ICVPB 2020

12th International Conference on Voice Physiology and Biomechanics
March 18-20, 2020, Grenoble, France

Towards an open-access database of 3D shapes of the vocal tract and their aero-acoustical properties

Mario Fleischer[1]*, Peter Birkholz[2]

[1]Department of Audiology and Phoniatrics, Charité – Universitätsmedizin Berlin, Berlin, Germany
[2]Institute of Acoustics and Speech Communication, Technische Universität Dresden, Dresden, Germany

Keywords: Vocal tract, Aero-Acoustics, Numerical Modelling, Speech dataset

Introduction

For a comprehensive and profound understanding of the origin and transmission of acoustic energy in the human vocal system, representative three-dimensional geometric descriptions of the vocal tract (VT) shape are essential.

Methods

We determined the air-filled cavities of the VT of one male and one female speaker by using 3D Magnetic Resonance Imaging. The subjects were asked to produce 22 German speech sounds each (16 vowels and 6 consonants). For all resulting 44 models, the teeth of the subjects were added to the raw data before segmentation.

Based on the processed 3D images, we built a complete set of volume meshes needed for numerical analyses and half-shell 3D-printable models (Figure 1).

All vocal tract models were acoustically characterized in terms of their transfer functions (0-10 kHz) that were determined both numerically with the 3D Finite Element Method, and experimentally by means of an external source excitation paradigm [1]. Furthermore, for the analyses of fricatives and in terms of human recognition of the acoustic outcome, all 3D-printed models were excited by using an artificial flow source at the glottis [2].

Results and Discussion

We found an excellent agreement (with an average difference of below 1%) of the experimentally and numerically determined frequency values of the first three vocal tract resonances (f_{R1}, f_{R2} & f_{R3}). Moreover, the first two resonances of the vowels were plausibly arranged in the f_{R2}-f_{R1}-plane and clearly separated from each other. That indicates that all vowels were produced differently from each other, despite the non-ideal recording conditions in the MRI scanner and the need to artificially sustain the vowels for 12 seconds.

We plan to make all raw data, models and results freely available on an open-access platform to the scientific community. The actual state of this project, including findings of flow-induced sound experiments, will be presented and discussed.

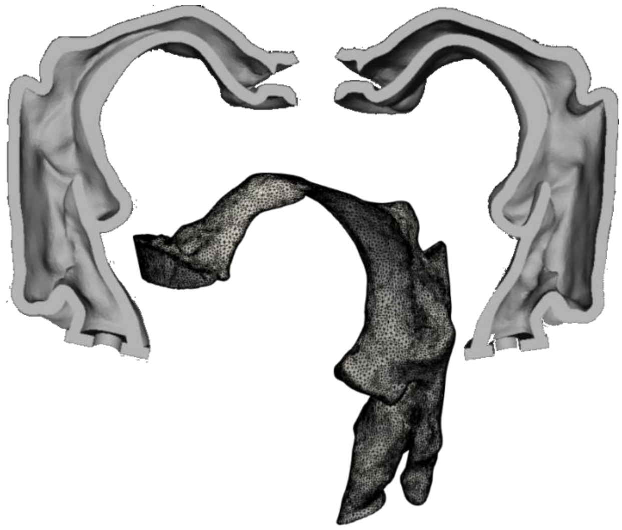

Figure 1: Example for a 3D-printable vocal tract model and the associated Finite Element Mesh

References

[1] Fleischer *et al*, PloS ONE, 13(3):e0193708, 2018.
[2] Birkholz *et al*, J Acoust Soc Am, 146(1):223-232, 2019.

*mario.fleischer@charite.de