Towards Speech Articulation Simulation with a Dynamic Coupled Face-Jaw-Tongue Model

Ian Stavness[1]*, John E. Lloyd[1] , Yohan Payan[2,3] , Sidney Fels[1]

[1]Electrical and Computer Engineering, University of British Columbia

[2]TIMC-IMAG Laboratory, UMR CNRS 5525, University Joseph Fourier

[3]GIPSA Laboratory, UMR CNRS 5216, Grenoble Institute of Technology

1. Introduction

Understanding the motor control strategies underlying human speech production is an active area of research. In order to separate features of observed speech movements that arise from central motor commands from those that are due to the mechanics of the peripheral musculoskeletal system, a number of researchers have developed detailed biomechanical models jaw, tongue, larynx, lips, and face. The finite-element method (FEM) is commonly used to characterize the non-linear mechanical properties of the face, lips, tongue and other soft-tissue structures of the upper airway. A number of models have been proposed, but typically focus on one single articulator. 3D FEM models of the face include Sifakis et al. (2006) and Nazari et al. (2010). Coupled jaw-tongue-hyoid models include an early 2D model by Sanguineti et al. (1998), a 3D mass-spring tongue and rigid-body jaw model by Fang et al. (2009), and a 3D dynamic FEM tongue by Buchaillard et al. (2009) coupled with a rigid-body jaw-hyoid model by Stavness et al. (2010).

We are developing a biomechanics simulation toolkit ArtiSynth (www.artisynth.org) in order to simulate coupled hard and soft tissue structures. Here we present the first biomechanical model that dynamically couples the face and lips with a jaw-tongue-hyoid model, as pictured in Figure 1.

2. Modeling requirements

Physiologically-based simulation of speech articulation presents a number of unique challenges as compared to traditional musculoskeletal modeling of limb movements. Firstly, the upper airway involves a combination of hard tissues (skull and jaw bones) and soft-tissues (face, lips, tongue, and soft-palate). While non-linear FEM models can be used to represent biological soft-tissue mechanics, the technique is computationally expensive and using detailed FEM models to represent the predominantly rigid movement of skeletal structures is particularly inefficient. For this reason, we have focused on developing constraint mechanisms in ArtiSynth to handle mechanical coupling between rigid and deformable bodies.

*stavness@ece.ubc.ca

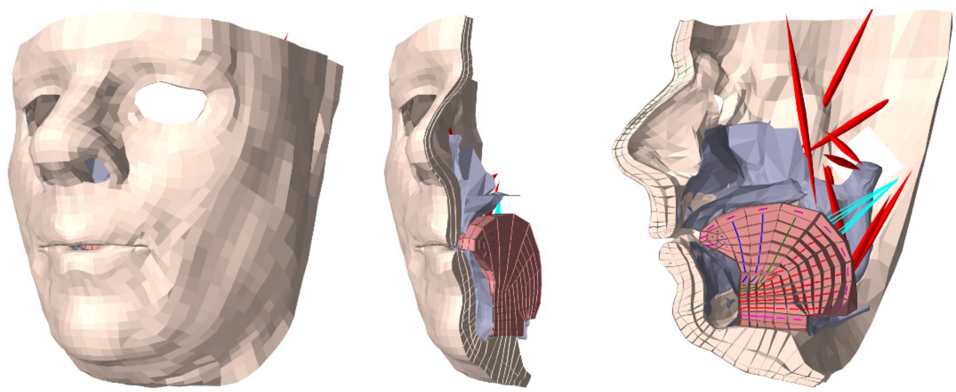

Figure 1. Oblique, oblique cut-away and sagittal cut-away views of the face-jaw-tongue model at rest.

Speech movements are typically observed only in the mid-sagittal plane (with electromagnetic articulography, dynamic MRI, ultrasound, or video fluoroscopic recordings). However, the 3D shape of the vocal tract is important factor in the modulation of acoustics during speech production. Also, the mid-sagittal shape of the tongue is determined by its 3D volume (lateral expansion and contraction of the tongue body) as the tongue is a nearly-incompressible tissue. For these reasons, 3D FEM modeling is important as well as incompressibility. ArtiSynth is a fully 3D simulation environment with tools for viewing and manipulating the model in 3D. ArtiSynth includes hyper-elastic FEM materials that are nearly-incompressible and includes the option to enforce FEM incompressibility as a constraint.

Another crucial component of speech production is contact. Contact is observed between different articulators and includes deformable to deformable body contact (between the upper and lower lip) as well as deformable to rigid body contact (between the lips and teeth and the tongue and palate). ArtiSynth supports mesh-based collision detection and contact handling using dynamic constraints. Sub-regions of an FEM model can be defined for purposes of handling contacts between two regions of an FEM model, e.g. between the upper and lower lips of the face mesh, and we are working on other techniques for robust self-collision detection.

3. Face-Jaw-Tongue model

Our face-jaw-tongue-hyoid model couples two previous reference models that were adapted and registered to Computed Tomography (CT) data for a specific subject, including a dynamic FEM face model (Bucki et al., 2010) and a jaw-tongue-hyoid model (Stavness et al., 2010). The volumetric face mesh and the surface jaw and skull meshes had been registered to the subject's CT data; however, the inner-surface of the face mesh did not exactly conform to the jaw and skull surfaces due to inaccuracies in the original adaptation. To improve the fit, we fixed the outer surface of the face mesh and used contact detection to conform the inner face surface to the jaw/skull surfaces. Once the contact-based morphing was complete, the outer-face-surface nodes were set

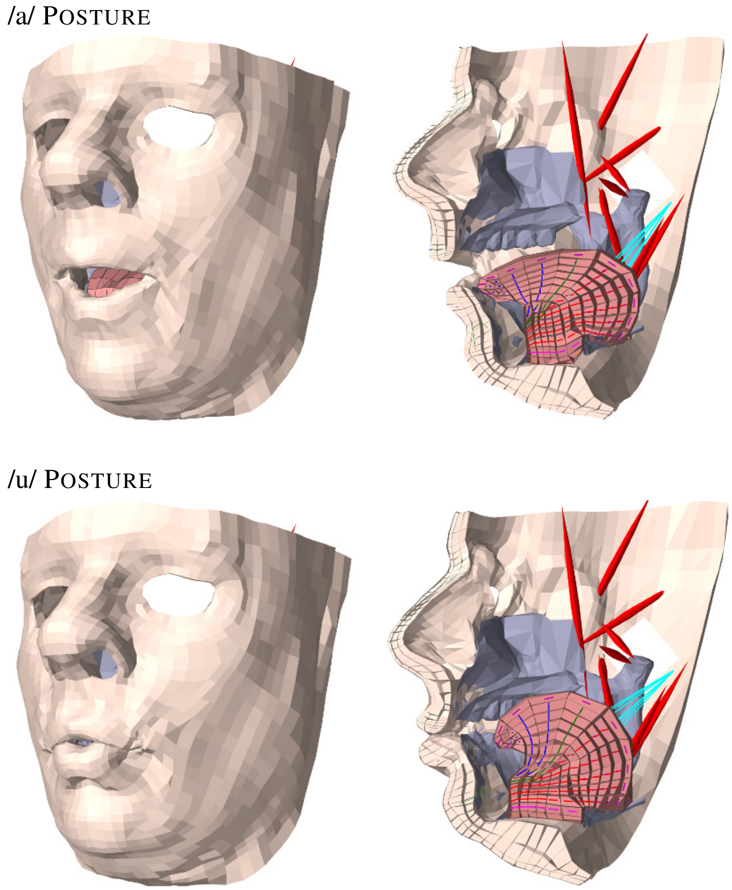

Figure 2. Oblique and sagittal cut-away views of the face-jaw-tongue model during muscle activated vowel articulations: /a/ (upper panels) and /u/ (lower panels).

free and a number of inner-face-surface nodes were attached to the jaw and skull, with the nodes around the lips and cheeks left free. Face muscles were defined by serial line segments in the original face model and were adapted to the subject along with the FEM mesh (Bucki et al., 2010). The adapted muscle paths were then manually smoothed to correct for registration artifacts.

4. Speech simulations

We have identified a few particular speech articulations that we are investigating with our new modeling platform. Example articulation simulations are shown in Figure 2. Contact between the tongue and the lower teeth occurs in the /a/ simulation. Also, contact between the upper tongue surface and the hard palate occurs during the alveolar stop /t/ and the French syllable /ta/. In our previous work we have simulated maximum voluntary tongue-palate pressure and shown it to be within a physiologically valid range. Muscle forces required to create the physiological level of pressure observed in speech articulations is of interest.

Labial phonemes are also of interest, including articulations that include lip-to-teeth contact, such as /fa/ and /va/, and bilabial lip-lip contact, such as /ba/ and /pa/. Lip rounding and protrusion have been simulated with the face model alone (Nazari et al., 2010) and we are interested in modulating the degree of jaw opening by activating jaw opening muscles in the coupled face-jaw-tongue model.

We will report our progress on this speech articulation simulations at the ISSP 2011 meeting. Also, the ArtiSynth modeling platform is open source and freely available for use by the ISSP community. Therefore, we hope to engage ISSP 2011 attendees to discuss the application of biomechanical modeling to the analysis of speech production mechanics and motor control.

5. Acknowledgements

We thank Pierre Badin at GIPSA-Lab, Grenoble for providing the CT data used in the subject-specific model. We also thank Mohammad A. Nazari and Pascal Perrier at GIPSA-Lab, Grenoble for their assistance with simulations of lip protrusion.

References

Buchaillard, S., Perrier, P., and Payan, Y. A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *Journal of the Acoustical Society of America*, 126(4):2033–2051, 2009.

Bucki, M., Nazari, M., and Payan, Y. Finite element speaker-specific face model generation for the study of speech production. *Computer Methods in Biomechanics and Biomedical Engineering.*, 13(4):459–467, 2010.

Fang, Q., Fujita, S., Lu, X., and Dang, J. A model-based investigation of activations of the tongue muscles in vowel production. *Acoustical Science and Technology*, 30(4): 277–287, 2009. ISSN 1346-3969.

Nazari, M. A., Perrier, P., Chabanas, M., and Payan, Y. Shaping by stiffening: a modeling study for lips. *Motor Control*, in press, 2010.

Sanguineti, V., Laboissière, R., and Ostry, D. J. A dynamic biomechanical model for neural control of speech production. *Journal of the Acoustical Society of America*, 103 (3):1615–1627, 1998.

Sifakis, E., Selle, A., Robinson-Mosher, A., and Fedkiw, R. Simulating speech with a physics-based facial muscle model. In *Proceedings of ACM Symposium on Computer animation*, pages 261–270, 2006.

Stavness, I., Lloyd, J., Payan, Y., and Fels, S. Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. *International Journal of Numerical Methods in Biomedical Engineering*, 27:367–390, 2010.