17th International Conference
ENGINEERING MECHANICS 2011
Svratka, Czech Republic, 9 – 12 May 2011

NUMERICAL SIMULATION OF ACOUSTIC CHARACTERISTICS OF PROFESSIONAL VOICE BASED ON MRI AND ACOUSTIC MEASUREMENTS

V. Radolf*, A. M. Laukkanen, R. Havlík***, J. Horáček***

Abstract: *Inverse method was used for numerical simulation of acoustic characteristics of a professional musical actor before and after vocal exercises. The geometrical data for a 1D model of the acoustic cavities of the vocal tract were evaluated from magnetic resonance images (MRI) registered during sustained phonation of vowel [a:] before and after a vocal warm-up. The numerically simulated voice signals are compared with the acoustic recordings and warm-up related changes in the vocal tract are discussed.*

Keywords: ***Biomechanics of voice, singer's and speaker's formant cluster, acoustic effects of vocal exercises.***

1. Introduction

In operatic singing, singers make use of a special voice quality in order to be heard over the orchestra without a microphone. The important acoustic component which determines the operatic quality of the voice, especially in male singers, is the so-called "singer's formant" (Sundberg, 2003; Titze, 2000). Similarly the voice quality of speakers, especially actors, is improved by the so-called "speaker's formant" (Leino et al., 2011). The epilarynx tube located just above the vocal folds is theoretically considered as a dominant source of clustering of formant frequencies especially when the cross-sectional area of this tube is approximately six times less than the area of the lower pharynx (Story, 2003). However, the physiological adjustment used in singers or in actors has not been known in sufficient details. The rationale of the present paper is to estimate the anatomical/geometrical adjustments and to model the acoustic changes that occur in the voice of a male professional musical actor after vocal warming up by vocal exercises.

2. Methods

2.1. MRI experiment

A Czech male musical actor (60 years, baritone) served as a subject in the magnetic resonance imaging investigation. Lying supine in the MRI machine he first produced the vowel [a:] in a naive technique and after ca 5 minute vocal exercising again in a professional "musical actors" manner, aiming at the best voice quality. Each sample was produced for at least 20 s to enable the MRI scanning. The samples were produced on a comfortable pitch, at approximately the same fundamental frequency. Being vocally trained, the subject was able to keep the articulation and phonation constant through a sustained vowel phonation. The subject's head position was stabilized with a support. Before starting the MRI measurement the subject phonated each sample with normal auditory feedback (i.e. without the MRI device on) and sustained the same phonation setting during MRI scanning. Due to noise and magnetic field no acoustic recording was possible during the MR imaging. For acoustic measurements the subject's voice was recorded during the same tasks afterwards in a

* Ing. Vojtěch Radolf, Ph.D. and Ing. Jaromír Horáček, DrSc.: Institute of Thermomechanics, Academy of Sciences of the Czech Republic; Dolejškova 1402/5; 182 00, Prague; CZ, e-mails: radolf@it.cas.cz and e-mail: jaromirh@it.cas.cz

** prof. Anne-Maria Laukkanen, Ph.D.: Speech and Voice Research Laboratory, School of Education, University of Tampere; FIN-33014, Tampere; Finland, e-mail: Anne-Maria.Laukkanen@uta.fi

*** MUDr. Radan Havlík, Ph.D.: AUDIO – Fon Centr. s.r.o.; Obilní trh 4; 602 00, Brno; CZ, e-mail: radan.ha@seznam.cz

sound-treated studio. The sampling frequency of 44.1 kHz was used. MRI scanning was performed at the Dept. of Medical Imaging, St. Anne's Faculty Hospital in Brno, using a 1.5 Tesla MRI device (Symphony Magnetom, Siemens).

The imaging parameters were as follows: Field of view 236 x 270 mm^2, slice thickness 1.5 mm, acquisition time 20.07 s, number of averages 1, repetition time 5.49 s, echo time 2.88 s, number of sagittal images 44, resolution 512 x 448 pixels. For viewing the MR images and for measuring the changes of the vocal tract a Syngo FastView software (Siemens AG) and image software (ImageJ 1.42q Wayne Raspand, National Institutes of Health, USA and ITK-Snap, version 2.0) were used.

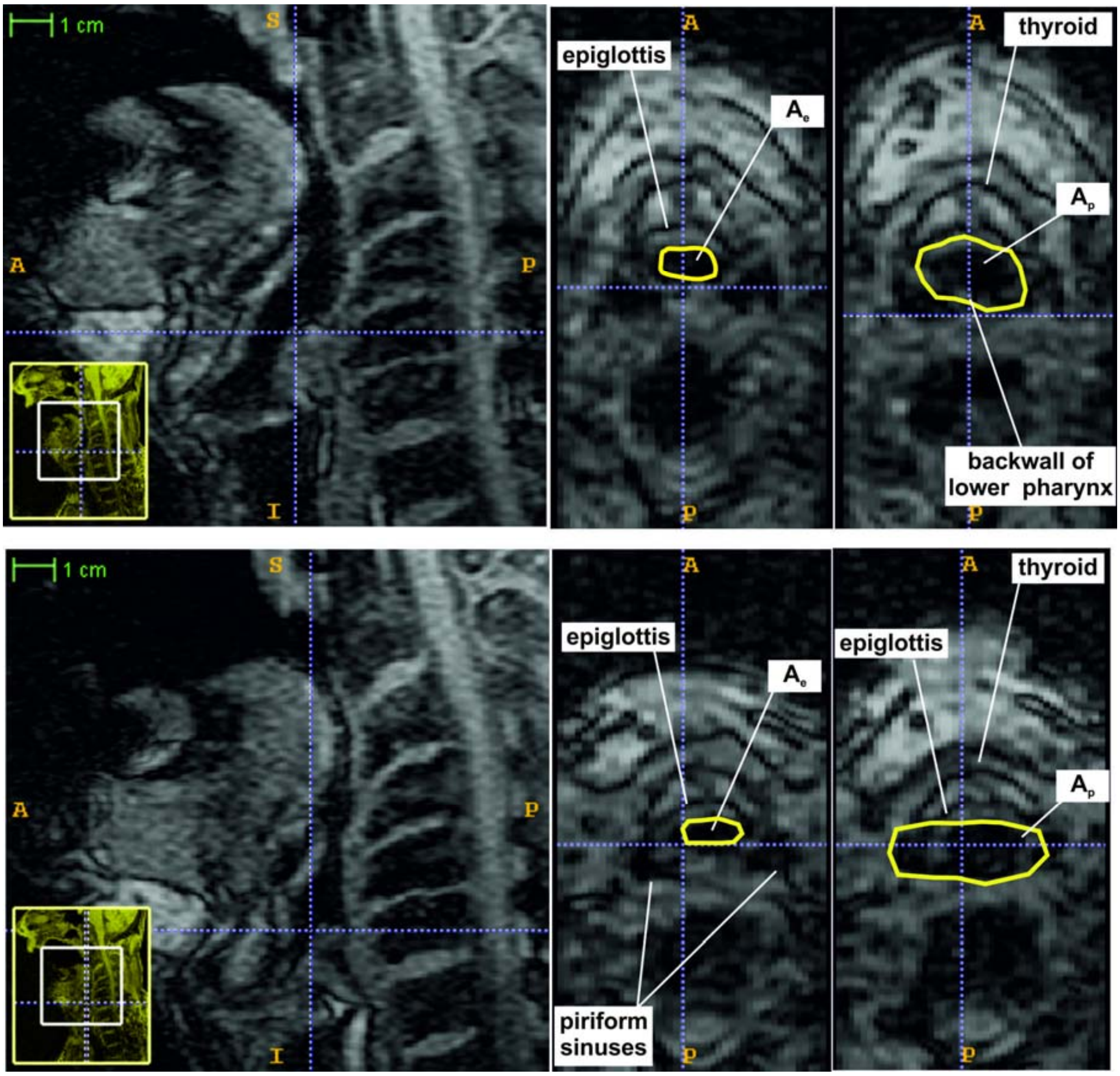

Fig. 1: MR images for the vowel [a:] before (upper panel) and after (lower panel) the vocal exercising and measurement of the transversal areas. Left: Midsagittal views of the corresponding locations. In the middle: Area A_e of the epilaryngeal outlet. Right: Area A_p of the pharyngeal inlet (just above the plane where the epilaryngeal cavity merges with the pharyngeal cavity).

Especially the areas of the outlet of the epilarynx and the inlet of the low pharynx were studied from transversal slices, see Figure 1. The area of the inlet to the pharynx was studied just above the collar of epiglottis, while the area of the outlet of the epilarynx was studied just below the collar of epiglottis, at the point where the epilaryngeal tube and the sinus piriformes are separated. The ratio of the inlet to the pharynx over the outlet of the epilarynx was calculated from the areas.

2.2. Analyses

The MR images revealed lowering of the larynx (about 19 mm), raising of the soft palate, prolongation of the vocal tract (ca 17 mm) and of the epilaryngeal tube (ca 4.5 mm) and narrowing of the epilaryngeal region and/or widening of the pharyngeal region in phonation after the vocal exercising.

The acoustic analyses showed that a cluster of two or three formants was formed in the range of F_3-F_5 after exercising (see Fig. 2 and Tab. 1) that leads to a stronger speaker's/singer's formant. A weak resonance between F_2 and F_3 can be caused by a slight nasality.

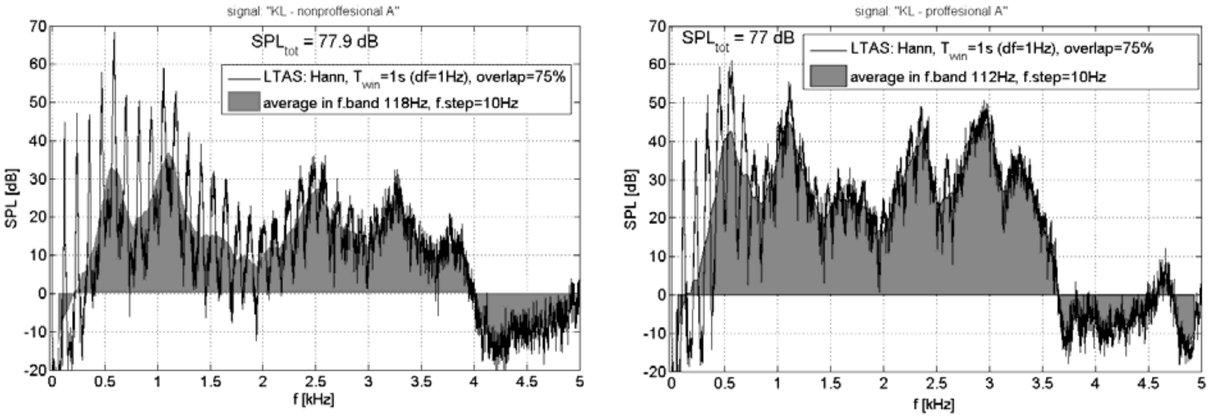

Fig. 2: Spectra (LTAS) of the acoustic signals before (left) and after (right) the vocal exercise.

Tab. 1: Fundamental frequency (F_0) and the five lowest formant (F_1 - F_5) frequencies evaluated from the acoustic recordings and obtained from the modeling.

[Hz]	F_0	F_1	F_2	F_3	F_4	F_5
Acoustic signal – before exercising	*118*	*560*	*1100*	*2510*	*3260*	*3780*
Model – before exercising		*560*	*1099*	*2507*	*3254*	*3787*
Acoustic signal - after exercising	*112*	*560*	*1100*	*2370*	*2940*	*3260*
Model - after exercising		*560*	*1099*	*2367*	*2935*	*3268*

2.3. Modeling approach

The possible vocal tract changes resulting in the formation of a speaker's (or singer's) formant cluster were studied using a 1D mathematical model of voice production. The 1D vocal tract model was developed from the 3D volume model obtained from the MR images (Vampola et al., 2008). The formant frequencies measured from the vowel [a:] recorded from the subject of the present study before and after exercising were prescribed to the model and by a tuning procedure (changing the vocal tract shape, i.e. the size of area cross-sections) the best fitting vocal tract configurations were obtained.

Length of the real vocal tract of the subject was measured using MRI data. The values L_{BEF} = 182.5 mm and L_{AFT} = 199.5 mm were used for total length of the vocal tract model before and after vocal exercising. Vocal tract channel was modeled as a system of conical elements of 4 mm in length except for the last element which was 2.5 mm long in the case "before" and the first and the last element in the case "after" which were 1 mm and 2.5 mm long respectively.

The mathematical model used is based on an analytical solution of 1D wave equation for conical acoustical elements in frequency domain. Radiation impedance at lips as well as viscous losses in the model are considered. Tuning procedure uses sensitivity functions as described in Leino et al. (2011). The iteration stops when the root of the sum of the squared differences between desired and instantaneous eigenfrequencies is less than a desired tolerance value. This value was set to 10 Hz in both cases. The speed of sound, the density, and the dynamic viscosity of the air were considered as follows: c_o = 353 ms^{-1}; ρ_o = 1.2 kgm^{-3}, μ = 1.8·10^{-5} kgm^{-1}s^{-1}. Number of iteration steps was 56 (for the case before) and 193 (for after). The results of the modeling are summarized in Fig. 3 and Tab. 1.

3. Conclusions

The area ratio of the lower pharynx over the epilaryngeal tube increased after the exercises. Results from modeling were A_p/A_e = 260mm^2/85mm^2 = 2.59 for phonation before (coordinates 16-28 mm) and

220mm^2/46mm^2 = 4.78 for after (coordinates 0 - 20 mm). This is in agreement with the measured areas from the MRI where A_p/A_e = 3.32 for before increased to 5.61 for after. The change in the areas is in agreement with the results found for operatic singers by Story (2003) as well as the changes in the formant frequencies measured in the present study. After the vocal exercising the formants F_1 and F_2 of [a:] remained nearly unchanged while the formants F_3 – F_5 were clustering, thereby considerably increasing the SPL level (ca 15 dB) in the frequency band of 2-4 kHz. The distance between F_3 and F_5 in the acoustic measurement decreased from 1270 Hz down to 890 Hz, and in the modeling the distance between these formants decreased from 1280 Hz to 901 Hz (see Tab. 1). Fig. 3 shows increased volume of the pharyngeal part, the changes in the oral cavity that can be realized by changes of the tongue position, while the lips opening remained the same as before the vocal exercising.

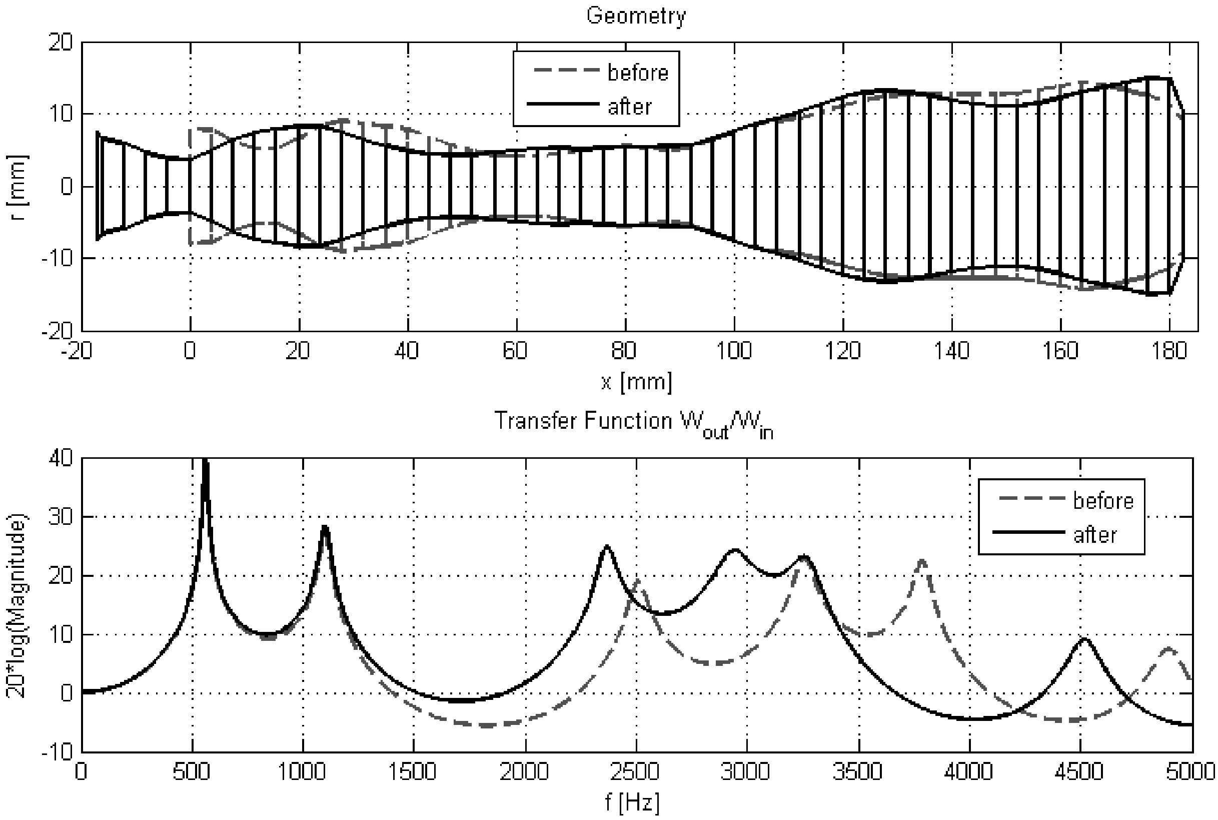

Fig. 3: Vocal tract geometry (top) and transfer function (below) between acoustic airflow rate at the vocal folds and the airflow rate at the lips – results of modeling.

Acknowledgement

The research was supported by the research plan AV0Z20760514 of the Institute of Thermomechanics and by the project GAČR 101/08/1155. The authors are also very grateful to Doc. MUDr. Petr Krupa from the Hospital U Svaté Anny in Brno for enabling the MRI measurements.

References

Leino, T., Laukkanen, A. M. & Radolf, V. (2011) Formation of the Actor's/Speaker's Formant: A Study Applying Spectrum Analysis and Computer Modeling. Journal of Voice, 25(2), pp. 150-158.

Story, B. H. (2003) Using imaging and modeling techniques to understand the relation between vocal tract shape to acoustic characteristics, in: Proc. of the Stockholm Music Acoustics Conf., Aug. 6-9, 2003 (SMAC 03), Stockholm, Sweden, pp. 435-438.

Sundberg, J. (2003) Research on the singing voice in retrospect. Speech, Music and Hearing, KTH Stockholm, TMH-QPSR Vol. 45, pp. 11-22.

Titze, I. R. (2000) Principles of voice production. Iowa City, IA: National Center for Voice and Speech.

Vampola, T., Horáček, J. & Švec, J. G. (2008) FE modeling of human vocal tract acoustic. Part I: Production of Czech vowels. Acta Acoustica united with Acoustica, 94, pp. 433-447.