Biplane ultrasound imaging of lingual grooving and sagittal displacement during the production of three sibilant sounds

Tim Bressmann[1, 2], *Marc Yarascavitch*[2], *Siew-Ging Gong*[2], *Bryan Tompson*[2]

[1] Department of Speech-Language Pathology, University of Toronto, Canada
[2] Faculty of Dentistry, University of Toronto, Canada

tim.bressmann@utoronto.ca, marc.yarascavitch@gmail.com, sg.Gong@dentistry.utoronto.ca, Bryan.Tompson@dentistry.utoronto.ca

Background: Biplane ultrasound imaging of the tongue allows to display the coronal and sagittal slices of the tongue concurrently. The goal of the current study was to investigate differences in coronal grooving and sagittal displacement of the tongue in the sibilant sounds /s/, /z/ and /ʃ/. The research was exploratory and not guided by specific hypotheses.

Methods: Three females and 3 males (mean age 29 years) participated in the study. The participants produced 5 repetitions of VCV syllables with the vowels /a/, /i/ and /u/ and the consonant sounds /s/, /z/, and /ʃ/. The participants sat with their forehead against a drum practice pad (RF-12G, HQ Percussion, Farmingdale, NY) on a cymbal stand (Pearl BC 1030, Pearl Cooperation, Nashville, TN). The participants' chin rested on an ultrasound transducer in a custom-made holder. A Philips X3-1 matrix array transducer with a center frequency of 5 MHz connected to a Philips iU22 workstation was used for the recordings (Philips Medical, Markham, ON). The video stream was recorded using a VRD-MC6 recorder (Sony Canada, Toronto, ON). The acoustic signal during the session was recorded to the same digital video disc using the audio output from a Q3 video recorder (Zoom, Tokyo, Japan).

The ultrasound movies were segmented with Screenblast 3.0 (Sony Canada, Toronto, ON). The tongue surface was traced using a Bamboo tablet (Wacom Technology, Vancouver, WA) and the Ultra-CATS software. The resulting data described the distance between the anchor point and the tongue surface in mm along radiating gridlines in 5° intervals in the sagittal and coronal planes. Concavity of the coronal tongue was calculated by subtracting the center measurement angle from the average of 15° right and left. (Bressmann et al., 2010). Values under -2 were interpreted as convex, values between -2 and 2 as flat, and values over 2 as concave. Cumulative sagittal movement at different measurement angles was calculated by smoothing the data with the 3RSSHS algorithm and then adding up the absolute differences between consecutive frames.

Results: To assess measurement reliability, two videos in a similar data set were retraced. Mean errors were 0.6 mm in the coronal and 0.95 mm in the sagittal plane. A mixed model ANOVA of the concavity scores was calculated with fricative sound, vowel sound and syllable position as the fixed effects and speaker as the random effect. The results showed significant differences between the fricative sounds ($F=6.154$, $df=2$, $p<0.01$), the vowel type ($F=40.718$, $df=2$, $p<0.01$) and syllable position ($F=9.313$, $df=2$, $p<0.01$). The fricative by vowel interaction was not significant.

The post-hoc Bonferroni tests for the concavity scores of fricative sounds showed a trend towards a statistically significant difference between /s/ and /z/ ($p<0.09$) and a significant difference between /s/ and /ʃ/ ($p<0.05$). The post-hoc Bonferroni tests for the vowels showed significant differences in concavity between all three vowels (all $p<0.01$). The post-hoc Bonferroni tests for the position showed a significant drop in concavity from the first to the second vowel ($p<0.05$) and from the consonant to the second vowel ($p<0.01$).

A mixed model ANOVA of the cumulative anterior displacement was calculated with fricative sound and vowel sound as the fixed effects and speaker as the random effect. The results showed significant differences between the fricative sounds ($F=6.4378$, $df=2$, $p<0.01$), as well as the vowel type ($F=19.564$, $df=2$, $p<0.01$). The fricative by vowel interaction was significant ($F=4.444$, $df=4$, $p<0.01$).

The post-hoc Bonferroni tests for the cumulative distances for the fricative sounds showed statistically significant differences between /s/ and /z/ ($p<0.01$) and between /z/ and /ʃ/ ($p<0.05$). The post-hoc Bonferroni tests for the vowels showed significant differences in sagittal displacement between all three vowels (all $p<0.05$). The interaction effect between fricatives and vowels was explained by significant differences between the vowels /a/ and /i/ in the context of /s/, the vowels /a/ and /u/ in the context of /z/ and the vowels /i/ and /u/ in the context of /z/. There were also significant differences between the fricatives /z/ and /ʃ/ in the context of /a/ and /s/ and /ʃ/ in the context of /i/.

Discussion: The goal of the present study was to assess the tongue shape during the production of the English sibilant sounds /s/, /z/ and /ʃ/ using a biplane ultrasound transducer. The results demonstrated differences in the magnitude of the lingual concavity and the cumulative displacement of the tongue.

Concavity scores were higher for /s/ than for the voiced cognate /z/. This finding may potentially be explained by the aerodynamic requirements of the voiceless fricative, which requires more airflow to produce a salient sibilant sound. Since the shape of the alveolar ridge is constant, this increased airflow may be achieved with a more concave tongue surface. Concavity was even lower for /ʃ/ but this sound could not be compared to its voiced cognate.

The degree of grooving of the tongue dorsum was influenced by the vowel context. It is likely that the more forward and raised tongue position for /i/ was the cause of the deeper midlingual groove because more genioglossus contraction is required to advance the tongue. The concavity measures for /u/ were lower.

The cumulative sagittal displacement measures for the VCV syllables were highest for the context of /u/, indicating that there was more displacement required to produce anterior consonants with this posterior vowel. Displacement was the lowest for /a/ with a relatively flat sagittal tongue shape for /s/ and /z/. However, to produce the /ʃ/ sound, the tongue moved more in the vowel context of /a/ than of /i/.

As a limitation of the research, the speakers' preferences for tongue tip up and tongue tip down /s/ were not taken into account. They were not asked about their preference, and it was not possible to discern possible differences from the ultrasound image.

Biplane ultrasound is still a relatively new tool for speech research. The measurement of data takes longer and is more difficult than in standard B-mode ultrasound. Like in all ultrasound investigations, it is difficult to stabilize the transducer, the image and the speaker. Nevertheless, the technique holds great promise. It would be of particular interest how it could be applied to clinical populations.

References

Bressmann, T., Flowers, H., Wong, W., Irish, J.C. (2010). Coronal view ultrasound imaging of movement in different segments of the tongue during paced recital: Findings from four normal speakers and a speaker with partial glossectomy. Clinical Linguistics and Phonetics, 24(8), 589-601.

Keywords:
- Sibilants; 3D ultrasound; Methodological research