An MRI Examination of Lingual Fricatives in Upper Sorbian

Phil J. Howson[1] & Peter Birkholz[2]

[1]Leibniz-Zentrum Allgemeine Sprachwissenschaft; [2]Technische Universität Dresden
howson@leibniz-zas.de; peter.birkholz@tu-dresden.de

ABSTRACT

Fricatives require precise gestural control to produce a narrow constriction in the vocal tract. This generates turbulent airflow that gives fricatives their distinctive acoustic properties. This study investigated the articulation of the lingual fricatives /s, ʃ, x/ of Upper Sorbian, which is an endangered language spoken in eastern Germany. We used MRI to examine the mid-sagittal plane for the fricatives. We also used the MRI tracings to simulate the fricatives' COG in VocalTractLab. 2 male and 2 female L1 Upper Sorbian speakers participated in the study. We performed 3D vocal tract scans using 1.2 x 1.2 x 1.8 mm^3 voxels with 44 sagittal slices.

The results revealed significant inter-speaker variation in constriction location and length, and tongue shape. The results thus implicate variation in gestural representations that deviate due to vocal tract size and shape to meet a specific acoustic target.

Keywords: Fricatives; Upper Sorbian; MRI; phonetics; endangered languages

1. INTRODUCTION

Fricative sounds involve articulations that produce turbulent airflow within the vocal tract. In sibilants, such as /s/, the airflow leaving the critical constriction is directed against the incisors for the generation of a specific noise pattern. Fricatives in general demand a great deal of precision to accurately produce the required spectral qualities. Even one millimetre difference in constriction location can impact the acoustics significantly [1, 5]. This makes them hard to produce and even harder to acquire.

The lingual fricatives of Upper Sorbian have not been studied in significant detail with MRI or acoustic measures, but the language shares a similar inventory as German, /s, ʃ, x/, and the impact from language contact is also likely to have caused similarities between the languages [7]. With respect to German fricatives, Jannedy & Weirich [9] found that for Hamburg and Jena dialects of German, there were high frequency spectral peaks for /s/, typically over 6,000 Hz, while /ʃ/ typically had a peak around 3,000 Hz, and /x/ had a peak often below 2,000 Hz. The overall distribution of noise differed significantly by segment and dialect, however, often spanning a range of 2,000 or 3,000 Hz. Gordon, Barthmaier, & Sands [6] examined seven languages that each had rich fricative inventories. Across the languages examined, /s/ had COG ranges from 4,400 Hz – 5,600 Hz, /ʃ/ had COG ranges from 3,900 Hz – 5,100 Hz, and /x/ had COG ranges from 3,900 Hz – 4,500 Hz. The wide and overlapping ranges of COG for the relevant fricatives suggests that there are a variety of distinct vocal tract shapes that distinguish each of the fricatives.

The purpose of this study was to examine the vocal tract shapes of lingual fricatives in Upper Sorbian as part of a documentation effort. Additionally, we compared and simulated COG to better understand how vocal tract geometry affects the acoustics for fricatives.

2. METHODS

2.1. Participants

Four L1 speakers of Upper Sorbian participated in this study (2 male and 2 female). Participants were all college-educated, ages between 20-24 years, and had no self-reported history of speech or hearing disorders.

2.2 Stimuli & Procedure

Participants were given a list of words with target segments prior to data collection and were instructed to practice producing them for an extended duration. The stimuli are presented in Table 1.

Segment	Stimuli	English
/s/	**s**adło	fat; grease
/ʃ/	**š**ach	chess
/x/	či**ch**awa	sneezing

Table 1. Stimuli with target segments in bold.

3D images of the vocal tract were collected with a Siemens 3T Trio, with a pixel size of 1.2 mm x 1.2 mm, and a sagittal slice thickness of 1.8 mm. 44 slices were taken in total and they were used to generate a 3D image of the vocal tract. In order to record 44 slices, participants had to produce a sustained articulation of a single segment for 14 seconds. To do this, participants were prompted with a power point presentation. A mirror was present in the MRI to reflect the power point screen. First, the target segment plus the example word were presented. Participants were asked to produce the word and the target segment through the intercom system. Following confirmation that the correct segment was being articulated, participants were prompted with the target segment and example word and a bar on the screen that would indicate when to take a deep breath, when to begin articulation, and when to stop. After each articulation, MRI images were examined and in the case of blurriness or poor image quality, the participant repeated the trial until a clear, high-quality image was obtained.

Audio recordings were additionally performed in a quiet room in Leipzig, Germany. Data were collected using a Tascam Linear PCM recorder. Data were recorded at a sampling frequency of 48,000 Hz. Participants were sitting for audio recording, rather than laying as they had in the MRI. Participants read the target words used in the original MRI stimuli and then produced 5 second static production of each of the target segments /s, ʃ, x/. Participants were instructed to produce the segments as similarly to the MRI productions as they could.

2.2 Analysis

Dicom files were converted into 44 bitmap files for each segment produced by each of the participants. Bitmaps were loaded into Image3D [2]. Image3D uses Catmull-Rom Splines so users can trace the tongue, lips, palate, and pharyngeal wall. We generated 2 independent splines, one for the lower lip and tongue and one for the upper lip, palate and pharyngeal wall. Splines were then exported to a scalable vector graphics (SVG) file and imported to Inkscape [8]. Midsagittal contours were then overlayed on top of each other to facilitate comparisons.

Acoustic analysis was performed in Praat [4]. We took a 30-millisecond window centred on the midpoint of the static articulation of each fricative. A spectral slice was extracted and then the Center of Gravity (COG) was calculated. Then VocalTractLab (Special Version for 3D acoustic simulations) [3] was used to simulate tongue shapes and acoustic outputs. The 3D acoustic simulation suite was used. VocalTractLab [3] provides control points to manipulate the 2D and 3D tongue contours. It can then be used to compute transfer functions, modes, and the acoustic field in order to simulate the acoustics of the vocal tract shape. We manipulated the control points and tongue side elevation to match the MRI tracings.

3. RESULTS

3.1 MRI Results

The midsagittal traces for speaker US01 are presented in Figure 1. The midsagittal profile indicated a constriction along the alveolar ridge for both /s/ and /ʃ/, which was around the postalveolar region. However, there was a longer constriction for /s/, which extended further forward towards the lower incisors. Additionally, there was more tongue tip advancement compared to /ʃ/. The tongue body was more raised and had a downward slope for /ʃ/, indicating a slight degree of palatalization. /ʃ/ also had a sublingual cavity. The posterior tongue was also more advanced for /ʃ/, compared to /s/. /x/ had a constriction in the midpalate, accompanied by a retracted tongue tip and tongue dorsum.

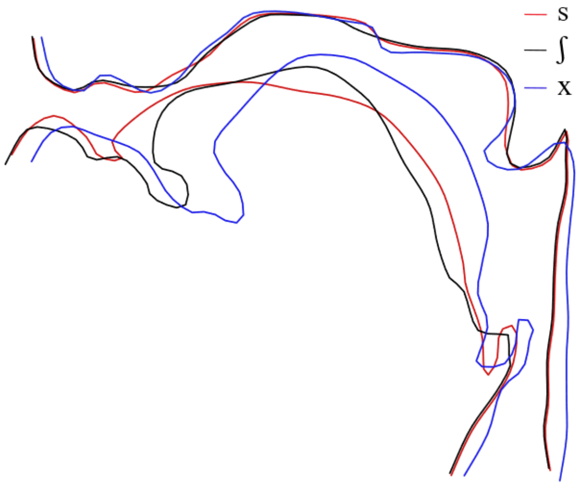

Figure 1: MRI traces of the midsagittal plane for lingual fricatives, /s/ (red), /ʃ/ (black), /x/ (blue) for US01.

The midsagittal traces for speaker US02 are presented in Figure 2.

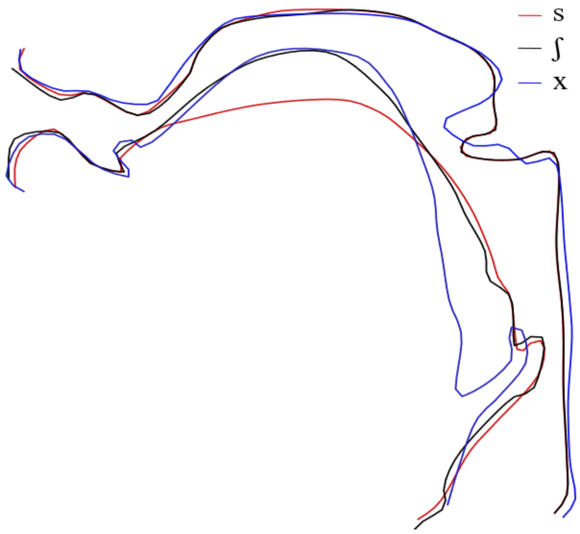

Figure 2: MRI traces of the midsagittal plane for lingual fricatives, /s/ (red), /ʃ/ (black), /x/ (blue) for US02

The constriction location for /s/ and /ʃ/ were similar. However, in the case of US02, the constriction for /ʃ/ was significantly longer than for /s/, spanning the entire post-alveolar ridge. /s/, on the other hand, had a short constriction and the most anterior portion of the alveolar ridge. /s/ had a low tongue body with a retracted tongue dorsum. /ʃ/ had a high degree of palatalization, exhibiting a much narrower constriction in the midpalate. /x/ had a similar contour to /ʃ/. The major differences included a slight curling of the tongue tip and a much more advanced tongue dorsum. The posterior tongue body was also slightly more retracted and raised. This suggested a similar type of constriction for /x/ and /ʃ/, but with marginal differences that impacted the acoustics significantly.

Figure 3 presents the midsagittal tracings for US03.

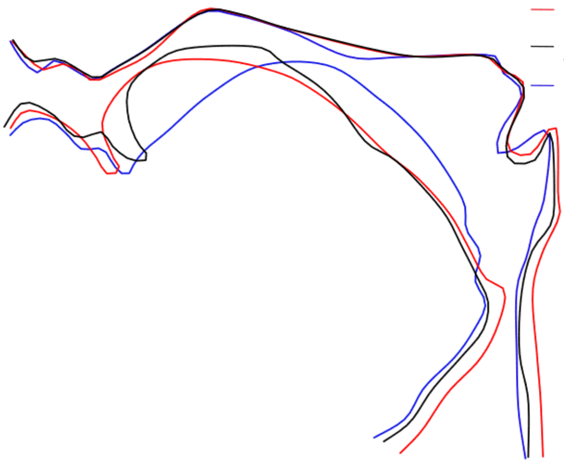

Figure 3: MRI traces of the midsagittal plane for lingual fricatives, /s/ (red), /ʃ/ (black), /x/ (blue) for US03.

Similar to US01 and US02, /s/ and /ʃ/ had overlapping constriction locations along the alveolar ridge, but /s/ had a more advanced constriction. /ʃ/ had a small, but noticeable sublingual cavity and an advanced tongue dorsum, while the tongue tip for /s/ was advanced towards the lower incisors. The tongue body for /ʃ/ was slightly raised along the alveolar ridge and hard palate, also indicating a small degree of palatalization. /x/ had a more retracted tongue dorsum and a constriction location along the soft palate. Additionally, for US03, /x/ did not have any observable sublingual cavity.

Figure 4 presents the MRI traces for participant US04. /s/ had a more advanced constriction right at the most anterior region of the alveolar ridge, compared to /ʃ/, which had a longer constriction along the posterior regions of the alveolar ridge. For both /s/ and /ʃ/, there was extensive tongue body raising, although for /s/ the raising was more in the mid-palate region and was achieved with the posterior tongue body. /ʃ/ had more anterior raising of the tongue body, congruent with the longer constriction along the posterior alveolar ridge. /x/ was produced with a long constriction along the mid-palate that extended towards the anterior portions of the hard palate. We also observed significant posterior tongue body and dorsum retraction into the velar and pharyngeal regions.

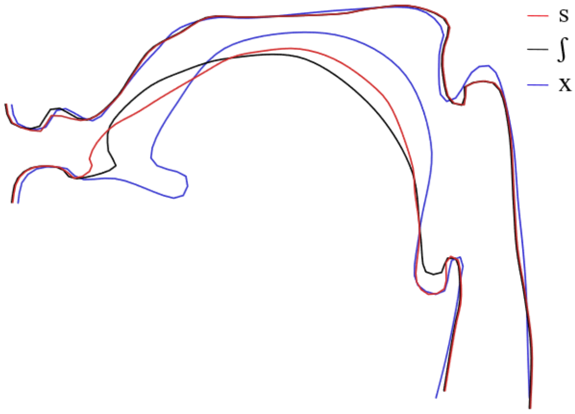

Figure 4: MRI traces of the midsagittal plane for lingual fricatives, /s/ (red), /ʃ/ (black), /x/ (blue) for US04.

3.2 Simulation Results

Table 2 presents the COG for each segment as produced by each speaker as extracted from the Praat analysis.

	/s/	/ʃ/	/x/
US1	8274	3711	1471
US2	9560	3328	4848
US3	7610	3725	3024
US4	8322	3589	2116

Table 2. COG (in Hz) for each of the segments, /s, ʃ, x/, for each of the participants US01, US02, US03, US04.

Table 3 presents the COG of the simulated data for each tongue shape for each segment using VocalTractLab. Figure 5 below presents an example simulation for US1 /ʃ/ with overlapped MRI tracing.

	/s/	/ʃ/	/x/
US1	6440	3749	1876
US2	6339	3456	4519
US3	6519	3966	3323
US4	5477	3282	2215

Table 3. Simulated COG (in Hz) for each of the segments, /s, ʃ, x/, based on MRI tongue contours for each of the participants US01, US02, US03, US04.

4. DISCUSSION

The MRI results revealed wide variation in the articulation of /s, ʃ, x/ across all speakers. Relative place of articulation across segments was consistent – that being that /s/ was most anterior, followed by /ʃ/, and finally /x/ – but that the precise location and length of the constriction along the palate varied by speaker. Significant overlap in constriction location was also observed between /s, ʃ/, implicating the shape of the constriction (i.e., grooved vs. domed) as an important predictor of acoustics [14]. The front cavity geometry also varied significantly with respect to length, presence and size of sublingual cavity, and anterior tongue shape, which all may be relevant in the case of sibilants in order to direct airflow at the incisors [10, 11, 12]. Additionally, laminality or apicality was found to vary by speaker [11]. Differences in the anterior cavity for /x/ are likely related to retraction of the tongue body and dorsum to produce a constriction in the posterior palate. One limitation is that the teeth were not visible during the MRI, which play an important role in fricative acoustics [1, 5, 13].

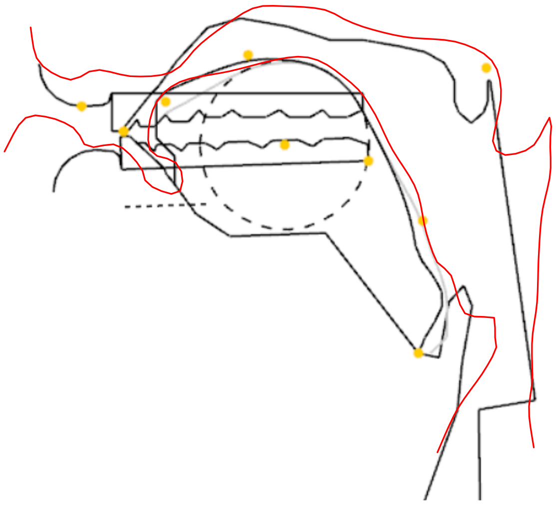

Figure 5: Example vocal tract geometry for simulation of US1's articulation of /ʃ/ (black) with MRI of their vocal tract (red).

Simulations with VocalTractLab implicated specific regions of importance for differentiation between /s, ʃ/. Overlapping constrictions with different degrees of anterior tongue body raising played an important role in lowering COG to resemble the modelled /ʃ/. Doming or grooving of the tongue for either of /s, ʃ/ did impact COG simulations, but tongue posture and constriction location and length differences for /s, ʃ/ might implicate a biomechanical necessity rather than specific achievement of an acoustic target.

With respect to /x/, we observed that the posterior cavity played a significant role in the simulated COG, which may be due to the effects on preconstriction airflow. One limitation of the simulations was differences in the shape of the alveolar ridge and hard and soft palate for the model compared to each individual participant.

Taken together, we suggest that variation in articulation is not due simply to speaker specific vocal tract geometry. Rather, it seems that specific acoustic targets can be achieved with significant differences in place of articulation, and the front and back cavity geometry.

5. ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 101018840. We would also like to thank the Institut für Sorabistik at Universität Leipzig for their help scheduling participants.

6. REFERENCES

[1] Badin, P. (1989). Acoustics of voiceless fricatives: Production theory and data. Speech Technol. Lett, 45-52.

[2] Birkholz, P. (2023). *Image 3D*. https://www.vocaltractlab.de/index.php?page=image3d-download.

[3] Birkholz, P. (2013). Modeling consonant-vowel coarticulation for speech synthesis. *PLoS ONE*, 8(4): e60603.

[4] Boersma, P. & Weenink, D. (2022). *Praat: doing phonetics by computer* [Computer program]. Version 6.3.03.

[5] Fujiso, Y., Nozaki, K., & Van Hirtum, A. (2015). Estimation of minimum oral tract constriction area in sibilant fricatives in aerodynamic data. *Journal of the Acoustical society of America* EL20-EL25.

[6] Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 142-174.

[7] Howson, P. (2017). Upper Sorbian. *Journal of the International Phonetic Association* 47, 1-9.

[8] Inkscape Project. (2020). *Inkscape*. Retrieved from https://inkscape.org.

[9] Jannedy, S. & Weirich, M. (2016). The acoustics of fricative contrasts in two German dialects. In Draxler C. & Kleber, F. (eds.), *Proceedings of Phonetics and Phonology in German Speaking Areas*, pp. 70-73. Ludwig-Maximilians-Universität München: Munich, Germany.

[10] McGowan, R. S. & Nittrouer, S. (1988). Differences in fricative production between children and adults: Evidence from an acoustic analysis of /ʃ/ and /s/. *Journal of the Acoustical Society of America* 229-236.

[11] Narayanan, S., Alwan, A., & Haker, K. (1995). An articulatory study of fricative consonants using magnetic resonance imaging. *Journal of the Acoustical Society of America*, 1325-1437.

[12] Shadle, C. H. (1990). Articulatory-acoustic relationships in fricative consonants. In Hardcastle, W. & Marchal, A. (eds.), *Speech Production and Speech Modelling*, pp. 187-209. Kluwer Academic: New York.

[13] Shadle, C. H. (1991). The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, 19(3-4), 409-424.

[14] Shadle, C. H., Berezina, M., Proctor, M., & Iskarous, K. (2008). Mechanical Models of Fricatives Based on MRI-derived Vocal Tract Shapes. *8th International Seminar on Speech Production*, 417-420.