Comparison of acoustic parameters of inhalations vs. exhalations with 3D-printed vocal tract models

UNIVERSITÄT DES SAARLANDES

DFG Deutsche Forschungsgemeinschaft

TECHNISCHE UNIVERSITÄT DRESDEN

Z A S

Raphael Werner[1], Susanne Fuchs[2], Jürgen Trouvain[1], Steffen Kürbis[3], Bernd Möbius[1], and Peter Birkholz[3]

[1]Language Science and Technology, Saarland University, Saarbrücken, Germany; [2]Laborphonologie, Leibniz Centre General Linguistics (ZAS), Berlin, Germany; [3]Institute of Acoustics and Speech Communication, TU Dresden, Dresden, Germany

1 Introduction

- inhalations differ from regular speech: no phonation, ingressive airflow
- inhalations showed similarities to some vowel formants and /k/-aspirations [1]
- acoustics of breath noises in speech under-researched
- how do in- and exhalation noises differ? details unknown so far

? What is the effect of reversing airflow direction on acoustic characteristics in the same vocal tract (VT)?

2 Methods

- 3D-printed vocal tract models (m, f) producing 8 sounds /aː, iː, uː, ə, x, ç, ʃ, s/ [2]
- imitate in- and exhalations: static airflow through glottis in 2 directions; 3 power levels; 10 s; recorded with microphone
- power spectral density for all 96 noises
- compared via Discrete Cosine Transform (DCT) 0-3 [3]
- `lme4` [4] for model fitting; `emmeans` [5] for pairwise post-hoc comparisons
- lmer(DCTi ~ direction * VTconfig + (1|speaker) + (1|condition)) with i being 0-3 (no interaction for DCT3)

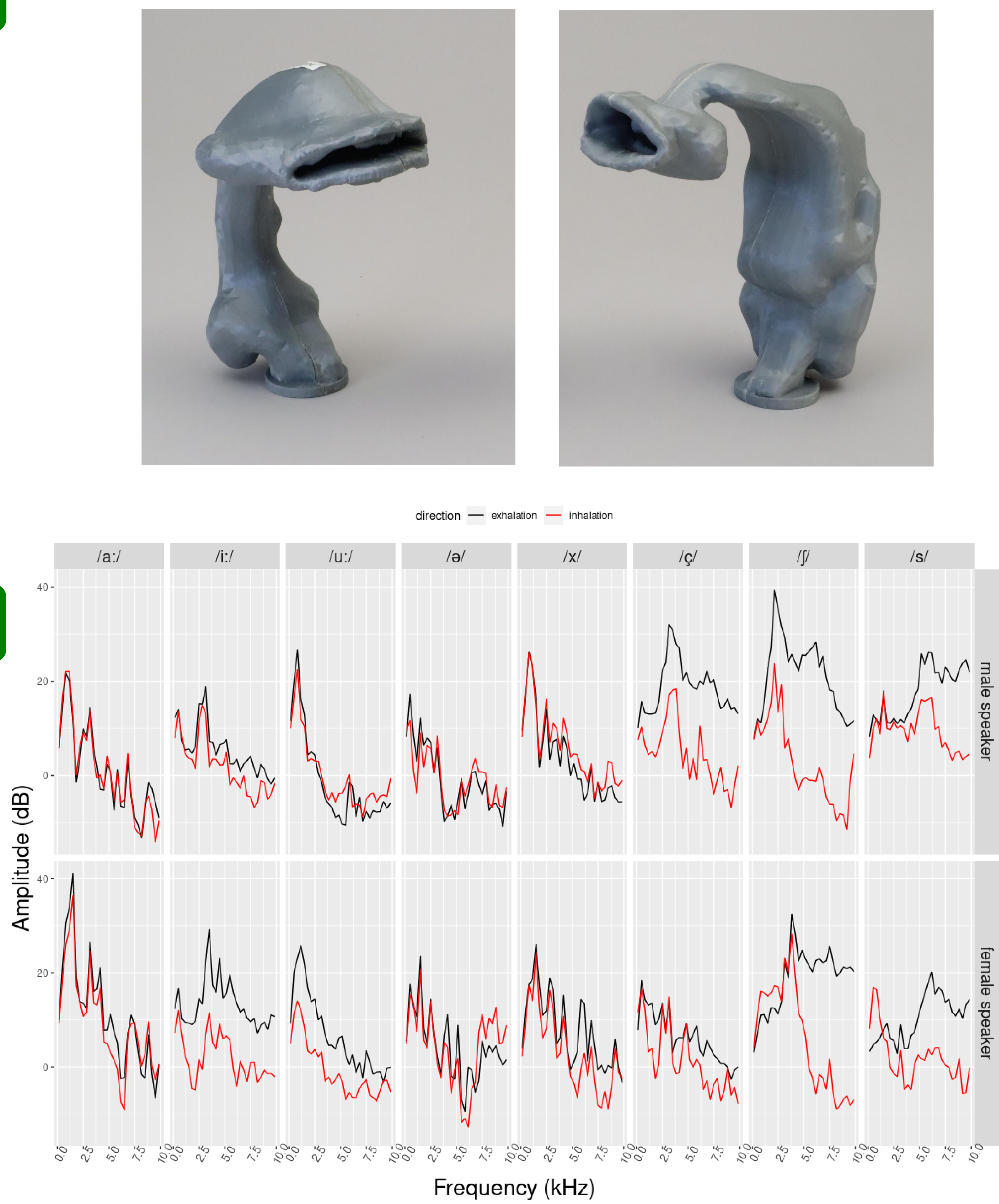

Figure: Top: two of the VT models corresponding to a male speaker producing the sounds /aː/ (left) and /ʃ/ (right); bottom: averaged spectra (0–10 kHz) for exhalation (black) and inhalation (red) by VT configuration and speaker (male and female).

3 Results

- no main effect of direction for any of the 4 statistical models
- post-hoc comparisons for significant direction contrasts by VT configuration:
  - DCT0: /iː, ç, ʃ, s/
  - DCT1: /ʃ, s/
  - DCT2: /ʃ/

4 Discussion & Conclusion

- no general effect of reversing airflow direction on spectrum, but specfic for VT config
- differences mostly found for sibilants (esp. /ʃ/) and for mean amplitude in configurations involving high tongue positions
- amplitude higher in 4 exhalations: concentrated airstream hitting incisors
- speakers/models differ for some VT config.
- implications for acoustic characterization of real inhalations: if VT relatively open there, direction not a problem for comparison with real speech sounds

5 References

[1] Werner et al. (2021). Inhalations in Speech: Acoustic and Physiological Characteristics, Interspeech, 3186–3190. **[2]** Birkholz et al. (2020). Printable 3D vocal tract shapes from MRI data and their acoustic and aerodynamic properties. Scientific Data, 7(1), 1–16. **[3]** Jannedy & Weirich (2017). Spectral moments vs discrete cosine transformation coefficients: Evaluation of acoustic measures distinguishing two merging German fricatives. The Journal of the Acoustical Society of America, 142(1), 395–405. **[4]** Bates et al. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. **[5]** Lenth (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. Version 1.6.1.

8th International Conference on Speech Motor Control – Groningen August 24–27, 2022