Preliminary gamified ultrasound visual biofeedback on rhotic syllable tongue movement for speech therapy

Sarah R. Li,[1] Kathryn J. Eary,[1] Sarah Dugan, PhD CCC-SLP,[2,3] Maurice Lamb, PhD,[4] McKenna Becker,[5] Reneé Seward, MGD,[5] Suzanne Boyce, PhD CCC-SLP,[2] Michael A. Riley, PhD,[3] & T. Douglas Mast, PhD[1]

[1]Biomedical Engineering; [2]Communication Sciences and Disorders; [3]Rehabilitation, Exercise, and Nutrition Sciences; [5]Design | University of Cincinnati

[4]University of Skövde, Sweden

Introduction

Ultrasound biofeedback

- Ultrasound imaging provides visual biofeedback of the tongue (Fig. 1) in real time [1], which is particularly useful for refining tongue movement towards a pattern (i.e., a desired speech sound production).
- This refinement is helpful for:
  - Speech sound disorder (SSD) remediation [2]: changing "wabbit" to be the desired "rabbit."
  - Accent expansion in second language learning [3].
- Understanding tongue movement from ultrasound images can be difficult due to the rapidly moving appearance of the tongue surface, possibly in different directions for different parts of the tongue [4].

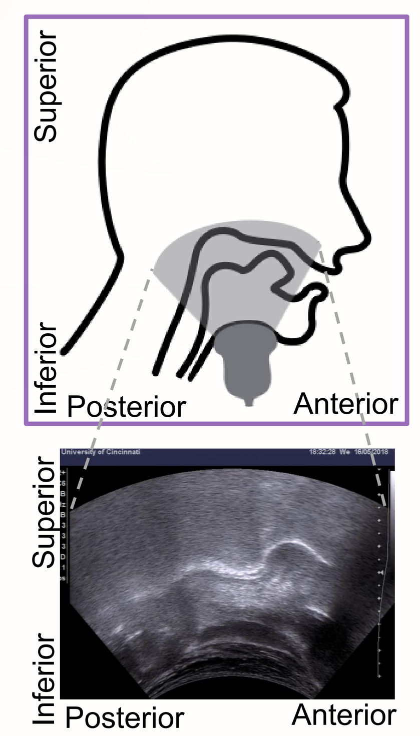


Figure 1: Ultrasound imaging of the tongue.

Tongue movement measurement

- Tongue movement from ultrasound images (Fig. 2) can be automatically measured as tongue part displacement trajectories [5, 6].
- These represent changes in vocal tract constriction due to the root, dorsum, and blade. (Positive displacement = narrowing of the vocal tract.)

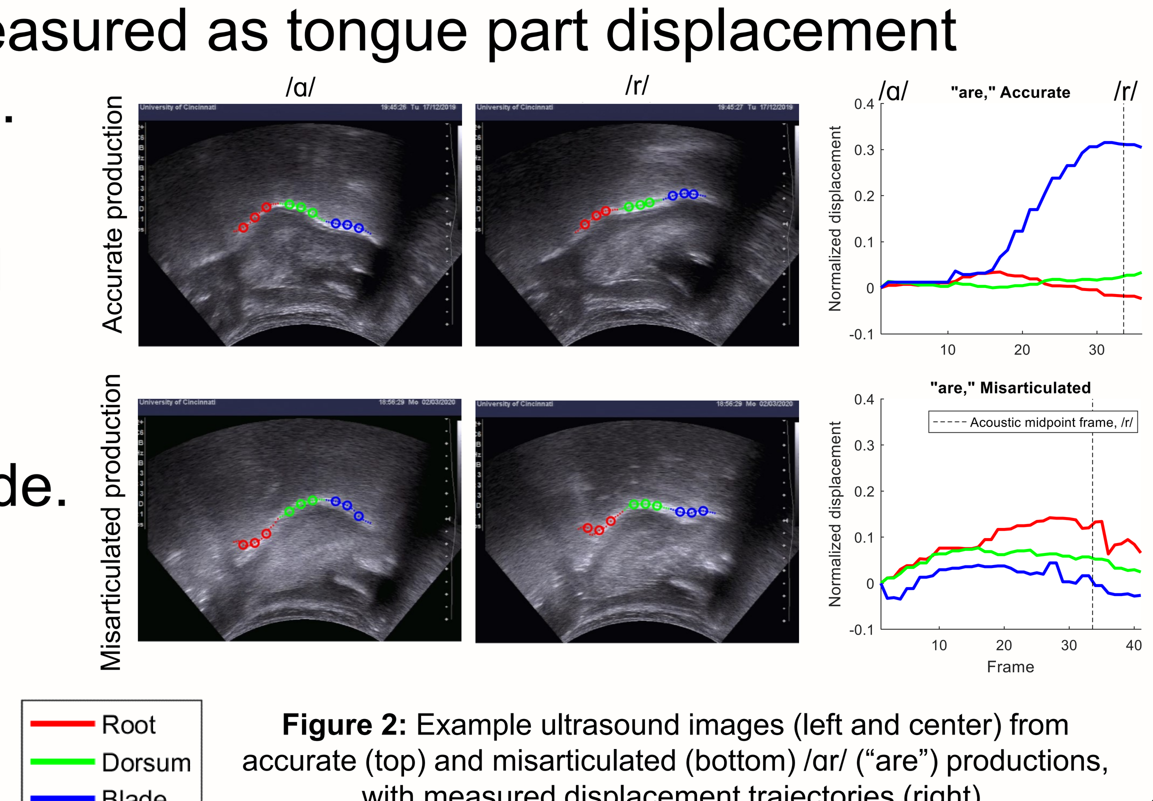


Figure 2: Example ultrasound images (left and center) from accurate (top) and misarticulated (bottom) /ɑr/ ("are") productions, with measured displacement trajectories (right).

Simple biofeedback parameter

- A linear combination of blade and dorsum displacement can be used as a parameter δ (Fig. 3) to represent the accuracy of a production, demonstrated previously for the syllables "ear," "are," "booer," "reap," "raw," and "rue" [5, 6].

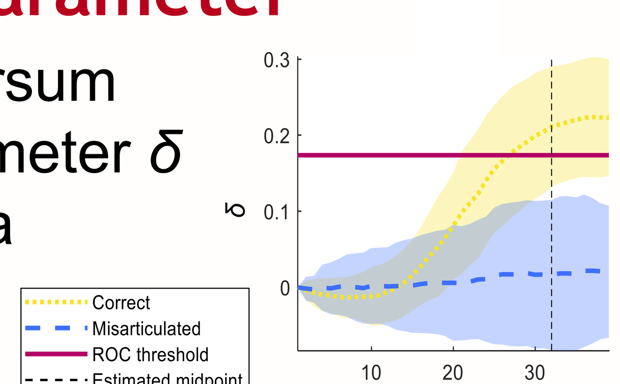


Figure 3: Averages (solid lines) and standard deviations (shading) of δ in "are" productions, with a discrimination threshold from a receiver operating characteristic curve (ROC).

Design

Design considerations

- A **gamified** presentation of this biofeedback may help elicit an external focus of attention, demonstrated in non-speech motor control studies to improve motor learning [7].
- **Difficulty level** affects motivation and the amount of information that an individual speaker can learn (see the Challenge Point framework) [8, 9].
- Difficulty level can be changed by adjusting **target width**. Fig. 4 shows a representation of a target with a width (between ····· lines) and amplitude (between 0 and – – –).
  - Fitts' Law relates index of difficulty (from target width and amplitude) to movement time (i.e., slower movement at harder difficulties retains accuracy) [10].
- **Feedback conditions** (e.g., delayed timing) impact motor learning [8].

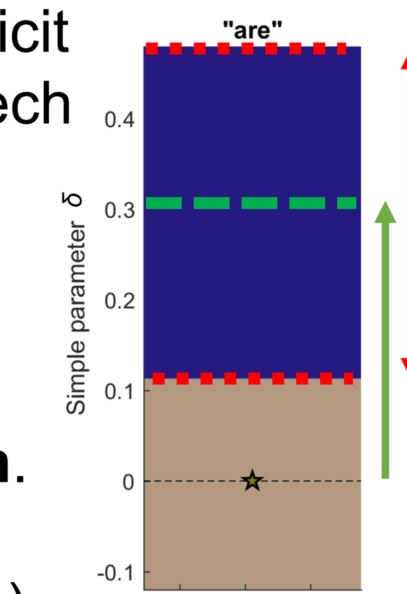


Figure 4: Representation of target; accurate /r/ is achieved by reaching the target area. The green arrow (left) shows the target amplitude and the red (right) the target width.

STAR and *GO, GO Goat!* games

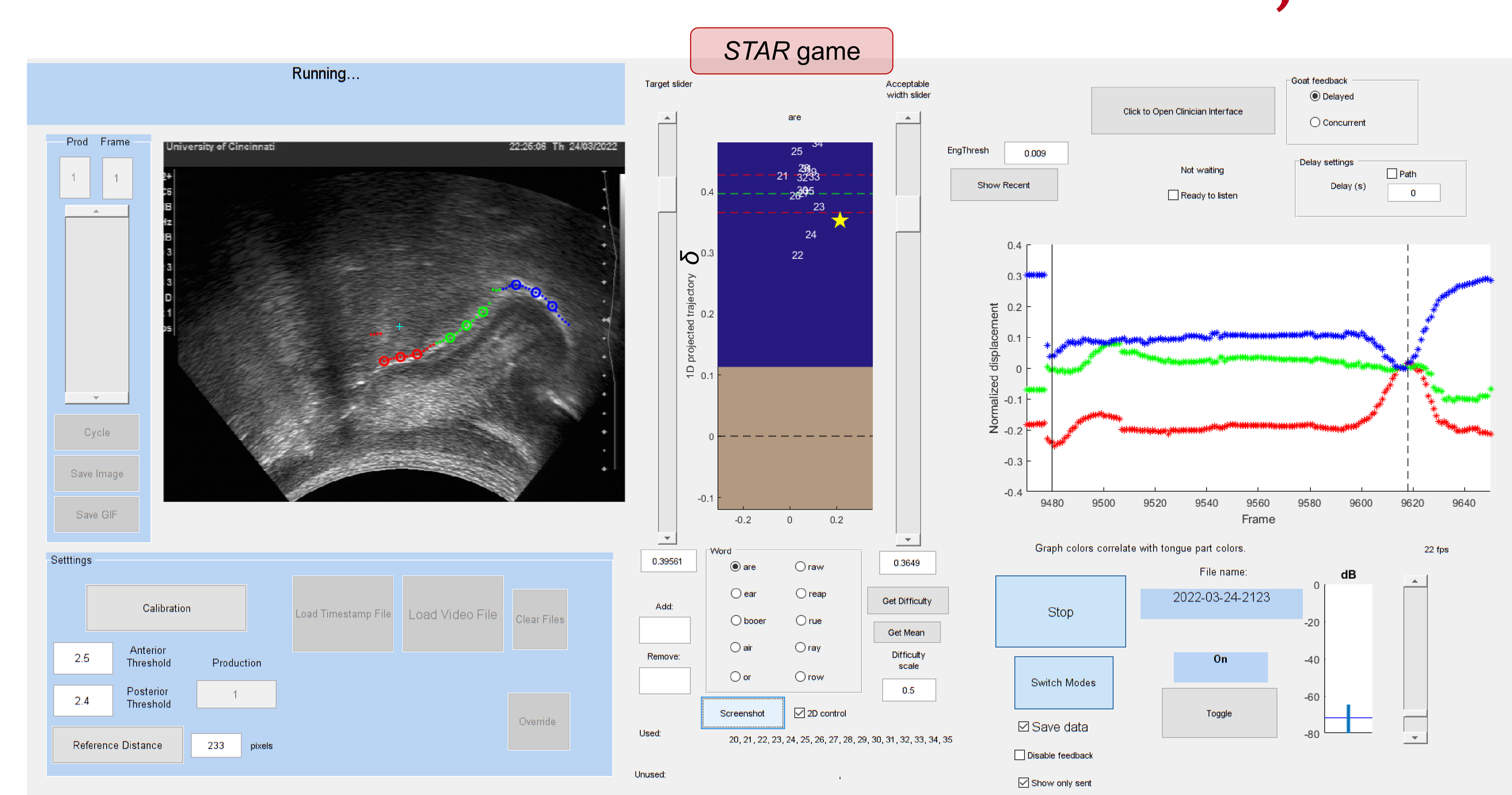


Figure 5: *TonguePART* Graphical user interface (GUI) for measuring tongue part displacement trajectories and *STAR* game (center). The *STAR* game shows the δ parameter during all the frames of a production as a moving star, with the representation of the target (here, the red and green dashed lines above the star) directly compared to the δ parameter.

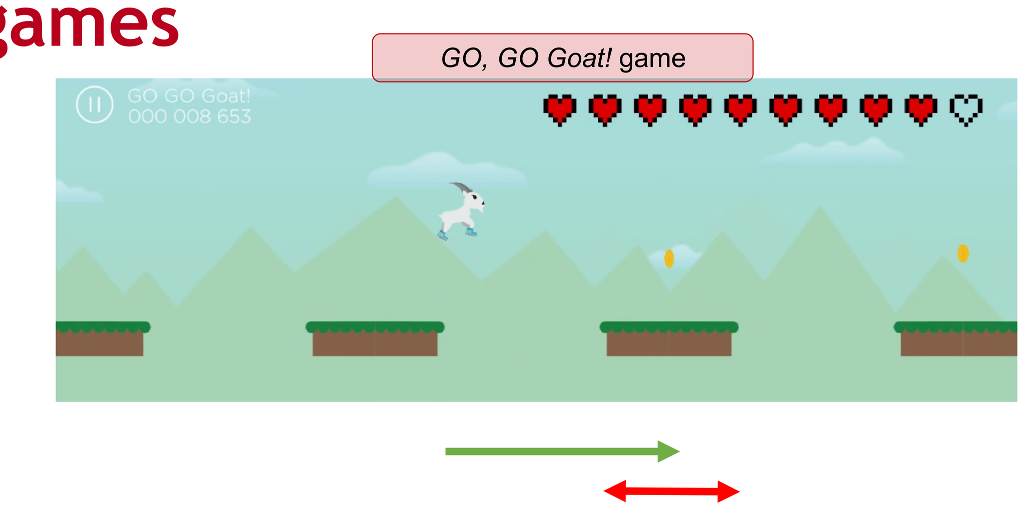


Figure 6: *GO, GO, Goat!* game. The green arrow (top) shows the mapping of target amplitude and the red (bottom) that of target width. After a production finishes, the goat jumps; where this goat lands is the δ parameter at one frame (estimated midpoint; i.e., calculated from the displacements between the initial and midpoint frames) of a production linearly scaled to the representation of the target as platforms in the game.

- Both games: are audio activated, use the simple parameter δ

Table: Differences in feedback conditions between the *STAR* and *GO, GO, Goat!* games.

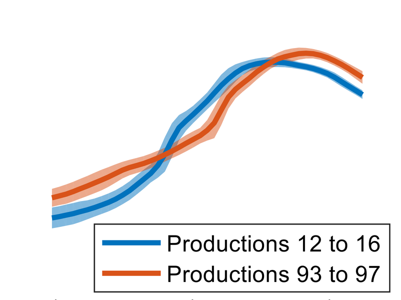
STAR	*GO, GO Goat!*
• Concurrent	• Delayed
• Greater amount (more frames)	• Lesser amount (fewer frames)
• Direct target	• Scaled target

Results

Pilot data (two participants)

- One participant with typical speech showed a change in articulation (Fig. 7).
- Participants found both games likeable, with the STAR game liked slightly better.

Figure 7: Comparison of tongue surface contours at end (/r/) of productions from an adult participant with typical speech, with the task of reaching a specified target for /ɑr/ ("are"). Averages (solid lines) and standard deviations (shading) are shown for a selection of productions. Contours were rotated so that mandible shadows on the images matched, and interpolation was used to standardize the horizontal sampling.



Conclusion: These gamified biofeedback interfaces allow for testing of motor learning principles and have potential use for speech therapy.

University of Cincinnati

References

[1] J. L. Preston et al., "Ultrasound images of the tongue: a tutorial for assessment and remediation of speech sound errors," *JoVE*, no. 119, pp. e55123, Jan. 2017, doi: 10.3791/55123.

[2] 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.

[3] T. L. Face, "Ultimate attainment of Spanish rhotics by native English-speaking immigrants to Spain," *Lengua y Migracion*, vol. 10, pp. 57–80, Jan. 2018.

[4] 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.

[5] S. R. Li et al., "An automatic, simple ultrasound biofeedback parameter for distinguishing accurate and misarticulated rhotic syllables," in *Proc Interspeech 2021*, 636-640, doi: 10.21437/Interspeech.2021-1749

[6] S. R. Li et al., Classification of accurate and misarticulated /ɑr/ for ultrasound biofeedback using tongue part displacement trajectories," *Clin Linguist Phon*, pp. 1–27, Mar. 2022, doi: 10.1080/02699206.2022.2039777.

[7] G. Wulf, C. Shea, and R. Lewthwaite, "Motor skill learning and performance: a review of influential factors: Motor skill learning and performance," *Med Educ*, vol. 44, no. 1, pp. 75–84, Jan. 2010, doi: 10.1111/j.1365-2923.2009.03421.x.

[8] E. Maas et al., "Principles of Motor Learning in Treatment of Motor Speech Disorders," *Am J Speech Lang Pathol*, vol. 17, no. 3, pp. 277–298, Aug. 2008, doi: 10.1044/1058-0360(2008/025).

[9] McAllister Tara, Hitchcock Elaine R., and Ortiz José A., "Computer-Assisted Challenge Point Intervention for Residual Speech Errors," *Perspect ASHA Spec Interest Groups*, vol. 6, no. 1, pp. 214–229, Feb. 2021, doi: 10.1044/2020_PERSP-20-00191.

[10] S. R. Kuberski and A. I. Gafos, "Fitts' Law in Tongue Movements of Repetitive Speech," *Phonetica*, 2019, doi: 10.1159/000501644.