Quantifying Tongue Tip Visibility in Ultrasound Images of /r/ Tongue Shapes Using Numerical Ultrasound Simulations

Sarah R. Li,[1] T. Douglas Mast, PhD,[1] & Suzanne Boyce, PhD CCC-SLP[2]
[1]Biomedical Engineering; [2]Communication Sciences and Disorders | University of Cincinnati

University of CINCINNATI

Introduction

- **Midsagittal ultrasound images** show the tongue surface from much of the root to the tip in real time, providing useful tongue shape information (e.g., for ultrasound biofeedback therapy (UBT) [1]).
- However, the **tongue tip may be obscured** by shadowing from the sublingual air space and/or mandible bone [2]. The amount of anterior tongue missing from the image is often not thoroughly understood.
  - This understanding may be of interest for American English /r/ due to the range of tongue shapes (i.e., the **bunched/retroflex** continuum [3, 4]), which is often split into categories for analysis [5].
- **Magnetic resonance images (MRI)** show the entire vocal tract and can thus be used to understand whether tongue shapes are miscategorized.
  - However, possible differences between tongue shapes in MRI and ultrasound [6] cause ambiguity (e.g., whether the difference is due to a difference in tongue shape or from ultrasound shadowing).
- A possible method to avoid ambiguity of tongue shape is to **generate ultrasound images via numerical simulations** of acoustic wave propagation [7], replicating scan lines recorded using ultrasound probes.
  - Tongue shapes segmented from MRI are used in ultrasound simulations to quantify the extent of tongue surface not visualized.

Hypothesis

The tongue tip is less visible for ultrasound images of retroflex /r/ than for bunched /r/ because of the larger front air cavity for retroflex shapes [8].
- This would be indicated by greater extent of missing anterior tongue contour in simulated ultrasound images, when compared to MRI.

Methods: Data

Data simulation

- **Simulated ultrasound**:

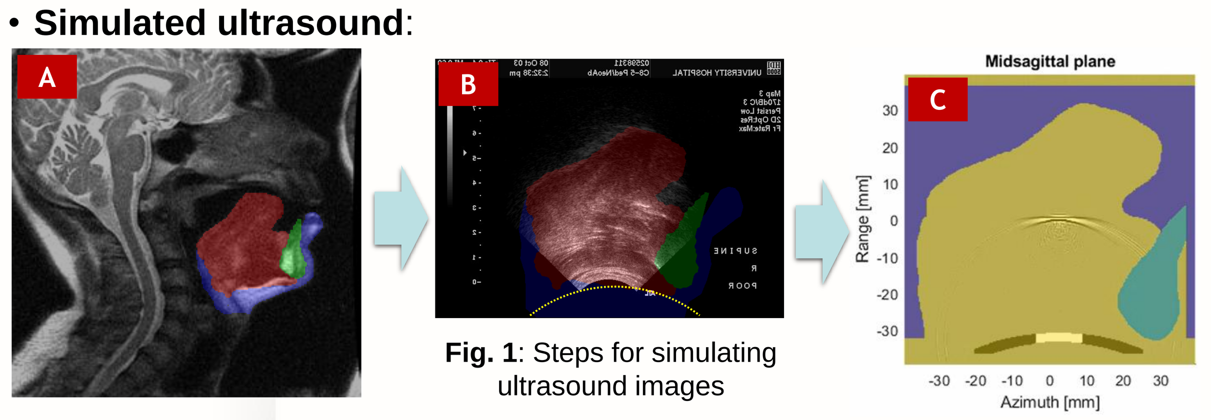

Fig. 1: Steps for simulating ultrasound images

SCAN ME

Link: https://github.com/SarahRLi/asa-may-2023

Step 1: (Fig. 1A) Tissue masks (tongue, mandible bone / teeth, and other soft tissue shown as red, green, and blue respectively) were manually segmented from MRI, completed for three sagittal slices from each speaker.
Step 2: (Fig. 1B) MRI were manually registered with measured ultrasound images (from the same speaker) to orient the tissue masks. The dashed yellow line illustrates the surface of the ultrasound transducer modeled in the simulation (may differ from measured).
Step 3: (Fig. 1C) Tissue masks were concatenated, smoothed, and then used to assign a map of acoustic properties for simulation. The k-Wave toolbox [7] was used to simulate acoustic wave propagation in 3D. Adjustments were made to maintain computational efficiency (see **5aBAb16**). This panel is a still frame from a video (see QR code) showing the acoustic pressure wavefront (currently located around 2 mm azimuth and 0 mm depth) for one scan line, with the transducer shown as the arc at bottom (active elements shown as bright yellow). The resulting simulated images were bandpass filtered at 4 MHz (Gaussian, 20-30% bandwidth).

Datasets used

- Speakers: Typical speakers of a rhotic American English dialect.
  - Adult dataset: 20 speakers
  - Child dataset: 3 speakers (aged 10-13)
- Stimuli and image parameters:
  - **MRI (for tongue tissue maps)**: in supine position; midline sagittal plane with two parasagittal slices, 240×240 mm^2 field of view (Fig. 2)

	Stimulus	*Slice thickness (mm)*	*Resolution (mm per pixel)*
Adult	Sustained /r/ in "pour" (~5 s)	5	0.938
Child	Sustained /r/ (~10 s for 7 frames)	3	1

  - **Measured ultrasound (for validating simulated ultrasound)**: midsagittal image of /r/ in "are" in upright position (child dataset) or sustained /r/ in "pour" in supine position (adult dataset)
    - Varying imaging configurations: curvilinear probes C5-2, C7-4, C8-5 on HDI 5000 machine or C6-2 (center frequency 4 MHz) on Siemens Acuson X300

Methods: Analysis

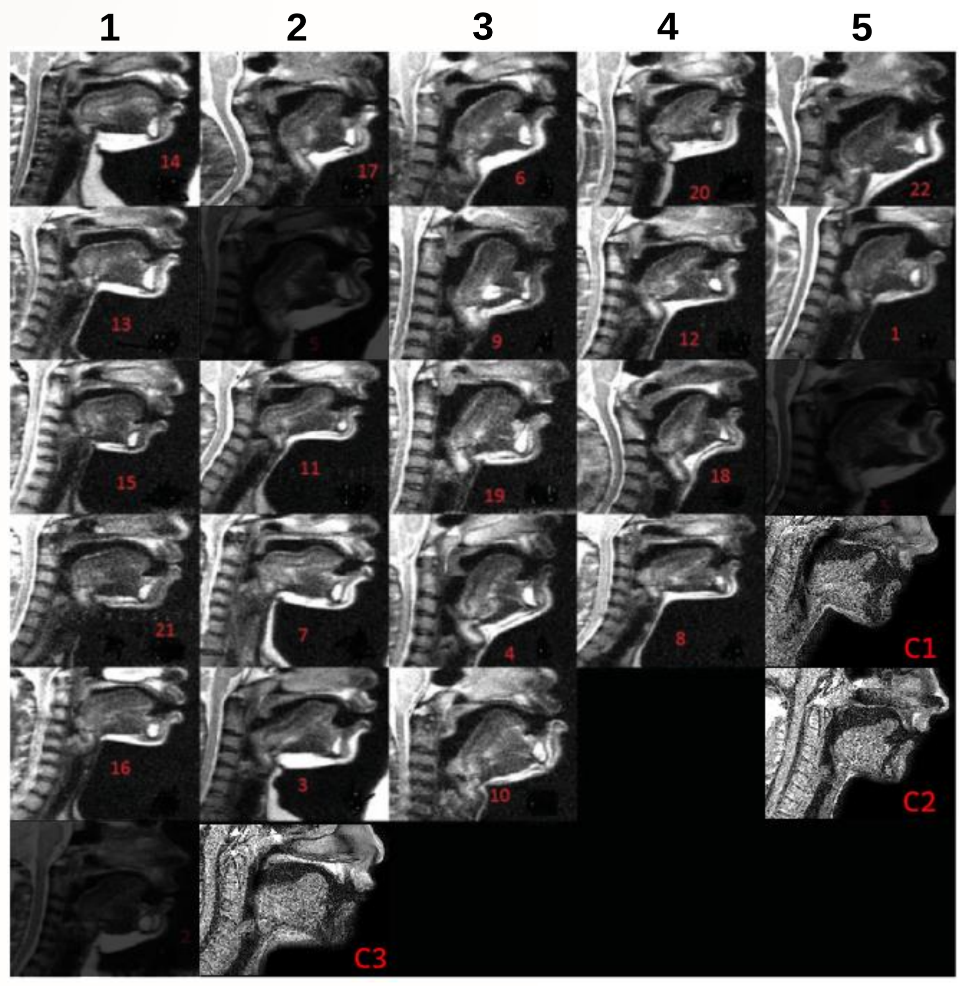

Fig. 2: Midsagittal MRI of speakers with columns indicating the manually assigned groups covering the range from most bunched (Group 1) to most retroflex (Group 5). Figure taken from [3], with additions of child speakers (C1, C2, C3) and with excluded adult speakers darkened.

Tongue contours

- Bunched/retroflex groups: A grouping of MRI tongue shapes into five categories (from [3, 4], shown as the columns of Fig. 2) across the bunched vs. retroflex continuum was used.
- Simulated ultrasound tongue contour: (Fig. 3, cyan line). Automatically found as local brightness maxima within vertical search windows (TonguePART [9]).
- MRI tongue contour: (Fig. 3, magenta line). Automatically found as the top vertical edge of the manually registered (Fig. 1) midsagittal tongue mask. Because this may include the anterior floor of the mouth (frenulum) below the tongue surface, parts of anterior contour with downward slope steeper than an empirical threshold were excluded.
- Calculation of missing anterior: The point (black star in Fig. 3) on the MRI contour closest to the anterior end of the simulated ultrasound contour was found. The length of the MRI tongue contour anterior to this point (Fig. 3, purple circles) was used as the length of missing contour, and the missing proportion was calculated as its ratio to the length of the entire MRI contour.

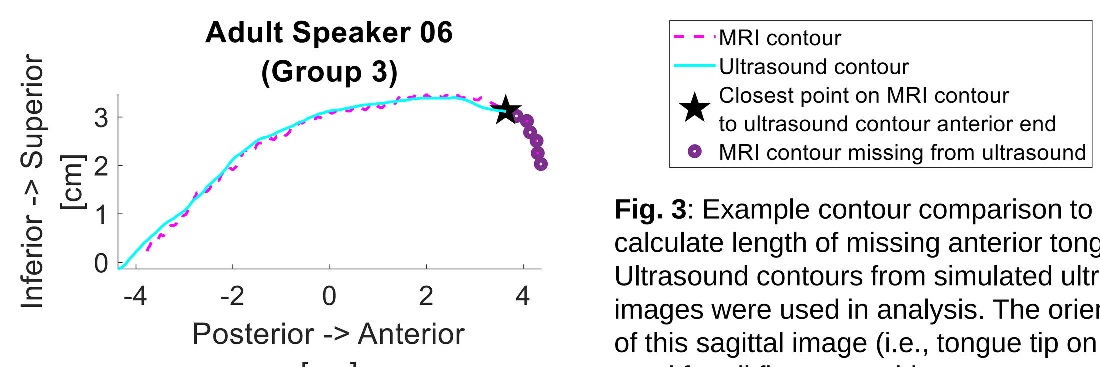

Fig. 3: Example contour comparison to calculate length of missing anterior tongue. Ultrasound contours from simulated ultrasound images were used in analysis. The orientation of this sagittal image (i.e., tongue tip on right) is used for all figures on this poster.

Acknowledgements

- NIH/NIDCD grants F31 DC020672, R01 DC05250, 8 UL1 TR000077-05, R01 DC013668, and R01 DC017301
- Additional mentors Dr. Steven Lulich and Dr. Shrikanth Narayanan for discussions on research project, as well as Dr. Maureen Stone for discussion on initial project idea.
- Siemens Medical Solutions for lending the Acuson X300 ultrasound scanner.

References

[1] E. Sugden, S. Lloyd, J. Lam, and J. Cleland, "Systematic review of ultrasound visual biofeedback in intervention for speech sound disorders," *Int J Lang Commun*, vol. 54, no. 5, pp. 705–728, Sep. 2019, doi: 10.1111/1460-6984.12478.
[2] M. Stone, "A guide to analysing tongue motion from ultrasound images," *Clin Linguist Phon*, vol. 19, no. 6–7, pp. 455–501, Jan. 2005, doi: 10.1080/02699200500113558.
[3] S. Boyce, "The Articulatory Phonetics of /r/ for Residual Speech Errors," *Semin Speech Lang*, vol. 36, no. 04, pp. 257–270, Oct. 2015, doi: 10.1055/s-0035-1562909.
[4] M. K. Tiede, S. E. Boyce, C. K. Holland, and K. A. Choe, "A new taxonomy of American English /r/ using MRI and ultrasound," *J Acoust Soc Am*, vol. 115, no. 5_Supplement, pp. 2633–2634, May 2004, doi: 10.1121/1.4784878.
[5] H. King and E. Ferragne, "Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English," *J Phon*, vol. 80, p. 100978, May 2020, doi: 10.1016/j.wocn.2020.100978.
[6] M. Stone *et al.*, "Comparison of speech production in upright and supine position," *J Acoust Soc Am*, vol. 122, no. 1, pp. 532–541, Jul. 2007, doi: 10.1121/1.2715659.
[7] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J Biomed Opt*, vol. 15, no. 2, p. 021314, 2010, doi: 10.1117/1.3360308.
[8] Z. Zhang, S. Boyce, C. Y. Espy-Wilson, and M. K. Tiede, "Acoustic Strategies for Production of American English ``Retroflex'' /r/," *in Proceedings of the International Congress of Phonetic Sciences XV (ICPhS-15)*, 2003, pp. 1125-1128.
[9] S. R. Li et al., "Classification of accurate and error tongue movements for /r/ in children using trajectories from ultrasound," *J Acoust Soc Am*, vol. 145, no. 3, pp. 1799–1799, Apr. 2019, doi: 10.1121/1.5101588.

Results, Discussion, and Conclusion

Results: Images

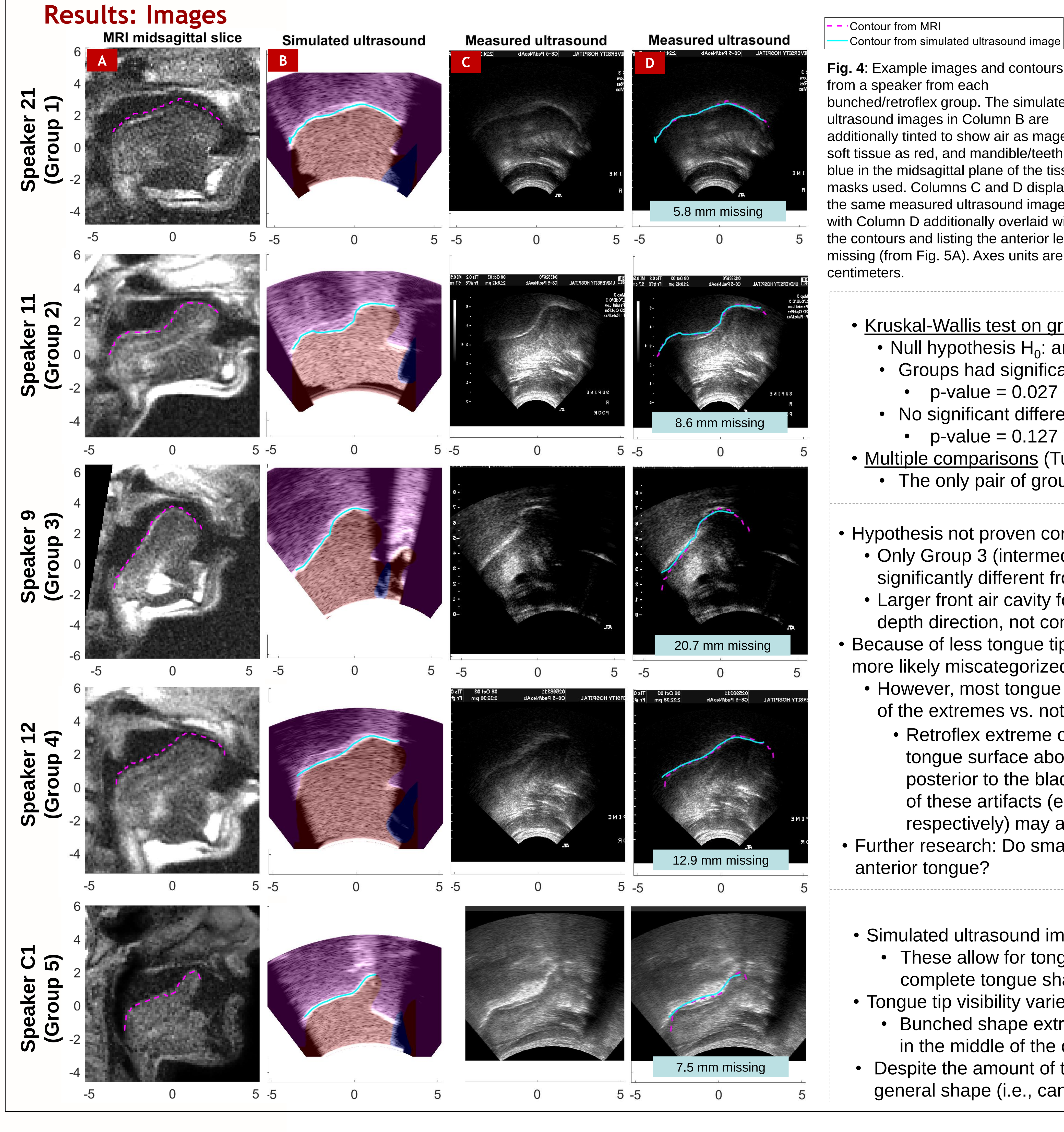

Fig. 4: Example images and contours from a speaker from each bunched/retroflex group. The simulated ultrasound images in Column B are additionally tinted to show air as magenta, soft tissue as red, and mandible/teeth as blue in the midsagittal plane of the tissue masks used. Columns C and D display the same measured ultrasound image, with Column D additionally overlaid with the contours and listing the anterior length missing (from Fig. 5A). Axes units are in centimeters.

Results: Quantification

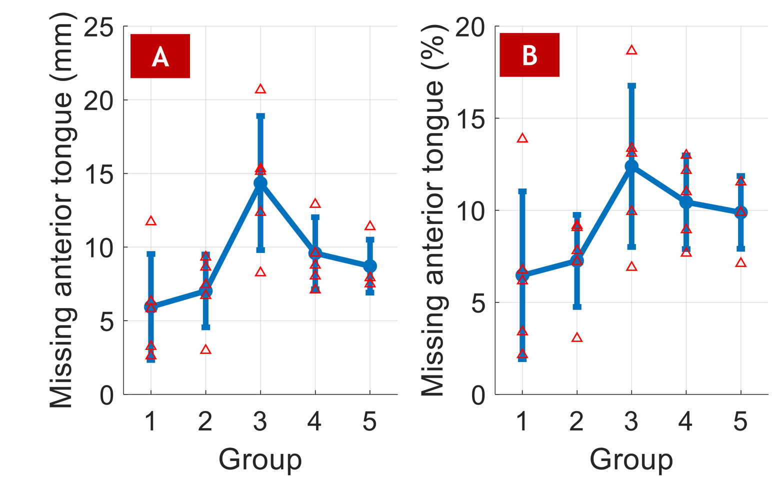

Fig. 5: Missing anterior tongue for each bunched/retroflex group. The mean and standard deviation for each group is shown as the blue dot and error bars respectively, with each red triangle showing the individual value for a speaker within the group. (A) indicates the missing length and (B) the proportion (via dividing by the entire MRI contour length).

Results: Statistical tests

- Kruskal-Wallis test on groups:
  - Null hypothesis H_0: amount of missing tongue in each group comes from the same distribution.
  - Groups had significantly different missing tongue length (mm):
    - p-value = 0.027 < 0.05
  - No significant difference among groups was found for proportion of missing tongue (%):
    - p-value = 0.127 > 0.05
- Multiple comparisons (Tukey-Kramer test comparing average group ranks):
  - The only pair of groups with significantly different averages is Groups 1 and 3 (for length).

Discussion

- Hypothesis not proven correct:
  - Only Group 3 (intermediate between retroflex and bunched) is significantly different from the bunched extreme (Group 1).
  - Larger front air cavity for retroflex extreme (Group 5) may be in depth direction, not contributing to obscuring the anterior tongue.
- Because of less tongue tip visibility, the middle group (Group 3) may be more likely miscategorized (e.g., Speaker 9 in Fig. 4).
  - However, most tongue shapes are often still recognizable as either of the extremes vs. not (Group 1 vs. Groups 2-4 vs. Group 5).
    - Retroflex extreme often has reverberation artifacts close to the tongue surface above the tongue dorsum and sometimes posterior to the blade/tip (Fig. 6), so the presence or absence of these artifacts (e.g., Speaker C1 and Speaker 12 in Fig. 4, respectively) may aid categorization.
- Further research: Do smaller probes (tighter curve) show more anterior tongue?

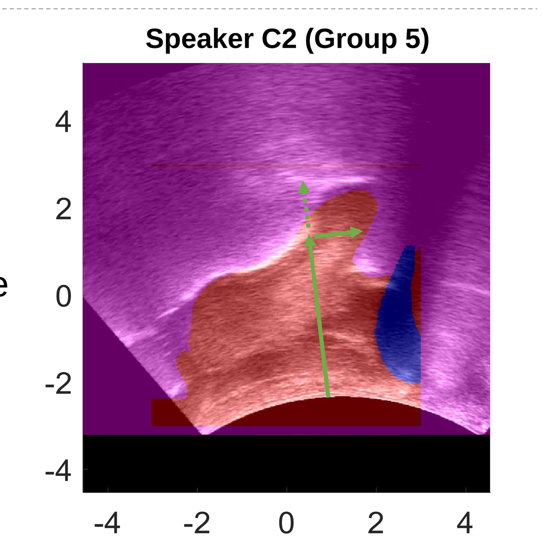

Fig. 6: Example measured ultrasound image with a reverberation artifact. The tint colors show midsagittal tissue masks from MRI. The probable path of the ultrasound wave for one scan line is shown as the green solid line with the artifact by the dashed line.

Conclusion

- Simulated ultrasound images replicate the amount of tongue tip missing.
  - These allow for tongue tip visibility to be quantified without ambiguity of not knowing the complete tongue shape during ultrasound image collection.
- Tongue tip visibility varies across groups along the bunched/retroflex continuum.
  - Bunched shape extremes (Group 1) have significantly less anterior tongue missing than shapes in the middle of the continuum (Group 3).
- Despite the amount of tongue tip missing, most ultrasound images of the tongue still show its general shape (i.e., can determine as Group 1 vs. Groups 2-4 vs. Group 5) for /r/.