# 5aSC1. A biomechanical model for infant speech and aerodigestive movements


**Connor Mayer[1] - Ian Stavness[2] - Bryan Gick[3]**

[1]Dept. of Linguistics, UCLA - [2]Dept. of Computer Science, USask - [3]Dept. of Linguistics, UBC; Haskins Labs

Correspondence: connormayer@ucla.edu


## Background and objective

**How do infants learn speech movements *rapidly* and with *limited input?***

**Hypothesis**: Core speech movements may build on preexisting aerodigestive movements [1, 2].

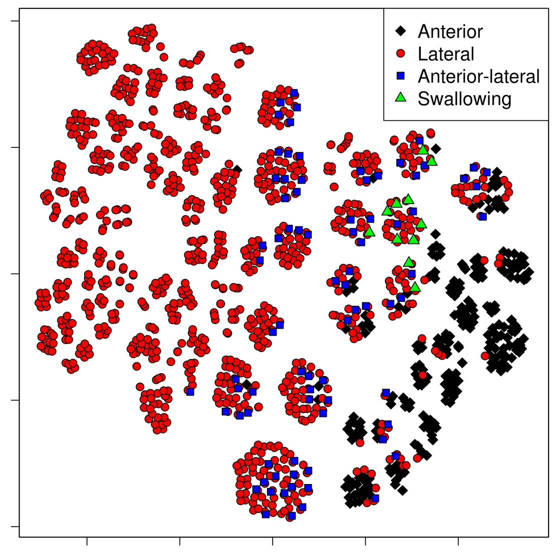
- e.g. swallowing, suckling, etc.
- There is neurological, clinical, and kinematic evidence relating speech and aerodigestion.
- Existing structure makes learning problem tractable.

**Problem:** Difficult to test from the perspective of motor control.

**Solution:** *Biomechanical modeling* provides a way to test some of the predictions of this hypothesis.

Biomechanical models of the adult vocal tract have been used to test hypotheses about infants [3].

- But adults and infants have different vocal tracts!

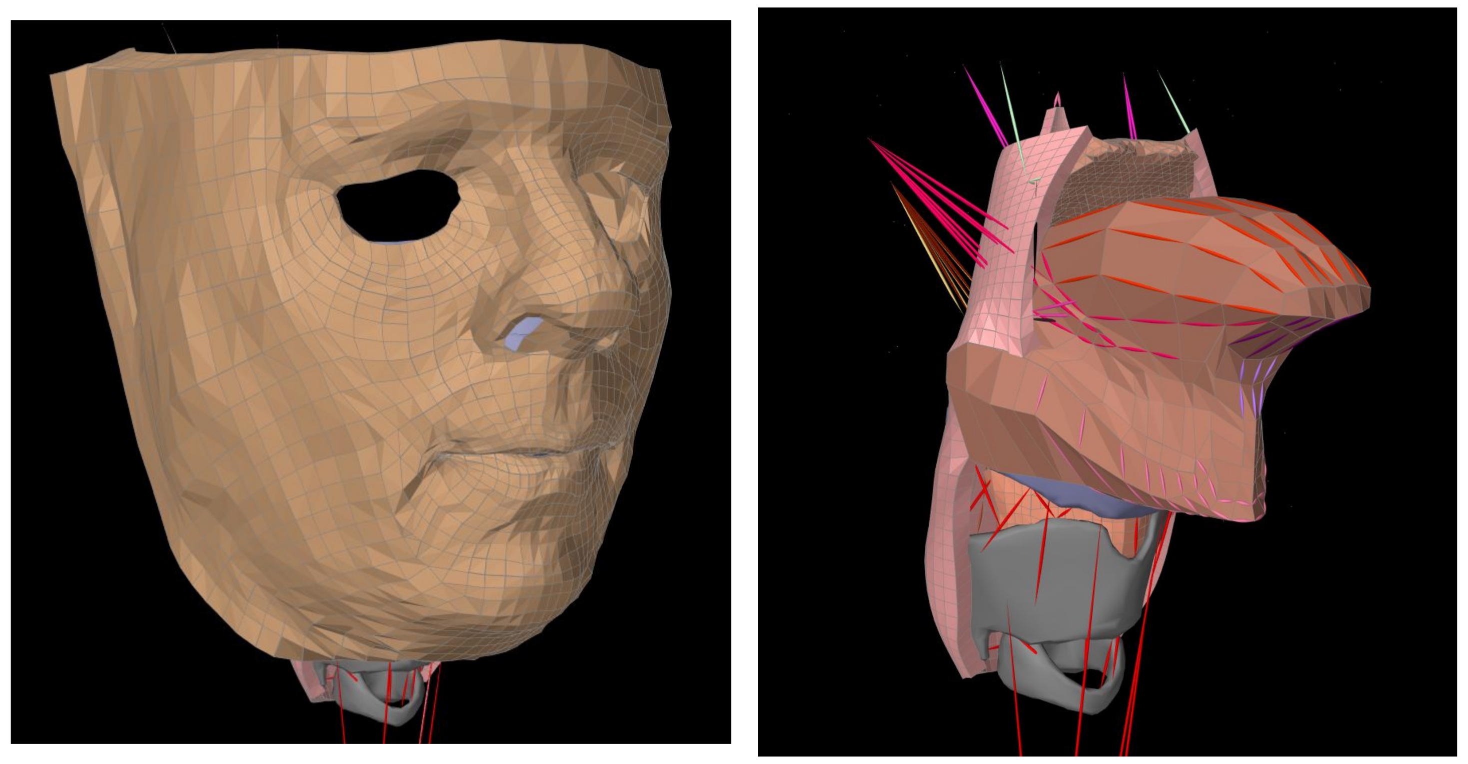



A t-SNE visualization of the muscle activation space for various speech and aerodigestive movements [3].

**Objective:** Create a biomechanical model of an infant tongue and palate using Artisynth [4, 5].

## The Frank vocal tract model

The biomechanical modeling platform *Artisynth* contains a model of the adult vocal tract, called *Frank*.

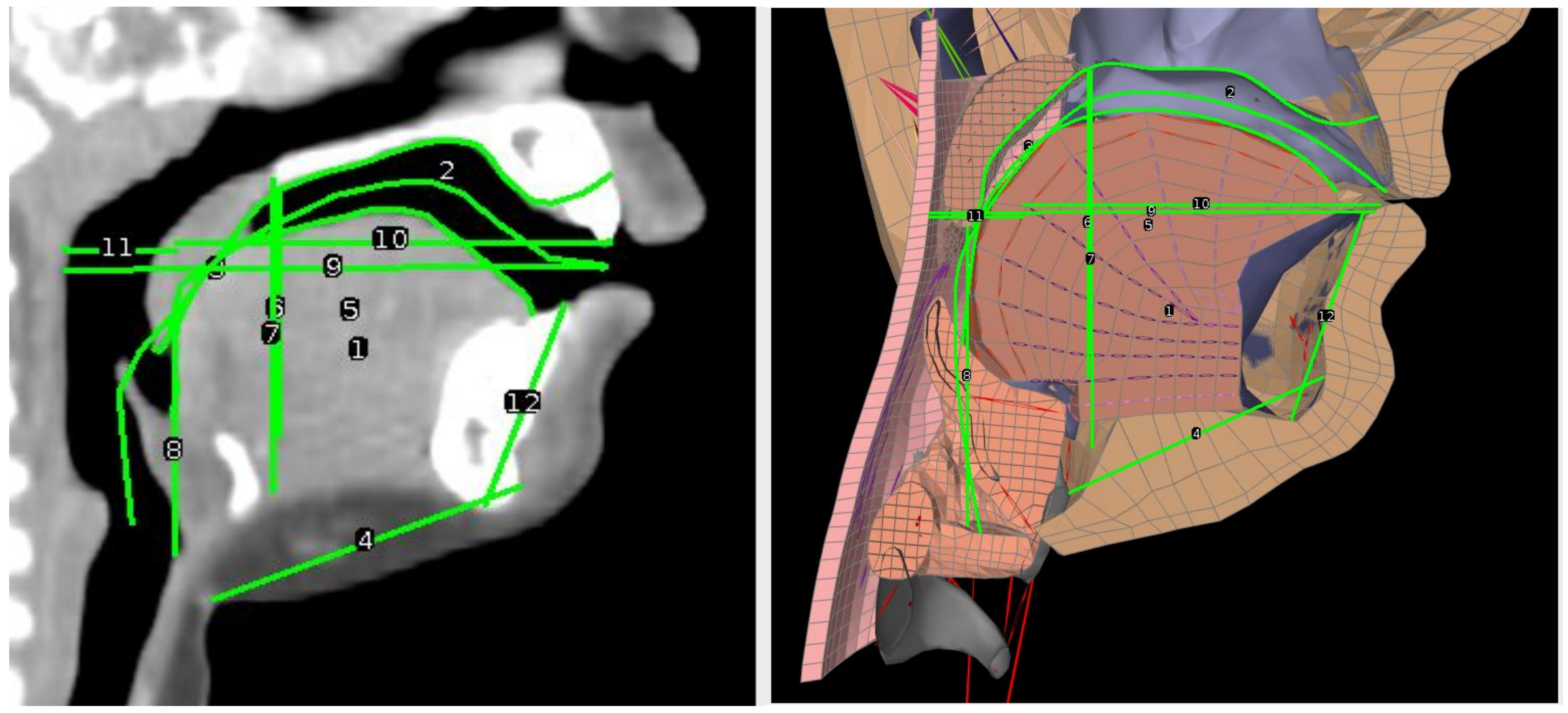


## Comparing the infant and Frank vocal tracts

We compared a midsagittal CT image of an 11-month male infant vocal tract with a midsagittal image of the Frank model.

Quantitative comparison was done using biometric measurements adapted from previous developmental studies of the vocal tract [6, 7].

1. **Vocal tract length**
2. **Hard palate length**
3. **Soft palate length**
4. **Mandibular length**
5. **Tongue length**
6. **Hyoid height**
7. **Larynx height**
8. **Oropharynx length**
9. **Vocal tract horizontal length**
10. **Anterior cavity length**
11. **Oropharynx width**
12. **Mandible height**



Each measurement is normalized based on the **mandible height** from the same image.

- Allows comparison between images with different scales

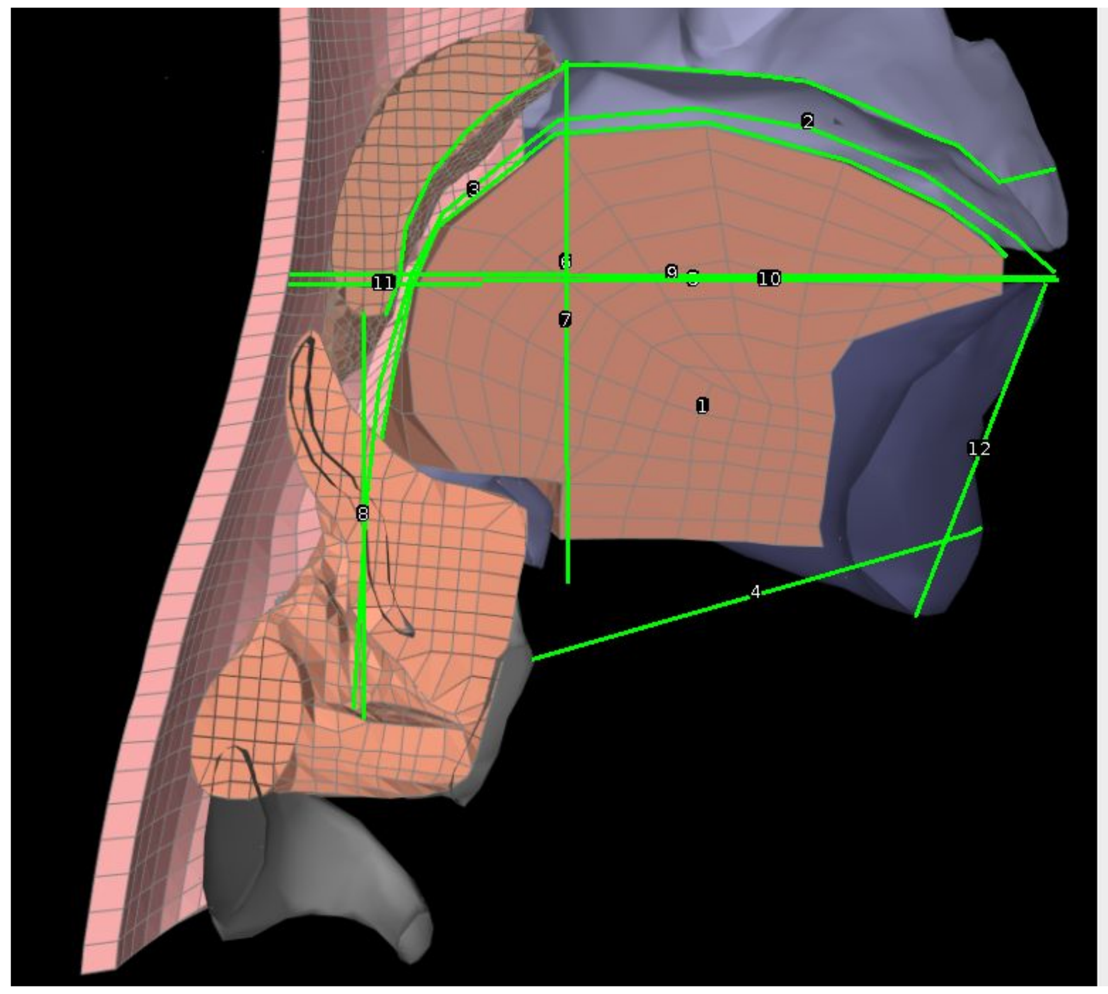
## Modifying the Frank vocal tract

Compared to Frank the infant has proportionally a:

- **Higher larynx**
- **Higher hyoid**
- **Longer anterior cavity**

We modified the structures in the Frank vocal tract with:

- Rigid body translations
- Affine transformations
- Removal of excess tissue including face



## Quantifying comparisons

We use the following measure to quantify the difference between the infant and Artisynth vocal tracts:

$$\theta = \sum_{i=1}^{m} I_i (M_i - I_i)^2$$

Where

- *i:* indexes over the set of measurements
- $M_i$: normalized measurement from model
- $I_i$: normalized measurement from infant

This measurement:

- Penalizes differences in grosser measurements
- Indicates a better fit with lower values (0 = perfect)

## Results

Original Frank model: **θ = 1.438**

Infant Frank model: **θ = 0.373**

The infant Frank model corresponds more closely to the proportions of the infant vocal tract

## Discussion and future work

This model will allow researchers to:

- Simulate aerodigestive and speech movements.
- Supplement evidence from other domains bearing on connection between speech and aerodigestion.

Generally useful for research on the infant vocal tract.

**Future directions:**

- Ensure muscle insertions are accurate.
- Replace skull with models generated from infant data.
- Simulate swallowing and speech sounds.

## References


[1] **MacNeilage, P.** (2008). The Origin of Speech. Oxford and New York: Oxford University Press. [2] **Studdert-Kennedy, M.**, & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. In M. Christiansen & S. Kirby (Eds.), Language Evolution. Oxford and New York: Oxford University Press. [3] **Mayer, C.**, Roewer-Despres, F., Stavness, I., & Gick, B. (2017). Do innate stereotypies serve as a basis for swallowing and learned speech movements? Behav. & Brain Sci., 40, e395. [4] **Stavness, I.**, Lloyd, J.E., & Fels, S.S (2012). Automatic Prediction of Tongue Muscle Activations Using a Finite Element Model. *Journal of Biomechanics, 45*(16), 2841-8. [5] **Gick, B.**, Anderson, P., Chen, H., Chiu, C., Kwon, H.B., Stavness, I., Tsou, L., & Fels, S. (2014). Speech function of the oropharyngeal isthmus: A modeling study. *Computer Methods in Biomechanics & Biomedical Engineering: Imaging & Visualization, 2*(4), 217-22. [6] **Vorperian, H.K.**, Kent, R.D., Lindstrom, M.J., Kalina, C.M., Gentry, L.R., & Yandell, B.S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *J. Acoust. Soc. Am., 117*(1), 338-350. [7] **Vorperian, H.K.**, Wang, S., Chung, M.K., Schimek, E.M., Durtschi, R.B., Kent, R.D., Ziegert, A.J., & Gentry, L.R. (2009). Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *J. Acoust. Soc. Am, 125*(3), 1666-1678.



Thanks to Houri Vorperian for the infant CT data, Sid Fels and C. Antonio Sánchez for useful suggestions on producing the model, and the members of the UBC Interdisciplinary Speech Research Lab for their comments and feedback. We acknowledge funding from NSERC and SSHRC.