ERRATA TO THE 1ST PRINTED EDITION OF SMAC 03 PROCEEDINGS

ver. 1.1, 2003-09-24

Page 6 was missing

Page 222, figure 3 has been shifted down

Page 224 was missing

Pages 459-461, some characters were missing

The on-line version of the proceedings (marked as 2nd edition) has some other small corrections, things such as missing characters or “strange” font substitutions were corrected.

```
PATH = / Science / Physics / Acoustics / Musical Acoustics

KEY = Coupled Piano Strings
URL = http://www.speech.kth.se/music/5_lectures/weinreic/weinreic.html

PATH = / Music / Computer Music / Software

KEY = Planet CCRMA Software Package | Planet CCRMA at Home | Planet CCRMA
URL = http://www-ccrma.stanford.edu/planetccrma/software/
DSC = Linux software packages used at CCRMA for sound and music applications

KEY = Synthesis Tool Kit | STK
URL = http://www-ccrma.stanford.edu/software/stk/
DSC = A C++ tool kit for rapid prototyping of sound synthesis algorithms.
```

Figure 1: *Open Dictionary submission syntax, sent via email to* `od@w3k.org`.

[28] C. Traube and J. Smith, "Estimating the plucking point on a guitar string," in *Proc. Int. Conf. Digital Audio Effects (DAFx-00), Verona, Italy*, Dec. 2000.

[29] A. Krishnaswamy, , and J. O. Smith, "Methods for simulating string collisions with rigid spatial objects," 2003, submitted to WASPAA-03.

[30] P. R. Cook, *Real Sound Synthesis for Interactive Applications*, A. K. Peters, L.T.D., 2002.

[31] J. O. Smith III, *Mathematics of the Discrete Fourier Transform (DFT)*, `http://www-ccrma.stanford.edu/~jos/mdft/`, 2003.

[32] J. O. Smith III, *Introduction to Digital Filters*, `http://www-ccrma.stanford.edu/~jos/filters/`, 2003.

[33] J. O. Smith, "The Open Dictionary," 2003, `http://www-ccrma.stanford.edu/~jos/od/`.

[34] A. Almeida, C. Vergez, R. Causs, and X. Rodet, "Applications of bioacoustics in physical modeling and the creation of new musical instruments," in *Proc. Int. Symp. on Musical Acoustics (ISMA-02), Mexico City*, 2002, available online at `http://-www.ircam.fr/equipes/instruments/almeida/articles/`.

[35] F. Avanzini, S. Serafin, and D. Rocchesso, "Modeling interactions between rubbed dry surfaces using an elasto-plastic friction model," in *Proc. COST-G6 Conf. Digital Audio Effects (DAFx-02), Hamburg, Germany*, September 2002, pp. 111–116.

[36] C. Erkut, *Aspects in Analysis and Model-Based Sound Synthesis of Plucked String Instruments*, Doctoral thesis, Helsinki University of Technology, Espoo, Finland, November 2002, available online at `http://lib.hut.fi/Diss/2002/isbn9512261901/`.

[37] C. Erkut, M. Karjalainen, P. Huang, and V. Välimäki, "Acoustical analysis and model-based sound synthesis of the kantele," *J. Acoust. Soc. of Amer.*, vol. 112, no. 4, pp. 1681–1691, October 2002.

[38] G. Essl, *Physical Wave Propagation Modeling for Real-Time Synthesis of Natural Sounds*, Ph.D. thesis, Computer Science Dept., Princeton University, November 2002, available online at `http://-ncstrl.cs.princeton.edu/expand.php?id=TR-659-02`.

[39] M. Karjalainen, V. Välimäki, and P. A. A. Esquef, "Efficient Modeling and Synthesis of Bell-Like Sounds," in *the 5th Int. Conf. Digital Audio Effects*, Hamburg, Germany, September 26-28 2002, pp. 181–186.

[40] P. Huang, S. Serafin, and J. Smith, "A waveguide mesh model of high-frequency violin body resonances," in *Proc. 2000 Int. Computer Music Conf., Berlin*, Aug. 2000.

[41] T. Hélie, C. Vergez, and X. Rodet, "Virtual musical instruments: Contribution to physical modeling and control of self-sustained instruments," in *Conférence Systemics Cybernetics and Informatics (SCI'2001), Orlando, Florida*, 2001, vol. X, pp. 547–550.

[42] L. Savioja and V. Välimäki, "Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques," *IEEE Trans. Speech and Audio Processing*, pp. 184–194, March 2000.

[43] M. Laurson, C. Erkut, V. Välimäki, and M. Kuuskankare, "Methods for Modeling Realistic Playing in Acoustic Guitar Synthesis," *Computer Music Journal*, vol. 25, no. 3, pp. 38–49, 2001.

[44] X. Rodet and C. Vergez, "Nonlinear dynamics in physical models: Simple feedback-loop systems and properties; from basic models to true musical instruments," *Computer Music J.*, vol. 23, no. 3, pp. 18–49, 1999.

[45] L. Trautmann, B. Bank, V. Välimäki, and R. Rabenstein, "Combining Digital Waveguide and Functional Transformation Methods for Physical Modeling of Musical Instruments," in *the AES 22nd Int. Conf. Virtual, Synthetic and Entertainment Audio*, Espoo, Finland, June 15-17 2002, pp. 307–316.

[46] V. Välimäki, M. Laurson, and C. Erkut, "Commuted waveguide synthesis of the clavichord," *Computer Music Journal*, vol. 27, no. 1, pp. 71–82, Spring 2003.

[47] M. van Walstijn and J. O. Smith, "Use of truncated infinite impulse response (TIIR) filters in implementing efficient digital waveguide models of flared horns and piecewise conical bores with unstable one-pole filter elements," in *Proc. Int. Symp. Musical Acoustics (ISMA-98), Leavenworth, Washington*. June 28 1998, pp. 309–314, Acoust. Soc. of Amer., available online at `http://www-ccrma.stanford.edu/~jos/tiirts/`.

[48] C. Vergez and X. Rodet, "New algorithm for nonlinear propagation of a sound wave application to a physical model of a trumpet," *Journal of Signal Processing*, 2000.

[49] G. P. Scavone, Ed., *CCRMA Overview, April 2003*, Stanford University Department of Music Technical Report STAN-M-112, Apr. 2003, (62 pages) available from `info@ccrma.stanford.edu` or `http://www-ccrma.stanford.edu/overview/`.

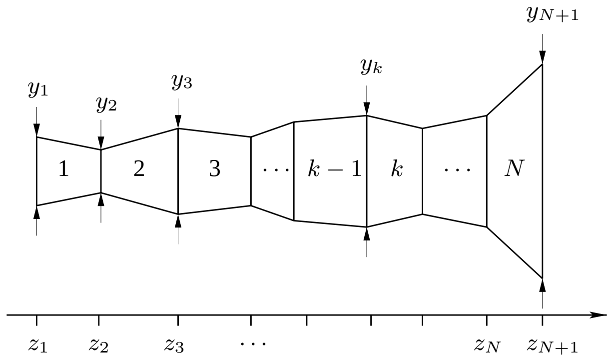

Figure 1: *A horn approximated as a series of truncated cones.*

2.2. Mathematical model of the horn

An established method for the computation of the input impedance Z_{in} of wind instruments is the one-dimensional transmission line analogy[2], where the instrument is approximated as a series of N truncated cones, as in Fig. 1. This model is accurate for moderately flaring horns. Denoting by

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \tag{4}$$

the transfer matrix of the instrument, the input impedance is given by

$$Z_{\text{in}} = \frac{H_{12} + H_{11} Z_L}{H_{22} + H_{21} Z_L}, \tag{5}$$

where Z_L is the radiation impedance. $\mathbf{H}$ is approximated by the product of the transfer matrices of the individual sections[6];

$$\mathbf{H} = \prod_{j=1}^{N} \mathbf{H}_j(y_j, y_{j+1}, z_j, z_{j+1}). \tag{6}$$

In this paper, an expression according to Beranek[7] is used to approximate Z_L.

2.3. Computation of ∇Z_{in}

Possible design variables for the horn approximated as in Eq. (6) are the end diameters $y_1, \ldots, y_{N+1}$, and the positions $z_1, \ldots, z_{N+1}$. The gradient of Z_{in} needs to be computed in an accurate and computationally inexpensive way. This precludes the use of finite differencing of the form

$$\frac{\partial Z_{\text{in}}}{\partial y_k} \approx \frac{Z_{\text{in}}(y_k + \varepsilon) - Z_{\text{in}}(y_k)}{\varepsilon}, \quad \varepsilon > 0,$$

since then the computational cost of the gradient grows as N^2. Fortunately, it is possible to differentiate Eq. (5) symbolically. This operation requires, in turn, differentiation of $\mathbf{H}$. Introducing

$$\mathbf{\Pi}_{j,k} = \mathbf{H}_j \mathbf{H}_{j+1} \cdots \mathbf{H}_k, \quad k > j,$$

the product rule for differentiation yields for $3 \leq k \leq N-2$

$$\begin{aligned} \frac{\partial \mathbf{H}}{\partial y_k} &= \frac{\partial}{\partial y_k}(\mathbf{\Pi}_{1,k-2}\mathbf{H}_{k-1}\mathbf{H}_k\mathbf{\Pi}_{k+1,N}) = \\ &= \mathbf{\Pi}_{1,k-2}\left(\frac{\partial \mathbf{H}_{k-1}}{\partial y_k}\mathbf{H}_k + \mathbf{H}_{k-1}\frac{\partial \mathbf{H}_k}{\partial y_k}\right)\mathbf{\Pi}_{k+1,N}. \end{aligned} \tag{7}$$

Similar expressions apply for $k = 1, 2, N, N+1$, and for differentiation with respect to z_k. In a computer implementation, the products $\mathbf{\Pi}_{1,1}, \mathbf{\Pi}_{1,2}, \ldots, \mathbf{\Pi}_{1,N-1}$, and $\mathbf{\Pi}_{N,N}, \mathbf{\Pi}_{N-1,N}, \ldots, \mathbf{\Pi}_{2,N}$ can be computed in a total of $2N-3$ matrix multiplications. All intermediate cumulative products are stored for later retrieval in the evaluations of Eq. (7). It is also necessary to differentiate Z_L with respect to y_N and y_{N+1}.

3. FINDING SMOOTH SOLUTIONS

3.1. Optimisation with $\{y_k\}$ as design variables

The most straightforward choice of design variables is simply the set $\{y_k\}$ itself. As Fig. 1 suggests, the number of design variables equals the number of segments. In a first experiment, the lowest three impedance peaks of a Bessel horn (Fig.2) of length 0.5 m were lowered by 10 Hz. y_1 and y_{N+1} were kept fixed, as well as the segment lengths. Although the inequality bound $y_k > 0$ is natural, it is not imposed in this, and the subsequent, experiments due to limitations of the used Levenberg-Marquardt implementation. This workes surprisingly well, presumably because narrow horn profiles usually make the objective function diverge quickly. This can often prevent the optimisation from continuing in the direction of negative values of y_k.

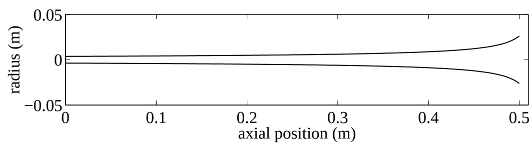

Figure 2: *Model horn of Bessel type.*

Figure 3 shows the shapes obtained for three values of N. Although a moderate impedance peak displacement was requested, the resulting shapes are infeasibly irregular. Refining the resolution seems to relieve the problem somewhat, but a considerable amount of ripple remains even for $N = 400$. The problems become even more severe if larger peak displacements are tried. An effort to shift the peaks as above, but by 20 Hz, fails due to a complete breakdown of the optimisation. It is thus not practical to optimise directly with respect to the segment end diameters.

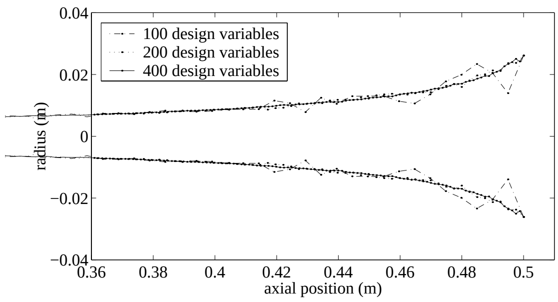

Figure 3: *Resulting shapes for an intonation experiment with optimisation directly with respect to the segment end diameters.*

to its first nine impedance peaks. The aim was then to reproduce the shape of this horn from an initial design shaped as a truncated cone with the same length and end diameters as the original Bessel horn. The number of segments N was 100. The properties of the start design (the cone) and the desired horn (the Bessel horn) are shown in table 1. After 121 iterations, the shape shown in Fig. 6 was obtained. Indicated are also the initial shape and the original Bessel horn . The optimiser has found a shape that is very close to the original one. Some wiggles of the contour remain, but on the whole it is remarkable that the specification of only nine impedance peak frequencies and magnitudes yields a shape so close to the original design.

start impedance		goal impedance	
freq. (Hz)	magn. (Ω)	freq. (Hz)	magn. (Ω)
290.3	$6.5 \cdot 10^7$	266.5	$2.0 \cdot 10^8$
591.3	$8.0 \cdot 10^7$	594.0	$1.4 \cdot 10^8$
904.4	$7.2 \cdot 10^7$	921.8	$9.5 \cdot 10^7$
1226	$5.9 \cdot 10^7$	1249	$6.6 \cdot 10^7$
1553	$4.8 \cdot 10^7$	1574	$4.6 \cdot 10^7$
1883	$4.0 \cdot 10^7$	1900	$3.3 \cdot 10^7$
2215	$3.4 \cdot 10^7$	2225	$2.5 \cdot 10^7$
2549	$3.0 \cdot 10^7$	2550	$2.1 \cdot 10^7$
2883	$2.7 \cdot 10^7$	2874	$1.8 \cdot 10^7$

Table 1: *Start and goal properties for a reconstruction experiment with nine peaks. The figures refer to the impedance peak frequencies and magnitudes.*

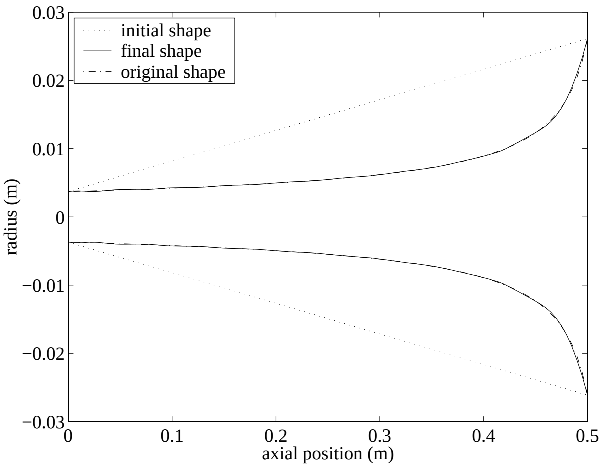

Figure 6: *Original shape, and the shape obtained after optimisation considering the frequencies and magnitudes of the nine first impedance peaks.*

4. COMPUTATIONAL PERFORMANCE

The computer code of this project has been written in MATLAB, utilising optimisation routines from the Optimization toolbox. Although the interpreting language used is far from optimal in terms of execution speed, a typical intonation experiment with 100 design variables takes only in the order of one minute on a personal computer. This figure can probably be reduced considerably by porting the code to a compiling language, such as `fortran`.

5. CONCLUSIONS

Gradient based minimisation is effective for brass optimisation, provided that a suitable choice of design variables is chosen. The demands on computational power can be satisfied with a modern desktop computer. Still, the algorithms work fast enough to make them tractable as a hands-on virtual laboratory, with possibilities to interactive experimentation.

Although the algorithms presented here are primarily intended for intonation optimisation that require rather small changes of the initial shape, they have been used successfully for more difficult design problems. For a systematic use of the presented methods in wind instrument design "from scratch" it would, however, be beneficial to combine them with global minimisation algorithms.

A limitation of the transmission line model is the degraded accuracy for high frequencies, where modal conversion in the rapidly flaring part of the bell is considerable. Future development of the algorithm presented here may involve the use of a finite element description of the outermost part of the bell. Gradient based optimisation for such a model has been reported in [9].

6. REFERENCES

[1] A. H. Benade and E. V. Jansson. On plane and spherical waves in horns with nonuniform flare. i. theory of radiation, resonance frequencies, and mode conversion. *Acustica*, 31:79–98, 1974.

[2] R. Caussé, J. Kergomard, and X. Lurton. Input impedance of brass musical instruments—Comparison between experiment and numerical models. *J. Acoust. Soc. Am.*, 75:241–254, 1984.

[3] V. Pagneux, N. Amir, and J. Kergomard. 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, 1996.

[4] W. Kausel. Optimization of brasswind instruments and its application in bore reconstruction. *Journal of New Music Research*, 30:69–82, 2001.

[5] J. J. Moiré. The Levenberg-Marquardt algorithm: Implementation and theory. In G. A. Watson, editor, *Numerical Analysis*, pages 105–116. Springer Verlag, New York, 1977.

[6] D. Mapes-Riordan. Horn Modeling with Conical and Cylindrical Transmission-Line Elements. *J. Audio Eng. Soc.*, 41:471–483, 1993.

[7] L. Beranek. *Acoustics.* McGraw-Hill, New York, 1954.

[8] A. H. Benade. *Fundamentals of Musical Acoustics*. Dover Publications, New York, 1990.

[9] E. Bängtsson, D. Noreland, and M. Berggren. Shape optimization of an acoustic horn. *Computer Methods in Applied Mechanics and Engineering*, 192:1533–1571, 2003.

VOCAL FOLD RESONANCES AT LOW AND HIGH PITCH TUNING
-preliminary results

Stellan Hertegård[1] and Hans Larsson[1], Svante Granqvist[2]

[1]Karolinska Institute, Department of Logopedics and Phoniatrics
Huddinge University Hospital,

[2]Department of Speech Music and Hearing
Royal Institute of Technology.
Stockholm Sweden

stellanhertegard@hotmail.com

ABSTRACT

Four trained singers were examined with high-speed recording during exernal acousic excitation of the vocal tract with the vocal folds kept in a phonatory position, but without phonating (so-called Kaneko manoeuvre). Automatic glottal area analysis was made and spectral analysis was performed on the obtained area waveform curves. The results showed resonance frequencies close to habitual phonation frequency in three out of the four subjects. Intended pitch at phonation seems to affect the resonance frequency as well as the voice category of the subjects.

INTRODUCTION

Previous stroboscopic examinations indicate that the mechanical properties of the vocal folds differ between high- and low pitched singing. There may also be an interaction with the vocal tract resonances. The mechanical properties of the vocal folds are pertinent to both the speaking and singing voice. These characteristics are usually investigated directly with endoscopy during phonation or indirectly by means of acoustical methods. The mode of phonation itself might affect the vibratory modes found at laryngoscopy. An alternative approach was suggested by Kaneko et al [1] and by Švec et al [2] who used external stimulation for excitation of the voice source when the vocal folds were kept in a phonatory position (a so-called Kaneko manoeuvre) during a frequency sweep. A similar experiment was performed by Granqvist et al [3] on a normal male speaker using acoustic excitation. In this experiment, a loudspeaker and a sweep tone was utilised for the excitation and the vocal fold response was studied with high-speed imaging. The results showed three different resonances for a relaxed phonatory setting. The aim of the present study is to examine the vocal fold resonances at low and high pitch tuning. A second aim was to study differences between 3 different voice categories, soprano, mezzo-soprano and barytone from the stimulation experiment.

Material and Task

Four subjects, one soprano, one mezzo-soprano and two barytones (age between 40 and 47 years) participated in the study. They were all healthy non-smokers with normal larynges. Three of the subjects were trained professional opera singers, whereas one subject (barytone) was an amateur singer with some voice training.

The task was to perform the Kaneko manoeuvre, i.e to phonate followed by a cease in phonation but keeping the phonatory vocal fold setting during the frequency sweep. The subjects were provided live feedback from the camera monitor in order to complete this rather difficult task. Two recordings were made of each subject, first phonations / manoeuvres at habitual pitch (chosen by the subject) and later phonations / manoeuvres at a higher pitch (approximately one octave higher).

Experiment Apparatus

For the experiment a loudspeaker was connected to a plastic tube with an inner diameter of 28 mm. A laryngeal endoscope (Storz 70 degree) was passed through a hole in the tube (which was sealed tight with plasticine to avoid losses) and connected to a Weinberger high-speed camera (image resolution 256x64pixels at 1904 frames/s). A custom made software was used to create the sound stimulus: For the females a sinusoidal sweep tone from 500-100 Hz was used and for the male subjects a tone from 250-50Hz. The sound pressure level was typically in the range 140-150 dB inside the tube and vocal cavity. The duration of the sweeps was selected to 3 seconds in order to fit within the 4 second capacity of the high-speed camera. The sweep tone was controlled from one computer and the high speed images during the sweep was recorded on another computer. The sound pressure level (SPL) inside the tube was measured by means of a Brüel & Kjaer (B&K) condenser microphone (4192) and a B&K microphone amplifier (2669) and recorded on the same computer that was used for the sweep generation. SPL was calibrated by a B&K calibrator, which generated a constant SPL of 94 dB.

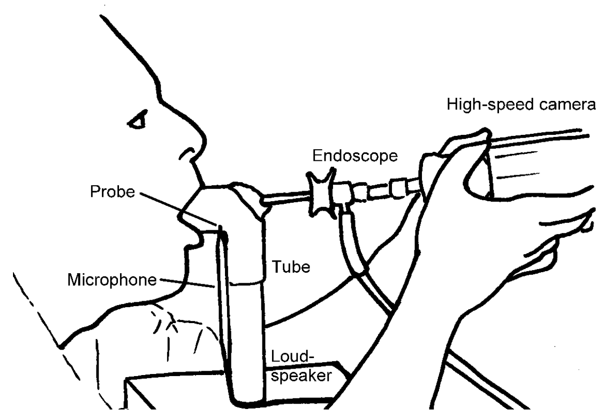

Figure 1: The experimental apparatus. The subject is phonating with the mouth closed around a tube where the acoustic excitation is fed from a loudspeaker. An endoscope connected to the high speed camera is passed through an airtight hole in the tube

Analyses

The software High speed toolbox was used for analysing the high speed recordings of the vocal fold vibration. The program can automatically detect the edges of the vocal folds and generate an area waveform, This waveform will correspond to the motion of the vocal folds as excited by the frequency sweep. The automatic area extraction sometimes had to be manually corrected. A special software was used for FFT analysis of the area and sound waveforms The software used a simple algorithm to detect F0 in the waveform files; the highest peak in the FFT spectrum was assumed to correspond to the F0. While this is not generally a good F0 extraction method for voice, it is sufficiently good for the purpose of this experiment, where the exciting signal was a sinusoid. Using quadratic interpolation, the F0 values could be detected with high precision even though the sampling rate was as low as 4 times the highest frequency of interest (1904 Hz and 500 Hz, respectively) The program used semi-automatic detection of the frequency sweeps in the area and sound waveforms. Since the sweeps were exponential, exponential regression was utilised on the part of the extracted F0 curve that contained the frequency sweep. Thus, two equations describing the location of the F0 within the two waveforms could be obtained These equations was then utilised to select a number of positions in the waveform files, re-do the FFTs for these positions and detect the height of the peak at F0. These heights correspond to the response of the sound (upper curve) and area (middle curve), respectively, and are plotted in the lower panel of the computer program. Also, the difference between these two curves is plotted (lower curve), and this curve corresponds to the transfer function from exciting sound pressure to glottal area. (Figure 2)

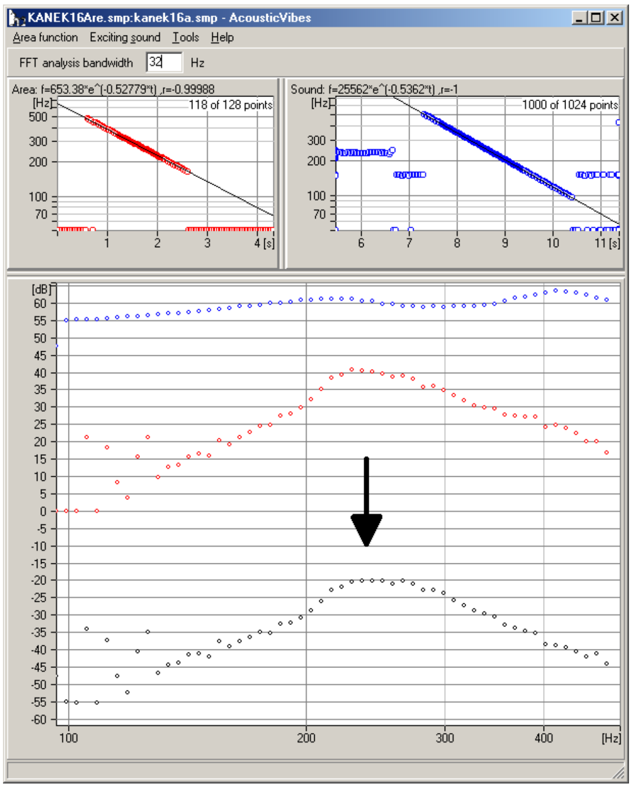

Figure 2 Analysis of resonance frequency in the glottal area variation during the sinusoid sweep. The sweeps are shown in the upper panels, and the response curves are shown in the lower panel. The lower curve in the bottom panel shows the compensated response curve of the glottal area; note the resonance at around 230Hz (arrow).

Results

The preliminary results show that all subjects had a resonance frequency close to the F0 of the phonations before the Kaneko manoeuvre. The subject barytone 1 also had a lower resonance at 80 Hz. The resonances of the high-pitched phonatory settings were difficult to analyse but seemed to be lower than the F0 of the phonations before the Kaneko manoeuvres. There were also clear differences between the different voice categories, as expected with the soprano and mezzo-soprano showing the highest habitual setting resonances. The preliminary results are summarised in table 1.

Subject	**Phonation 1** Resonance (habitual pitch setting)	**Phonation 2** Resonance (higher pitch setting)
Soprano	230Hz (230Hz)	300Hz (345Hz)
Mezzosoprano	200 Hz (210Hz)	300Hz (420Hz)
Barytone 1 (pro)	75Hz (115Hz)	80Hz and 180Hz (220-225Hz)
Barytone 2 (amateur)	120Hz (120-140Hz)	180Hz (190Hz)

Table 1: *Resonances found for the habitual phonatory settings and for the phonatory settings with higher pitch, within brackets:F0 for the phonations before the Kaneko manoeuvre.*

DISCUSSION and CONCLUSIONS

The method of exciting the vocal folds by means of an acoustic signal seems to work well, and thus it seems possible to study vocal fold resonances with the technique used. The area analysis is rather time-consuming, in particular when the automatic detection needs manual correction. The technique also requires subjects who are able to perform the Kaneko manoeuvre. The results indicate that the method can be used for studying the vocal fold's small-signal resonances. These resonances are probably important factor in the description of vocal fold mechanics and tissue characteristics.

In the future, it may be worthwhile to utilise some other signal than of the glottal area, e.g. a single point on the edge of the vocal fold. In this way, it would be possible to analyse the properties of the vocal folds separately, and also to analyse the response at different locations along the vocal folds.

REFERENCES

[1] Kaneko, T, Masuda, T, Shimada, A., et al. Resonance characteristics of the human vocal fold in vivo and in vitro by an impulse excitation. Laryngeal function in phonation and respiration, in Vocal Fold Physiology Eds Bar, Sasaki, Harris., Collge Hill Press. Boston MA pp 349-365, 1987.

[2] Švec, J, Horacek, J, Sram, F., et al. Resonance properties of the vocal folds: in vivo laryngoscopic investigation of the externally excited laryngeal vibrations. J Acoust Soc Am. 2000 Oct 108(4):1397-407.

[3] Granqvist, S, Hertegård, S. In vivo examination of resonances in the larynx using external acousitc stimulation. Presentation at Pevoc 4 Stockholm Sweden August 23-28 2001