PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

[Acoustic characterization of an averaged vocal tract model based on the MRI data of professional tenors]

Judith PROBST(1), Alexander LODERMEYER(1), Sahar FATTOUM(2), Matthias ECHTERNACH(3), Stefan BECKER(1), Stefan KNIESBURGES(2)

(1)Department of Process Machinery and Systems Engineering, Friedrich-Alexander-University Erlangen-Nürnberg, Cauerstrasse 4, 91058 Erlangen, Germany; pro@ipat.uni-erlangen.de (J.P); lo@ipat.uni-erlangen.de (A.L.); sb@ipat.uni-erlangen.de (S.B)

(2)Division of Phoniatrics and Pediatric Audiology at the Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital Erlangen, Medical School at Friedrich-Alexander-University Erlangen-Nürnberg, Waldstraße 1, 91054 Erlangen, Germany; sahar.fattoum@uk-erlangen.de (S.F.); stefan.kniesburges@uk-erlangen.de (S.K.)

(3) Division of Phoniatrics and Pediatric Audiology at the Department of Otorhinolaryngology, Head and Neck Surgery, University of Munich, Campus Grosshadern, Marchioninistrasse 15, 81377 München, Germany; matthias.echternach@med.uni-muenchen.de

Abstract

Human speech is created in the larynx and the vocal tract. For vowels, the basic sound is generated by the oscillating vocal folds and filtered by the vocal tract to produce the typical formants. In the present study, we investigate the influence of the vocal tract on phonation. Based on Magnet Resonance Imaging (MRI) data, the acoustic behavior of the vocal tract can be characterized using numerical and experimental models. However, previous studies only analyzed single vocal tracts that were directly obtained from individuals. To investigate the basic influence of the vocal tract on phonation, a vocal tract model is necessary that shows preferably little individual but the typical acoustic characteristics of the specific vowel. Hence, we computed an average vocal tract geometry based on six simplified individual vocal tracts that were extracted from MRI images of professional tenors. The resulting vocal tract model was evaluated in an experimental setup including artificial vocal folds. The comparison with the tenors' vocal recordings shows that the model reproduces the acoustic properties well. Hence, the average vocal tract model presented here is suitable for future investigations of the influence of the vocal tract on phonation.

Keywords: voice production modeling; vocal tract acoustics; artificial vocal fold models

1 INTRODUCTION

The human voice is generated by the flow induced oscillation of the vocal folds. This acoustic primary signal, which is composed of the fundamental frequency of the vocal fold oscillation, the higher harmonics and the broadband sound, is filtered through the vocal tract and thus formed into the different sounds of a language [1]. The human vocal tract can in an idealized way be considered as resonator. The resonance frequencies, which form in the vocal tract and depend on its geometry, are called formants. The first two formants are used to distinguish the different vowels. Higher formants, on the other hand, determine the timbre of a voice [2].

Since the acoustic properties of the vocal tract are determined by its geometry, geometric data based on MRI images are used for the characterization of the vocal tract [3, 4, 5]. For simplification, the vocal tract is often expressed by an area function that describes the cross-sections along the centerline of the vocal tract [3]. This simplification is valid from an acoustic point of view, since the sound propagation up to a frequency 5 kHz can be regarded as one-dimensional [6].

As the geometry is obtained by imaging methods, previous studies are based on the data of individual subjects. In order to investigate the influence of the vocal tract on the fluid-structure-acoustic interaction of phonation and the radiation of sound, however, a more general model, which does not show strong individual influences, is

necessary. Hence, the aim of the study is to create a simplified model of the vocal tract that shows preferably little individual but the typical acoustic characteristics of the specific vowel. This is achieved by averaging over six simplified individual vocal tract geometries based on MRI images of professional tenors.
In order to consider the influence of the flow that is present in human phonation, the vocal tract models are investigated in an experimental setup with artificial vocal folds.

2 METHODS

2.1 Vocal tract models

The vocal tract geometries used are based on MRI images of professional tenors taken at the Institute Musicians' Medicine of the University Hospital Freiburg by Echternach et al. [4, 7]. Based on these data, a vocal tract model with realistic geometry and one with simplified form was created for each of the six tenors. The realistic geometries are generated by a manual segmentation of the MRT data in the open source software *3D Slicer*. The simplified models are based on the area functions of the tenors by Echternach et al. [7] and are represented by straight models with circular cross-sections.
In order to investigate the influence of the vocal tract on phonation and the underlying physical processes, a more general, simplified model of the vocal tract was created in addition to the individual vocal tract models. This averaged vocal tract geometry was calculated by averaging the area functions of the six tenor vocal tracts. As shown in figure 1, the vocal tracts were normalized to mean length and the cross-sectional areas were averaged.
The vocal tract models were produced using additive manufacturing methods.

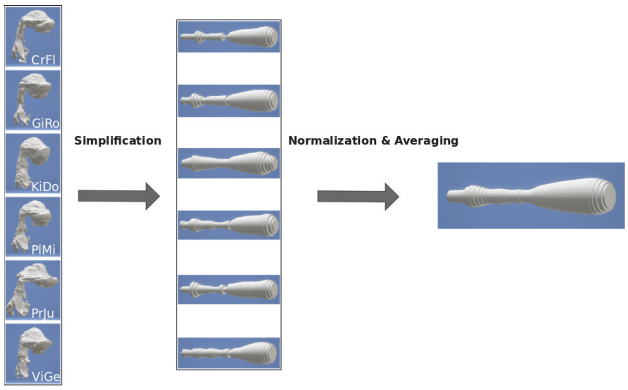

Figure 1. Method for averaging the vocal tract geometry

2.2 Experimental Setup

The investigation of the acoustic properties of the vocal tract is done in an experimental setup based on the setup by Kniesburges et al. [8, 9] and Lodermeyer et al. [10, 11], as depicted in figure 2.
It consists of a unit for providing and conditioning the airflow, an artificial larynx, a subglottal channel, and the vocal tract replica. The artificial larynx consists of an artificial pair of silicone vocal folds that are based on the M5 model by Scherer et al. [12] and Thomson et al. [13], from the silicone *Ecoflex 30* (Smooth-On, USA) and have a Young's modulus of E = 4,4 kPa.

2.3 Measurement setup and evaluation methods

The sound pressure measurements were done in an anechoic chamber with a 1/2” free field microphone *Type 4189* (Brüel & Kjaer, Denmark) placed at a distance of 90 cm and an angle of 45° from the exit of the vocal tract. The recorded signal was amplified with the *Nexus Conditioning Amplifier* (Brüel & Kjaer, Denmark) and sampled by the multifunctional module *PXIe-6356* (National Instruments, USA) with a sample rate of 44.1 kHz. The evaluation of the measurement data and the calculation of the sound pressure level of the acoustic pressure is done in *Matlab*. (Mathworks, USA). Based on the resulting spectra, the formant frequencies of the individual vocal tracts were detected using inverse filtering.

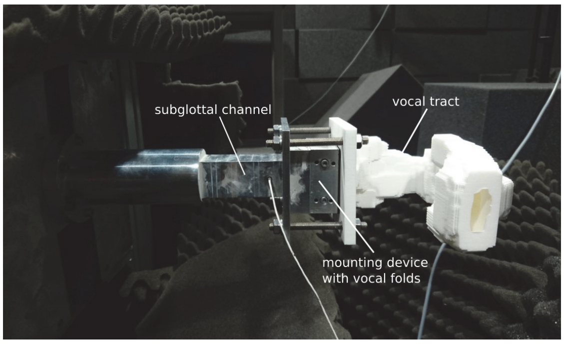

Figure 2. Experimental setup with subglottal channel, artificial vocal folds, and MRI-based vocal tract

3 Results and Discussion

3.1 Analysis of formant frequencies

Figure 3 shows the obtained spectra of the sound pressure level for the MRI-based and simplified models of the six subjects. Each spectrum shows the fundamental frequency, the harmonics and the broadband noise. The comparison between the spectra of the MRT-based and simplified models shows, however, that the broadband sound differs between the MRI-based and the simplified geometry. These differences indicate that the flow field and the resulting flow sound in the vocal tract are changed by the simplification.

In addition, the formant frequencies, as detected by inverse filtering are marked with dashed lines in figure 3. For comparison the physiological formants that where determined from audio recordings of the subjects by Echternach et al. [7] are marked as well. The formant frequencies are depicted in figure 4, which shows the determined formant frequencies of the MRI-based and the simplified vocal tract model for each subject and of the mean vocal tract model compared to the physiological formants determined by Echternach et al. [7]. For all models four formant frequencies can be determined in the range between 50 Hz and 5 kHz. The mean deviations of the formant frequencies from the reference values for the MRI-based models are less than 18 % and less than 16 % for the simplified models.

However, as a vowel is not characterized by a discrete formant frequency, but varies for each individual [14], the vowel-specific filter properties of the vocal tract are reproduced with sufficient accuracy.

3.2 Analysis of the averaged vocal tract

To analyze the vocal intelligibility of the mean vocal tract, the first two formant frequencies of the averaged vocal tract are plotted in a formant chart, as presented by Peterson and Barney [14], together with the physiological values by Echternach et al. [7], see figure 5. The comparison shows that the formants of the averaged vocal tract differ by 10 % from the mean value of the first formants and by 3 % from the mean value of the second formants. Despite the deviation of 10 % for the first formant frequency it can be seen, that the

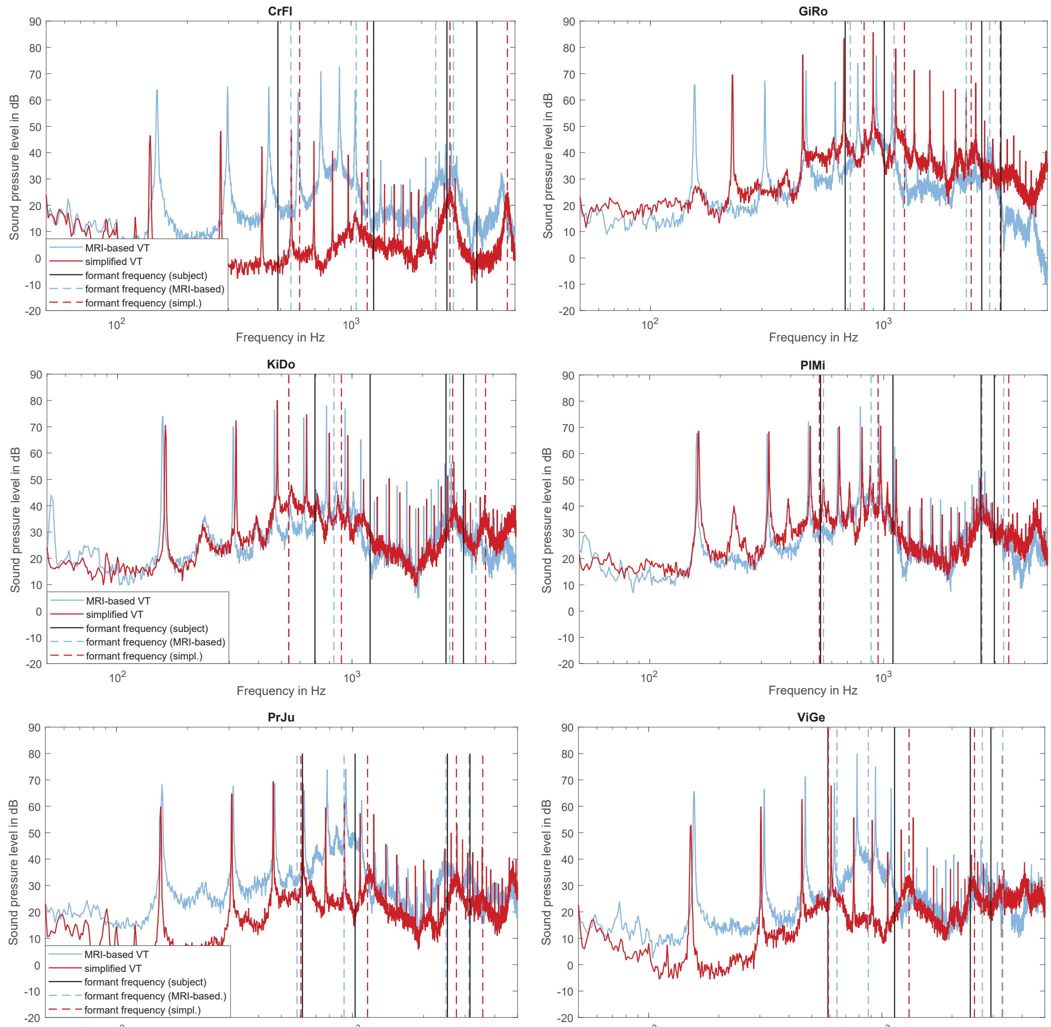

Figure 3. Sound pressure level of the MRI-based (blue) and simplified (red) models for all subjects. The formant frequencies, as detected by inverse filtering are marked with vertical lines

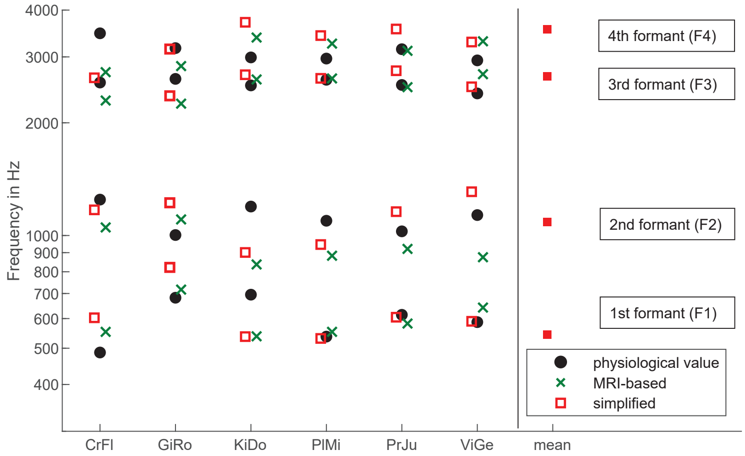

Figure 4. MRI-based and the simplified vocal tract model for each subject and of the mean vocal tract model

formants produced by the mean vocal tract are in the same frequency range as the subjects' formants. Due to the fact that a vowel is characterized by a frequency band, this shows that the vowel characteristics can be preserved. Hence, the averaging over area functions of the vocal tracts is a suitable procedure to reduce individual influences while preserving the essential acoustic properties of the vocal tract.

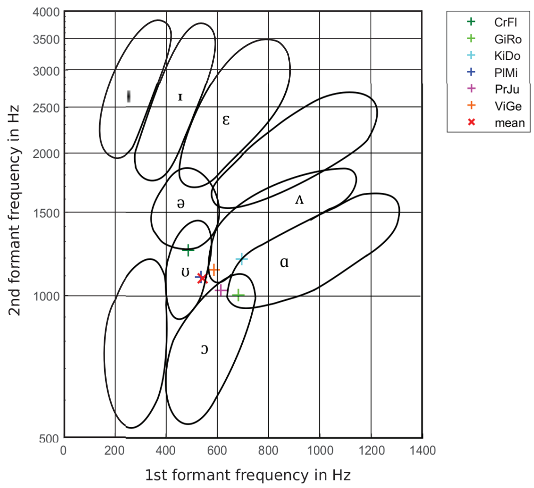

Figure 5. The formant chart, according to Peterson and Barney [14] shows the formant frequencies of the mean vocal tract compared to the physiological formants of the tenors.

4 CONCLUSIONS

In this study, the acoustic properties of the vocal tract were investigated in an experimental setup with an artificial larynx and MRI-based and simplified vocal tract models, respectively. In addition a simplified mean

vocal tract was generated to reduce the individual influences. The comparison with the tenors' vocal recordings show that the vowel characteristics can be reproduced with all vocal tract models. However, contrary to the current literature, a change in the broadband sound level due to the simplification of the geometry is detected if the oscillation of the vocal folds and the resulting flow are taken into account. This indicates that the simplification of the geometry leads to a change of the flow field in the vocal tract and the resulting flow-induced sound.
The validation of the mean vocal tract model based on the physiological reference values shows that the acoustic properties and the given vowel sound are reproduced. Hence, the average vocal tract model presented here is suitable for future investigations of the influence of the vocal tract on phonation.

ACKNOWLEDGEMENTS

This research was funded by Else Kröner-Fresenius Stiftung under grant agreement number 2016 A78. The authors also gratefully acknowledge the support of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative.

REFERENCES

[1] I.R. Titze: Principles of voice production. Prentice Hall, 1994

[2] Johan Sundberg. Vocal tract resonance in singing. The NATS Journal, 44:11-19, 1988.

[3] Story, B.H.; Titze, I.R.; Hoffman, E.A. Vocal tract area functions from magnetic resonance imaging. The Journal of the Acoustical Society of America 1996, 100, 537–554.

[4] Echternach, M.; Sundberg, J.; Markl, M.; Richter, B. Professional opera tenors' vocal tract configurations in registers. Folia Phoniatrica et Logopaedica 2010, 62, 278–287.

[5] Aalto, D.; Aaltonen, O.; Happonen, R.P.; Jääsaari, P.; Kivelä, A.; Kuortti, J.; Luukinen, J.M.; Malinen, J.; Murtola, T.; Parkkola, R.; others. Large scale data acquisition of simultaneous MRI and speech. Applied Acoustics 2014, 83, 64–75.

[6] Arnela, M.; Dabbaghchian, S.; Blandin, R.; Guasch, O.; Engwall, O.; Van Hirtum, A.; Pelorson, X. Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds. The Journal of the Acoustical Society of America 2016, 140, 1707–1718.

[7] Echternach, M.; Sundberg, J.; Baumann, T.; Markl, M.; Richter, B. Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers. The Journal of the Acoustical Society of America 2011, 129, 3955–3963.

[8] Kniesburges, S.; Hesselmann, C.; Becker, S.; Schlücker, E.; Döllinger,M. Influence of vortical flow structures on the glottal jet location in the supraglottal region. Journal of Voice 2013, 27, 531–544.

[9] Kniesburges, S.; Lodermeyer, A.; Becker, S.; Traxdorf, M.; Döllinger, M. The mechanisms of subharmonic tone generation in a synthetic larynx model. The Journal of the Acoustical Society of America 2016, 139, 3182–3192.

[10] Lodermeyer, A.; Becker, S.; Döllinger, M.; Kniesburges, S. Phase-locked flow field analysis in a synthetic human larynx model. Experiments in Fluids 2015, 56, 77.

[11] Lodermeyer, A.; Tautz, M.; Becker, S.; Döllinger, M.; Birk, V.; Kniesburges, S. Aeroacoustic analysis of the human phonation process based on a hybrid acoustic PIV approach. Experiments in Fluids 2018, 59, 13.

[12] Scherer, R.C.; Shinwari, D.; De Witt, K.J.; Zhang, C.; Kucinschi, B.R.; Afjeh, A.A. Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. The Journal of the Acoustical Society of America 2001, 109, 1616–1630.

[13] Thomson, S.L.; Mongeau, L.; Frankel, S.H. Aerodynamic transfer of energy to the vocal folds. The Journal of the Acoustical Society of America 2005, 118, 1689–1700.

[14] Peterson, G.E.; Barney, H.L. Control methods used in a study of the vowels. The Journal of the acoustical society of America 1952, 24, 175–184.