Vocal fold vibration and voice source aperiodicity in phonatorily distorted singing

Zangger Borch, D. and Sundberg, J. and Lindestad, P-A. and Thalén, M.
Vocal fold vibration and voice source aperiodicity in phonatorily distorted singing

Zangger Borch D*, Sundberg J**, Lindestad PA***, Thalén M*

* University College of Music Education in Stockholm (SMI).
** KTH Voice Research Centre, Dept of Speech Music Hearing, KTH
***Dept of Logopedics and Phoniatrics, Karolinska Institutet, Huddinge University Hospital

Corresponding author:
Johan Sundberg, KTH Voice Research Centre, Dept of Speech Music Hearing, KTH, SE-100 44 Stockholm, Phn +468 790 7873, Fax +468 790 7854, email pjohan@speech.kth.se

Abstract
The acoustic characteristics of so-called distorted tones, commonly used in singing rock music, are analysed in one experiment. The results revealed that dist tones contain aperiodicity. SPL @ 0.3 m varied between 90 and 96 dB, produced by a subglottal pressure in the range of 20 to 43 cm H2O, approximately. A doubling of subglottal pressure yielded, on average, an SPL increase of 2.3 dB. In another experiment, the associated vocal fold vibration patterns were recorded by digital high-speed imaging. 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 periodically vibrating vocal folds. The supraglottic mucosa vibration can be assumed to be driven by the high airflow produced by the elevated subglottal pressure.

Introduction
Vocal technique in singing differs greatly depending on the musical style (Sundberg & Thalén, 2001). 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 distorted or “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 (Johnson, 2000; Coverdale, 1994). 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 (Granqvist et al., 2003).

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 (Figure 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. Hertegård, 1994). The flow and oral pressure signals were stored on separate channels of a TEAC PCM signal recorder.
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 Figure 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 300W 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 in a PC.

A custom made program (Granqvist & Lindestad, 2001) 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 (Svec & Schutte, 1996). 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 (Larsson et al., 2000). The program also allows Fourier analysis of the vibrations in an image sequence, such that the vibration frequency of the various parts of the image can be determined (Granqvist & Lindestad, 2001).

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 (1994), was selected for analysis from experiment 1. The song has an ambitus of one octave, A#3 – A#4. Figure 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 the average, a doubling of Ps yielded a 2.3 dB increase of SPL. Similar values have been found in country singers (Cleveland et al., 1997), while considerably greater values are typically observed in classically trained singers (Schutte, 1980; Cleveland & Sundberg, 1985; Titze & Sundberg, 1992).
were compared. Such a comparison is shown in Figure 3. 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, i.e. the portion of the pitch period during which the glottis is closed, see Figure 4. 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.

An interesting question now is what vibrational characteristics cause this variation in pulse amplitude. Figure 5 shows a kymogram from the high-speed recording of a dist tone. The figure reveals vibrations of the supraglottal mucosa, apparently including 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. Underneath this pattern, an apparently periodic opening and closing of the glottis can be observed as an alternation between black and slightly brighter vertical stripes. The inverse filtered flow signal 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. 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, the latter corresponding to F0, Figure 6. These frequencies are harmonically related, since 3*160=480. 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.
In the dist tone analysed 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 quasi-periodic F0 in many cases. 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 (Lindestad et al., 2001; Fuks, 1999). 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, and it is 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 many concerts per week, year after year, without 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 airstream, which brings these structures to vibration. A condition would then be the magnitude of this airstream. The high subglottal pressures associated with dist tones would cause a forceful airstream which, combined with an approximation of the arytenoids, may bring them to vibration.

Conclusions

Dist tones are loud tones, produced with relatively high subglottal pressures, in the range of 20 to 45 cm H2O. 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