of the ventricular folds can be seen.

This suggests that 'dist' tones are produced with a vibration of the supraglottic mucosa that may be aperiodic or periodic.

The vibratory pattern in 'dist' tones shows interesting similarities with that observed during so-called Mongolian throat singing (14, 15). In both cases vibrations occur in the supraglottic mucosa. In the throat singing, however, these vibrations always seem harmonically related to the vibrations of the vocal folds, while in 'dist' singing this does not necessarily apply.

The hygienic aspects of these results are certainly interesting. First it should be recalled that 'dist' is a timbral ornament rather than a phonatory baseline condition, used primarily to add expressivity to high and loud tones. Second, according to reports from many professional pop- and rock singers, 'dist' singing is indeed taxing to the voice and voice rest is considered recommendable after a concert including an abundance of 'dist' tones. On the other hand, this may not necessarily apply to all singers. Some singers appear capable of continuing with 'dist' singing during

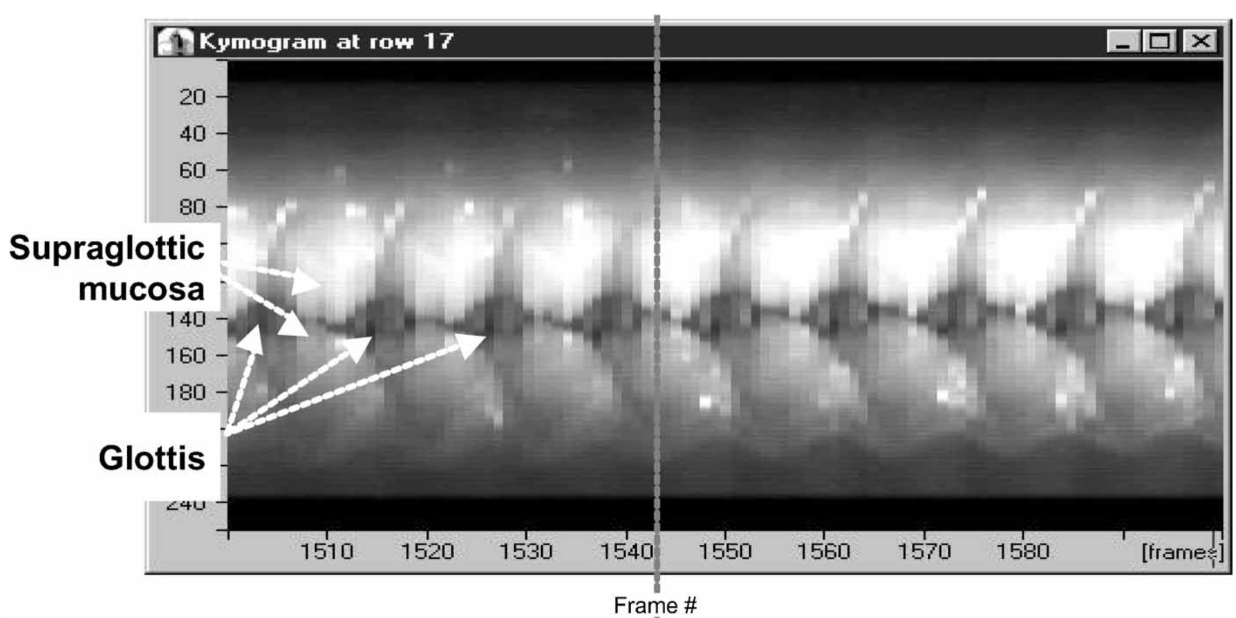

Fig. 6b. Kymogram of the middle part of the glottis during frames 1500–1600 of the same tone as shown in Fig. 6a. Note the vibrations in the supraglottic mucosa.

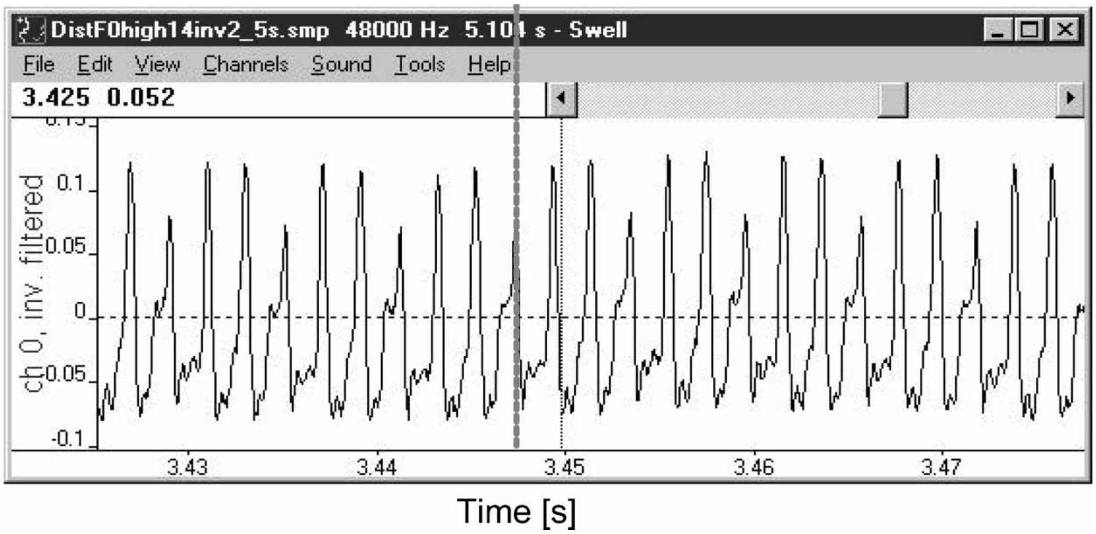

Fig. 6c. Synchronous, inverse filtered flow signal for the same time interval as in the kymogram in Fig. 6b.

many concerts per week, year after year, without interleaved periods of voice rest. This suggests that the vocal technique used to produce 'dist' tones may vary from one individual to the other. This does not seem unlikely. If the supraglottic mucosa is used to superimpose a secondary variation on a periodic glottal pulse train, the periodicity of this superposition should depend on the vibrating mass and the motility of the mucosa covering these structures. These properties are likely to show inter-individual variation. It is also possible that the periodic vocal fold vibrations are an essential part of the 'dist' tone technique from the point of view of vocal hygiene.

Our investigation did not aim at identifying the mechanism exciting the supraglottic mucosa in 'dist' tones. It is however tempting to speculate that it is the air-stream, which brings these structures to vibration. A condition would then be the magnitude of this air-stream. The high subglottal pressures associated with 'dist' tones would cause a forceful air-stream which, combined with an approximation of the arytenoids, may bring them to vibration.

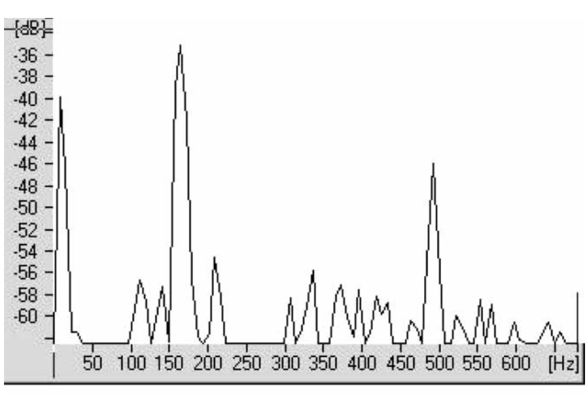

Fig. 7. Fourier analysis of the brightness variations in the kymogram in Fig. 6b, showing vibration frequencies at 160 Hz and 480 Hz, approximately.

CONCLUSIONS

'Dist' tones are loud tones, produced with relatively high subglottal pressures, in the range of 20–45 cm H_2O. They seem characterized by a periodic vocal fold vibration combined with periodic or aperiodic vibration of supraglottal mucosa. The latter obstructs the free passage of the glottal air pulses and imposes variation of pulse amplitude and closed quotient. Professional artists generally recommend vocal rest after concerts in which 'dist' tones have been used extensively.

ACKNOWLEDGEMENTS

This investigation was first presented at PEVOC V, the fifth Pan European Voice Conference in Graz, Austria, Aug. 2003. Co-author MT's participation in this investigation was supported by the University College of Music Education in Stockholm (SMI).

REFERENCES

1. Sundberg J, Thalén M. Describing different styles of singing. A comparison of a female singer's voice source in 'Classical', 'Pop', 'Jazz' and 'Blues'. Logoped Phoniatr Vocol 2001; 26: 82–93.
2. Schutte HK, Miller DG. Belting and Pop. Nonclassical approaches to the female middle voice: Some preliminary considerations. J Voice 1993; 7: 142–50.

3. Johnson B. AC/DC, Stiff upper lip, Stiff upper lip. Elektra/Asylum 2000.
4. Coverdale D. Crying in the rain, Whitesnake. Whitesnake/Greatest hits. C.C Songs Ltd/WB Music Corp. 1994.
5. Granqvist S, Hertegård S, Larsson H, Sundberg J. Simultaneous analysis of vocal fold vibration and transglottal airflow: Exploring a new experimental technique. J Voice 2003; 17: 319–30.
6. Hertegård S. Vocal fold vibrations as studied with flow inverse filtering [dissertation]. Stockholm: Department of Logopedics and Phoniatrics, Karolinska Institute; 1994.
7. Granqvist S, Lindestad P-Å. A method of applying Fourier analysis to high-speed laryngoscopy. J Acoust Soc Amer 2001; 110: 3193–7.
8. Švec JG, Schutte HK. Videokymography: High-speed scanning of vocal fold vibration. J Voice 1996; 10: 201–5.
9. Larsson H, Hertegård S, Lindestad P-Å, Hammarberg B. Vocal fold vibrations: high-speed imaging, kymography, and acoustic analysis: A preliminary report. Laryngoscope 2000; 110: 2117–22.
10. Cleveland T, Stone RE, Sundberg J, Iwarsson J. Estimated subglottal pressure in six professional country singers. J Voice 1997; 11: 403–9.
11. Schutte HK. The Efficiency of Voice Production. San Diego, CA: Singular Publishing Group; 1980.
12. Cleveland T, Sundberg J. Acoustic analyses of three male voices of different quality. In: Askenfelt A, Felicetti S, Jansson E, Sundberg J, editors. SMAC 83. Proceedings of the Stockholm International Music Acoustics Conference. Vol. 1. Stockholm: Royal Swedish Academy of Music 1985; No. 46(1): 143–56.
13. Titze I, Sundberg J. Vocal intensity in speakers and singers. J Acoust Soc Amer 1992; 91: 2936–46.
14. Lindestad P-Å, Södersten M, Merker B, Granqvist S. Voice source characteristics in Mongolian "throat singing" studied with high-speed imaging technique, acoustic spectra, and inverse filtering. J Voice 2001; 15: 78–85.
15. Fuks L. From Air to Music [dissertation]. Stockholm: Department of Speech Music Hearing, KTH (Royal Institute of Technology), 1999.