The interplay between dynamics and muscle activation in the tongue during speech

Carolyn Essex-Torcaso[1], William S. Levine[1], Emi Z. Murano[2], Maureen Stone[2]

[1]University of Maryland, Department of Electrical and Computer Engineering
AV Williams Building, College Park, MD 20740, USA

[2]University of Maryland Dental School, Depts of Biomedical Sciences and
Orthodontics, 650 W. Baltimore St, Baltimore, MD, 21201, USA

`essex@isr.umd.edu, wsl@eng.umd.edu, emurano@gmail.com, mstone@umaryland.edu,`

Abstract. *The human tongue is used in speaking, an activity unique to humans. The tongue is unusual in that it does not contain any rigid structure, such as a bone, for its muscles to act against. In an effort to better understand the biomechanics and control of the tongue, a mathematical model has been created. This model is in the form of a nonlinear controllable incompressible elastic structure that undergoes large deformations. The model was then used to solve a simplified inverse problem. The trajectories of arbitrary points within the tongue were obtained by means of tagged cine MRI. This data was then used as input to a two-dimensional simplification of the model. The output was the set of muscle activations that produced the best approximation to the observed motion in a least squares sense. This results in information relevant to two biological questions—How many independently controllable motor units are there in the tongue? How important are the dynamics of the tongue in determining its shape?*

1. Introduction

Humans use their tongue to produce intelligible speech and to process food in their mouth. Thus, understanding the tongue's dynamics and control would be very helpful to understanding speech production, treatment of lingual tumors, macroglossia and sleep apnea. Speech is both unique and common to all humans. Thus, a deeper understanding of the dynamics and control of the tongue as it is applied to speech might lead to insight into the operation of the human brain.

The biomechanics of the tongue is also special. It is one of a small group of biomechanical systems that do not have rigid skeletons, known as muscular hydrostats [Smith & Kier, 1989]. The rich collection of movements of these systems results from the interaction of actively contracting orthogonal muscles with the incompressibility of the overall structure.

We have been developing a mathematical model of the human tongue as part of a larger project to study and understand the role of the tongue in speech.

2. Background

Our previous work on the mathematical modeling of the tongue concentrated first on the simplest possible model, a linear small deformations model [Iny & Levine, 1999]. This work led to useful insights [Iny et al., 2000] but we ultimately showed that a small deformations model cannot adequately describe many tongue movements [Levine et al, 2005]. This led us to create our new nonlinear incompressible large deformations model.

In most finite element models the tongue is replaced by the elements, which are chosen to simplify the inclusion of muscles in the model. This greatly simplifies the development of the model. However, it means that there is no underlying mathematical description of the tongue. There is no way to refine these models by continually increasing the number of finite elements. Mathematical questions, such as the existence and uniqueness of solutions are not addressed. By creating an underlying smooth mathematical model of the tongue we make it possible to address the mathematical, computational, and physiological questions.

3. The Model

Because water comprises approximately 85% of its tissues, it is well-approximated by an incompressible elastic structure that undergoes large deformations. This means that the material form for the dynamics is preferable and that the core of the model is the standard equations for an incompressible elastic solid that undergoes large deformations. The precise details of the model can be found in [Levine et al, 2005]. The mathematics is thoroughly explained in [Gartner & Hiatt, 1997].

3.1 The Muscle Model

Muscles consist primarily of contractile fibers encased in a lattice of collagen fibers [Zajac, 1985]. There are also blood cells, nerve cells, and fat cells but we believe these can be ignored at the level of accuracy of our model. The lattice of collagen fibers is aligned to the muscle fiber direction so that one part of the lattice is parallel to the contractile fibers. Both collagen and muscle fibers offer no resistance to either compression or shear other than that resulting from their incompressibility.

Isolated collagen fibers have a smooth, monotonic increasing force/length curve in extension. At zero force, the collagen fibers curl. Small forces uncurl the collagen. Once straight, collagen behaves like a linear spring. For simplicity, we approximate the collagen fibers by a linear, homogeneous, and isotropic stress/strain relation that applies at every point in the tongue. As with all our simplifying assumptions, this is subject to change if it can be shown to lead to measurable differences between our analytical and experimental results.

The stress due to muscle fibers is added to that due to collagen. This is a standard assumption in the literature on muscle models [Gerard et al, 2003]. The model of the contractile fibers in a muscle and of the muscles in the tongue consists of a geometric component and a stress/strain component. Geometrically, a muscle consists of a set of points in R^3. If we denote the points of the ith muscle by M_i and the points of the tongue by B, then

$$M_i \text{ ï } B \text{¿½} R^3 \text{¿} \tag{1}$$

Each point of a muscle has a contractile fiber running through it. Mathematically, a contractile fiber is just a curve extending from one point of the muscle to another. These curves can be taken to be as smooth as necessary. Physically, there can only be one such fiber through any single point of the tongue. Fibers from different muscles are often closely interleaved, even when the fibers are nearly perpendicular. Mathematically, we approximate this by allowing a point to belong to several muscles and to have one fiber from each muscle pass through it. Lastly, the stress-strain characteristics of each fiber are described by a distributed version of the Zajac lumped muscle model [Monti et al, 2001] aligned along the muscle fiber. The precise details can be found in [Levine et al, 2005].

4. The Inverse Problem

Skeletal muscle fibers are sometimes in series and act functionally as a single fiber that extends the full length of the muscle, terminating in either bone or tendon [Slaughter et al., 2005]. Such a fiber is either activated or not. Because such muscles are used to produce force, there is no reason to activate only one or two fibers in the series. The muscles of the tongue are different. They can be in series, in parallel or radial and may originate and/or terminate in soft tissue. It is reasonable to expect different activation of fibers in series because of the need to create a shape as well as a force. There is anatomical evidence for such differential activation along the length of in-series tongue muscles and along the width of parallel and radial muscle fibers [Slaughter et al, 2005, Miyawaki, 1975]. On the other hand, it is extremely unlikely that the muscles of the tongue are controlled fiber by fiber by the brain. That would require entirely too much calculation. This leaves the question: To what extent are the muscle fibers in the tongue independently controllable.

To answer this question, our model consists of 6 muscles projected to the midsagittal plane: Styloglossus, Hyoglossus, Verticalis, Inferior Longitudinalis, Superior Longitudinalis (SL), and Genioglossus (GG). Many of the muscles are known or hypothesized to have independently activated subregions. In particular, it is known that SL has independent lengthwise activations [Slaughter et al, 2005] and GG has independent radial activations [Miyawaki, 1975]. We have also been collecting tagged cine MRI images of the tongue during speech. The tags decay after approximately 1.2 seconds limiting the duration of the speech sound. While the tags are present we obtain the position, at each of a sequence of discrete values of time, of any material point within the tongue. Using this data as input to our model of the tongue allows us to define a well-posed inverse problem. Given the trajectories of any number of material points in the tongue, compute the control inputs that created these trajectories.

The first step in solving this problem is to create a finite element approximation to our tongue model. In this initial work, one subject makes two sounds, one after the other. They are the vowels /i/ and /u/. As long as the subjects do not make extreme deformations, our model is a linear system of algebraic equations with a nonlinear, but constant, constraint due to incompressibility, and the incompressibility constraint can be approximated by a value of Poisson's ratio close to ½. This results in an unconstrained linear system of algebraic equations of the form

$$Au(t_k) = M\ddot{d}(t_k) + C\dot{d}(t_k) + Kd(t_k) \tag{2}$$

where $\underline{u}$ is the vector of muscle activations and M, C, and K are known. The vector $d(t_k)$ is the position of each point of the FEM of the tongue at time t_k. These positions are determined from the MRI. The derivatives are approximated from the positions t successive times.

Of course, $\underline{A}$ is an nï¾m matrix with n m so Eqn. (2) does not generally have a solution. The simplest and most obvious way to obtain a solution is by least squares. This is what we have done for several different collections of independent activations under two different approximations. The first is assuming that the derivatives are zero —the quasi static approximation common in tongue research. The second includes the approximate derivatives.

5. Results and Discussion

We addressed the following three issues: the finite element (FE) grid refinement needed to reasonably approximate a solution to our mathematical model, the effect of dynamics, and the number and location of independently activated sub regions of muscles. Initially, we created a FE grid with 91 nodes and 155 FEs. We then refined the grid by dividing each element into 4 new FEs resulting in a refined grid with 336 nodes and 620 FEs. Figure 1 shows the refined FE grid depicting GG and SL, respectively.

An important biological question relating to tongue muscles is—to what extent can sub regions of these muscles be independently controlled? To address this question, we considered four different sets of possible controls. Figure 1 shows the largest number of independently controllable components of GG and SL, respectively, that we have considered.

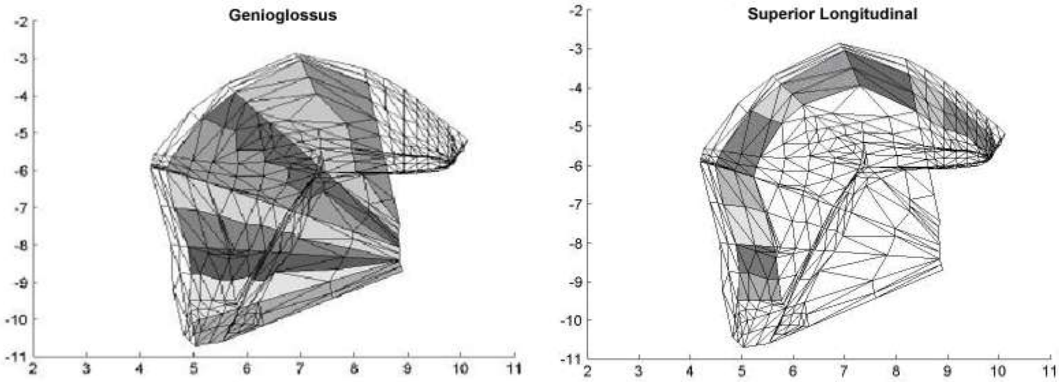

Figure 1: Refined model of midsagittal plane of the tongue showing the maximum number of muscle activations for GG (9) and SL (10).

Theoretically, the least squares approximation to the actual displacements should improve as the number of independent controls increases. However, it is expected that the rate of improvement will sharply decrease at some number. We propose this point as a good estimate of the number of independently activated muscle components used for this particular movement. This is in fact the case as demonstrated in Table 1. Note that the errors are smaller when the accelerations and velocities are included.

Figure 2 depicts the velocity and acceleration trajectories between frames 2-3 and 4-5, the /i/, and the transition to /u/, respectively. The origin of the vector is at the node and the direction and magnitude of velocity or acceleration are indicated by the length and direction of the blue lines. From time 2-to-3 the velocity and acceleration are in the same direction. From time 4-to-5 the acceleration is opposite to the velocity.

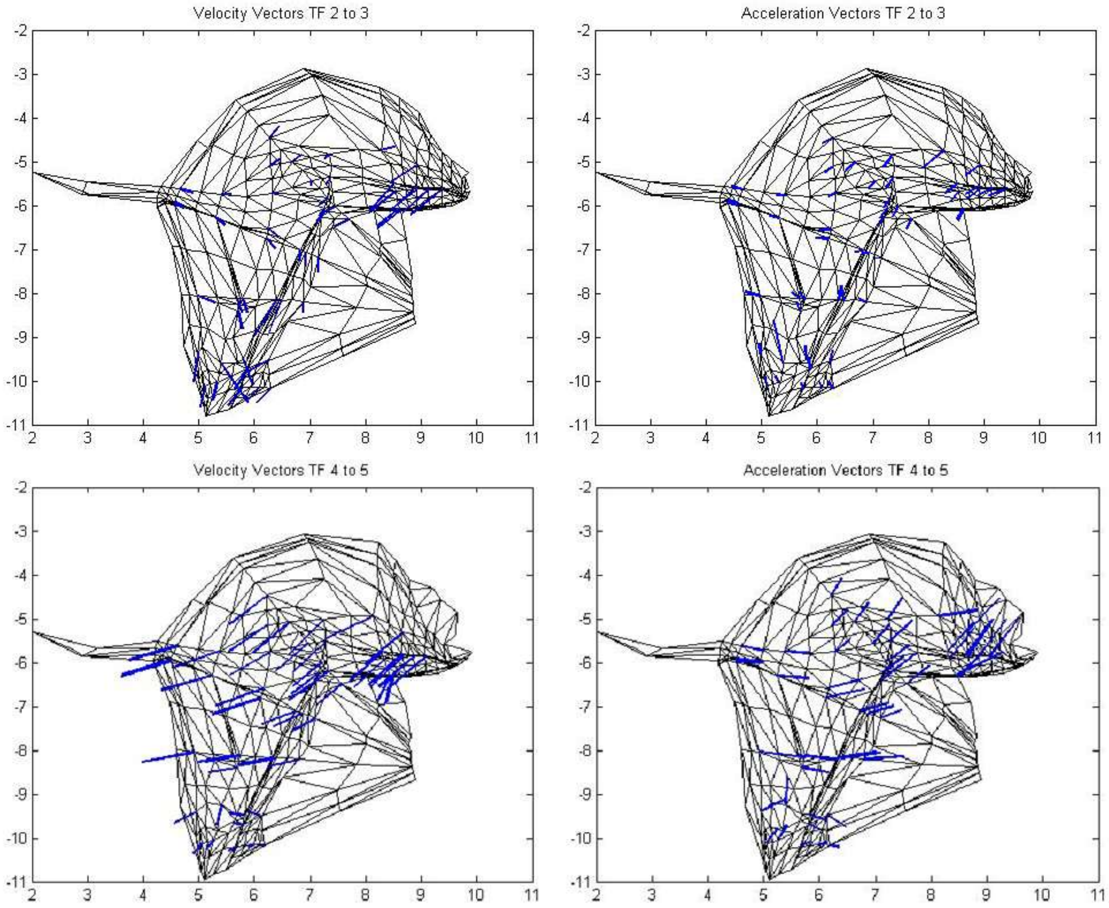

Figure 2: Velocity and Acceleration at end of /i/ (2-3) and during transition to /u/ (4-5).

Velocity and acceleration values are calculated from the tMRI data collected for this subject. Velocity fields and principal strains are shown in Figure 3 for time-frames 4-5. At this point in time the tongue is moving down and back (lower tongue) and straight back (upper tongue). The principal strains indicate that compression is occurring mostly in the anterior and middle tongue, motion in the posterior tongue is due primarily to translation (in this plane). Table 1 shows the least squares error for the static and dynamic models. The errors are greater for time-frames 4-5 where there is large motion, than for 2-3 where there is very little motion. The dynamic model has marginally less error than the static. However, the activations are considerably different in the dynamic vs the static model, as seen in Table 2.

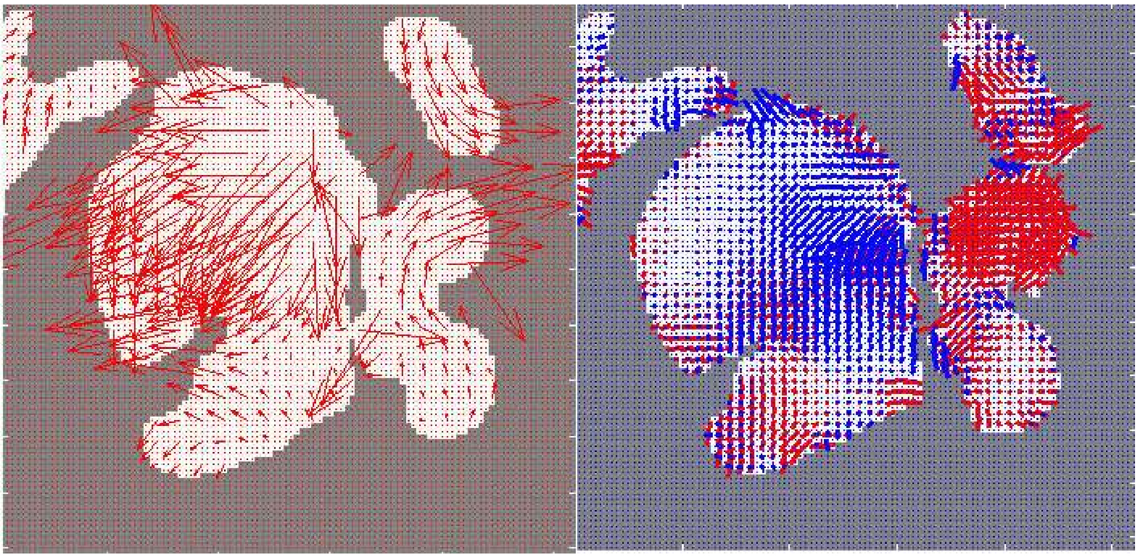

Figure 3: Tagged Cine-MRI data for time-frame 4-5. Velocity fields (left) indicate direction of motion of local region of the tongue and lips. Lips on right. Principal strains (right) show local compression (blue) and expansion (red).

Table 1. The least squared error for different approximations to the acceleration and velocity

Number of muscle sections	Dynamic TF 2 to 3	Static TF 1 to 3	Dynamic TF 4 to 5	Static TF 1 to 5
8	1.1499	1.1511	3.8134	4.1215
12	1.1298	1.1303	3.6245	3.8654
18	1.067	1.0783	2.9464	3.356

First, consider the different time frames: 2-3 (little motion) vs 4-5 (considerable motion) in Table 2. At time frames 2-3 activation of IL, HG and SG indicates onset of backward motion . At frames 4-5, IL and HG substantially increase their activity, while SG becomes inactive. This modeled difference reflects the tMRI data very well (Figure 3). This speaker uses a very low tongue position for /u/, consistent with much HG and no SG. His final /u/ position is not directly posterior to /i/, but back and downward, only slightly above /a/, with very protruded lips.

Next, consider the difference between the muscle activations in the static and dynamic models. The advantage of the dynamic model is the inclusion of velocity and acceleration, allowing less overestimation of muscle activity by including the effects of inertia. Thus we would predict that for faster motions the dynamic model might have smaller activations than the static model. This did in fact occur. Instances of greater activation for the static model occurred 8 times for time-frames 4-5; the reverse occurred once. It can also be seen that in Figure 3-right, tip retraction is executed by both backward and downward compression. The dynamic model attributes more of this compression to IL than the static model, and less to HG.

Table 2 Muscle activations for 8 vs 18 muscles. Numbers are in percentages.

	8 activations		18 activations		8 activations		18 activations	
Independent	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic
muscles	TF 1 to 3	TF 2 to 3	TF 1 to 3	TF 2 to 3	TF 1 to 5	TF 4 to 5	TF 1 to 5	TF 4 to 5
SLtip	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SL1			0.00%	0.00%			0.00%	0.00%
SL2			0.00%	0.00%			0.00%	0.00%
SL3			0.00%	0.00%			0.00%	0.00%
SL4			0.00%	0.00%			0.00%	0.00%
GG1	0.31%	0.27%	0.00%	0.00%	0.00%	1.82%	0.00%	2.91%
GG2			0.33%	0.29%			0.00%	0.00%
GG3			1.22%	1.28%			0.00%	0.00%
GG4	0.36%	0.33%	0.02%	0.00%	1.23%	1.18%	1.53%	0.92%
GG5			0.00%	0.00%			0.00%	0.00%
IL	0.07%	0.61%	1.41%	1.86%	14.22%	16.76%	12.63%	14.26%
Vtip	0.00%	0.00%	0.00%	0.00%	1.36%	0.97%	0.00%	0.00%
V1			0.17%	0.21%			1.42%	0.00%
V2			0.00%	0.00%			0.96%	0.00%
V3	0.00%	0.00%	0.00%	0.00%	2.32%	0.00%	4.43%	2.42%
V4			0.00%	0.00%			2.96%	0.00%
HG	0.96%	0.68%	0.93%	0.83%	7.04%	3.11%	7.59%	3.13%
SG	1.76%	1.24%	0.71%	0.47%	0.00%	0.00%	0.00%	0.00%

Finally, consider the effect of number of activations (model-8 vs. model-18). In TF 2-3, the end of static tongue position for /i/, there is a big difference between 8 and 18 activations. Model-8 calculates activation for SL, GGa and GGp, IL, Va, Vp, HG and SG. Model-18 subdivides these regions. Model-8 indicates similar activity at both GGA and GGp. The additional degrees of freedom in model-18 are able to localize the activity to GG2-3. Because Model-8 cannot account for activity at the boundary between GGa and GGP (i.e., GG3), it distributes the activity equally into both regions. Another difference between the two models is that model-18 retrudes the tongue tip using the more local muscles IL and V1, rather than the more distributed SG activated in model-8. In TF 4-5 HG and IL show the largest activity, with none in SG or SL. Model-18 localizes GG activity to GG1 and GG4 and attributes less activity to IL and more to V3 and 4.

One must be skeptical about these results, especially TF4-5, which involves large motion. We believe that measurement and/or modeling? inaccuracy influences these results, especially for model-18. For example in TF 4-5 the static model-8 shows less activation for Va than Vp (1.36, 2.32). Static model-18 shows the same thing, but the data are not continuous in value (0.0, 1.42, 0.96, 4.43, 2.96), which is inconsistent with our intuition.

Conclusions

The results show continuing improvement in the match between model and data as we refine the finite elements, as we increase the number of independently activated muscle components, and as we add the dynamics. This indicates that we should continue to do all of these things as these are the simplest ways to improve our model. The limitations to these improvements are the increase in the amount of computing needed to solve the augmented problems and the accuracy of the available data. It would be very desirable to decrease the time between successive MRI images and to improve our estimates of velocity and acceleration. There are certainly also issues related to the accuracy of our model. Many of these can be resolved by studying perturbations to the model as well as by implementing the abovementioned improvements to both data and computation.

References:

Antman, S. S., *Nonlinear Problems of Elasticity*, Springer-Verlag, 1995.

Gartner, L. P. & Hiatt, J. L., Extracellular Matrix, chapter 4, pp. 61-63, In Color Textbook of Histology, W. B. Saunders Company, (1997).

Gerard, J. –M., Wilhelms-Tricarico, R, Perrier, P., & Payan, Y., A 3D dynamical biomechanical tongue model to study speech motor control, Recent Research Developments in Biomechanics 1, pp.49-64, (2003).

Iny, D. & Levine, W.S., Mathematical modeling of mammalian muscle and the human tongue, Proc. of MTNS-98, pp.1035-1038, (1999).

Iny, D., Lundberg, A. Stone, M. & Levine, W.S., The dynamics and control of he tongue during speech, Proc. ACC 2000, pp.246-251, (2000).

Levine, W.S., Essex-Torcaso, C. & Stone, M., Controlling the shape of a muscular hydrostat: A tongue or tentacle, Lecture Notes in Control and Information Sciences, Volume 321, New Directions and Applications in Control Theory, (Eds: W. P. Dayawansa, A. Lindquist, Y. Zhou), pp.207-222, (2005).

Miyawaki, K., A preliminary report on the electromyographic study of the activity of lingual muscles, Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, University of Tokyo, 9: pp. 91-106, (1975).

Monti, R. J. , Roland, R. R., & Edgerton, V. R., Role of motor unit structure in defining function, Muscle Nerve, 24: pp. 848-866, (2001).

Slaughter, K., H. Li and Sokoloff, A.J. , Neuromuscular organization of the Superior Longitudinalis muscle in the human tongue I: Motor Endplate Morphology and Muscle Fiber Architecture. Cells Tissues Organs. 181(1): pp. 51-64. (2005).

Smith, K.K. & Kier, W.M., Trunks, tongues and tentacles: moving with skeletons of muscle, American Scientist, 77:pp. 29-35. (1989).

Zajac, F.E., Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control, Critical Rev. Biomed. Eng., 17(4), pp.359-411. (1985).