Towards a 3D Tongue model for parameterising ultrasound data

Alan A Wrench[1,2] Peter Balch[3]

1 Queen Margaret University 2 Articulate Instruments Ltd 3 Analogue Information Systems Ltd
awrench@articulateinstruments.com

ABSTRACT

This paper describes the process and aims of the manual construction of a 3D mesh modelling the tongue, hyoid and mandible. The mesh building process includes the ability to assign muscles to mesh struts which can be independently contracted to nominal lengths in order to test how the mesh deforms. In this way the behaviour of the mesh can be easily and quickly observed and structures can be amended or enhanced.

One such mesh is described which is based on a laminar structure and where the genioglossus is divided into five functionally independent compartments. The model is capable of being deformed to fit midsagittal MRI data of a wide range of distinct articulations by a single speaker and by carefully identifying landmark features and orientation, can also be fitted to ultrasound images for that same speaker.

Keywords: 3D Model Tongue Hyoid ultrasound MRI

1. INTRODUCTION

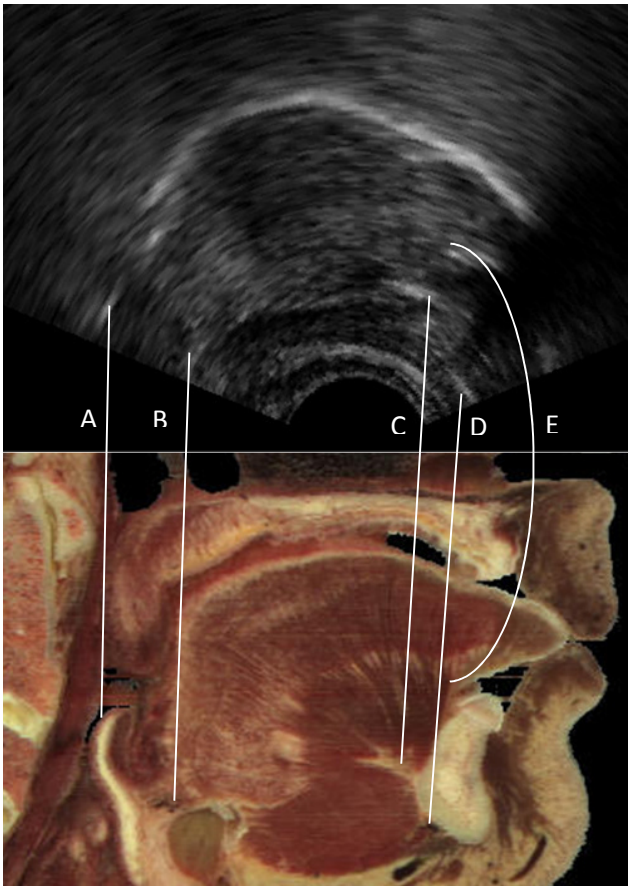
Estimating complete tongue surface contours from ultrasound images is difficult because the images are noisy, parts of the contour are not visible and image artefacts caused by internal reflections can sometimes be mistaken for part of the tongue surface contour.

This is true for manual estimation and even more so for automated contour detection. The classic means of estimating a signal from noisy partial data is to apply a model. For example, when a normal distribution curve is fitted to experimental results.

To date, the initial processing step applied to ultrasound tongue images has been non-parametric data reduction: e.g. by performing manual or semi-automatic edge detection or by principal component analysis to the whole or part image. A second stage of data reduction is then applied using curve comparison of some kind or displacement of the contour along a one dimensional intersecting axis.

Without a model however, it is sometimes difficult to separate noise and imaging artefacts from the underlying tongue contour and not possible to estimate missing data such as the tongue tip contour. Our purpose in developing a 3D tongue model is to provide a set of parameters, constrained by properties such as volume preservation and muscle stiffness, which describe all observed midsagittal tongue contours. If we are successful then the objective of any image processing then becomes one of estimating these parameters by fitting the midsagittal contour of the 3D model to the ultrasound image. It is worth pointing out that such model parameters which correspond to muscle contractions may also turn out to be very useful in analysing speech production.

Figure 1: Midsagittal ultrasound image compared with Visible Human Project® female. A. Tip of epiglottis B. Epiglottic valecula C. Short tendon D. Mandible



2. FEATURES OF THE ULTRASOUND TONGUE IMAGE

Interpretation of ultrasound tongue images and extraction of features typically focusses on the bright regions of the image near where the tongue surface is expected to be. To fit a model to the ultrasound it will be necessary to identify other structures evident in the

image and not just the surface contour. Careful comparison with midsagittal slices of female speaker from the Visible Human Project® (VHP) 0 reveals other important features. A key salient feature in many if not all ultrasound images is the short tendon (Fig 1C) that acts as an extended insertion point for the muscle bundles of the genioglossus and connects to the mandible at a pair of protuberances known as the superior mental spines. Anterior and inferior to the mental spine there is very often another bright reflection in the ultrasound image just beneath the shadow of the mandible (Fig 1D). This appears to correlate with the movement of the mandible but the exact cause of this reflection is not clear. There is often a reflection from the floor of mouth. The most posterior end of the floor of mouth (Fig 1E) is a guide to the most anterior part of the genioglossus.

Ultrasound images from children (such as Fig. 1) do not have such a pronounced hyoid shadow. This may be because the hyoid grows and ossifies with age, absorbing more ultrasound and producing a larger shadow source. This being the case, the following features may be obscured by the hyoid shadow in adult images: Fig 1A shows the tip of the epiglottis; Fig 1B shows a bright reflection from the base of the tongue, the valecula. Since the tongue surface at the root is almost parallel to the ultrasound beam there is very little reflection from the tongue surface itself and only the valecula and/or hyoid shadow indicates the location of the base of the tongue.

3. DEVELOPING A 3D TONGUE MODEL

3.1 Background

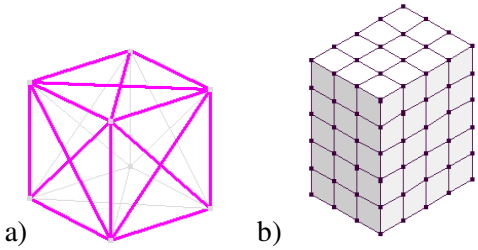
Various 3D biomechanical models of the tongue have been constructed. Some models [3][9][13][14] consist of geometrical structures with a limited number of elements to reduce the computational load when running simulations. In the mid-nineties Wilhelms Tricarico [26] modelled the properties of muscle tissue and the VHP female provided a source for tracing muscle fibres and fitting a hexahedron mesh structure. Others [4][5][6][8][24] have developed models with a very similar mesh structure.

Finite element modelling (FEM) is the favoured dynamic deformation modelling approach as it allows physical tissue properties such as incompressibility and the muscle contraction function to be defined in detail. FEM breaks a 3D structure into a finite number of discrete volumes (elements) and assigns properties to these elements. In practice, the fidelity of this approach depends on the computational implementation, on how fine the elements are and how many nodes define each element. For computationally efficient FEM modelling, there is a tendency for the model to become stiff and lock up [20], not due to the modelled physical properties but because of numerical discretisation.

3.2 Mesh building

Our mesh starts out very simply as a set of cubes with *atoms* at each of the 8 vertices. *Struts* connect the *atoms* along each edge and additional cross bracing *struts* connect atoms on each face (Fig 2).

Figure 2: a) cubic unit with 8 vertices 12 edge *struts* and 12 additional cross bracing *struts*, 2 on each face. b) initial 3D array of hexahedrons



Atoms can subsequently be manually moved in an unconstrained manner either individually or in groups. [Note that this means that each quadrilateral face of the originally cubic polyhedron can become two triangular faces]. Extra layers can be added to any extremity of the mesh array, cloning the shape of the layer from which it is extended. In this way it is possible to start from a single layer and “grow” the mesh layer by layer. *Atoms* can be merged such that two separate arrays can be joined together. So for example, the mandible can be created separately to the tongue and then connected.

Once the mesh is formed, any *strut* can be assigned to a “muscle”, which is defined to have a linear range of elasticity defined by an "at rest" and a maximum contraction (30% rest length) elasticity.

There is an optional symmetry plane. If this is invoked then the structure and deformation is mirrored across this plane.

3.3. Deformation Modelling

To deform the mesh we have chosen not to use FEM. Our aim (at least initially) is to create an environment in which a 3D mesh can be constructed easily and tested to see how it deforms as muscles are contracted. We are less concerned with the dynamics of the model under contraction but more focused on the steady state equilibrium that results from those contractions. Consequently we have chosen a simple and robust 3-stage method based on iterative movement of the position of the mesh vertices to satisfy firstly a set of *strut* length constraints,

secondly a polyhedron volume preservation constraint and thirdly a rigid body constraint. Note that the vertices of the mesh do not represent lumped masses and we do not calculate velocities. We iterate these three stages until the model reaches a stable equilibrium for the specified muscle target lengths.

The 1st stage deformation is based on Hooke's law but where the elasticity constant varies linearly as the target length of the *strut* changes (Eqn 2). The Hookean force defined in Eqn 1 is divided by 2 because we assume each *strut* shares its force equally between the two *atoms* on either end of the *strut*.

$$(1)\quad \boldsymbol{F_{Atom}} = \sum_{n=1}^{N_{Atom}} e_{strut\,n} \frac{\Delta \boldsymbol{l}_{strut\,n}}{2}$$

$$(2)\quad e_{strut\,n} = e_{min\,n} + \left(e_{rest\,n} - e_{min\,n}\right) \frac{(l_t - l_{min})}{(l_{rest} - l_{min})}$$

$$= elasticity\ of\ nth\ strut\ at\ target\ length$$

$$\boldsymbol{F_{Atom}} = Vector\ force\ applied\ to\ atom$$

$$N_{Atom} = Number\ of\ struts\ connected\ to\ atom$$

$$\Delta \boldsymbol{l_{strut\,n}} = (l_c - l_t)\widehat{U}_{strut\,n}$$

$$l_{rest} = rest\ length\ ; \qquad e_{rest_n} = elasticity\ at\ rest$$

$$l_{min} = 30\%\ rest\ length\ ;\ e_{min\,n} = elasticity\ at\ l_{min}$$

$$l_c = current\ length\ ; \qquad l_t = target\ length$$

$$\widehat{\boldsymbol{U}}_{strut\,n} = unit\ vector\ in\ direction\ of\ strut\ n$$

$$(3)\quad \Delta \boldsymbol{p_{Atom}} = -\boldsymbol{F_{Atom}} / \sum_{n=1}^{N_{Atom}} e_{strut\,n}$$

$$= Vector\ position\ change\ to\ equalise\ force$$

Application of Hooke's law as defined in equations 1-3 does not guarantee volume preservation as required by a muscular hydrostat. To counteract a tendency for the model to compress when *struts* are contracted, a separate volume pressure force is applied to each polyhedron in order to maintain its original volume. For each polyhedron, the difference between the original and current volume is calculated and a positional change is determined for each *atom* in order to restore the original volume; that vector is added to the one arising from stage one. To improve the stability of the simulation, an ad hoc limit is set on the amount of pressure based positional change in any one direction for each iteration.

The hyoid, short tendon and mandible form three separate rigid bodies integral to the movement of the tongue. Modelling of full 3D rigid body movement is complex and computationally expensive. An easier proposition is to calculate the movement in 2D. To do this we implement a symmetry plane corresponding with the midsagittal plane and restrict the rigid bodies to being symmetrical. Each "rigid" *atom* is constrained to move in its own fixed plane which is parallel to the midsagittal plane. The tongue model as a whole is then necessarily also symmetrical about the midsagittal plane.

3.4 Biomechanical structure

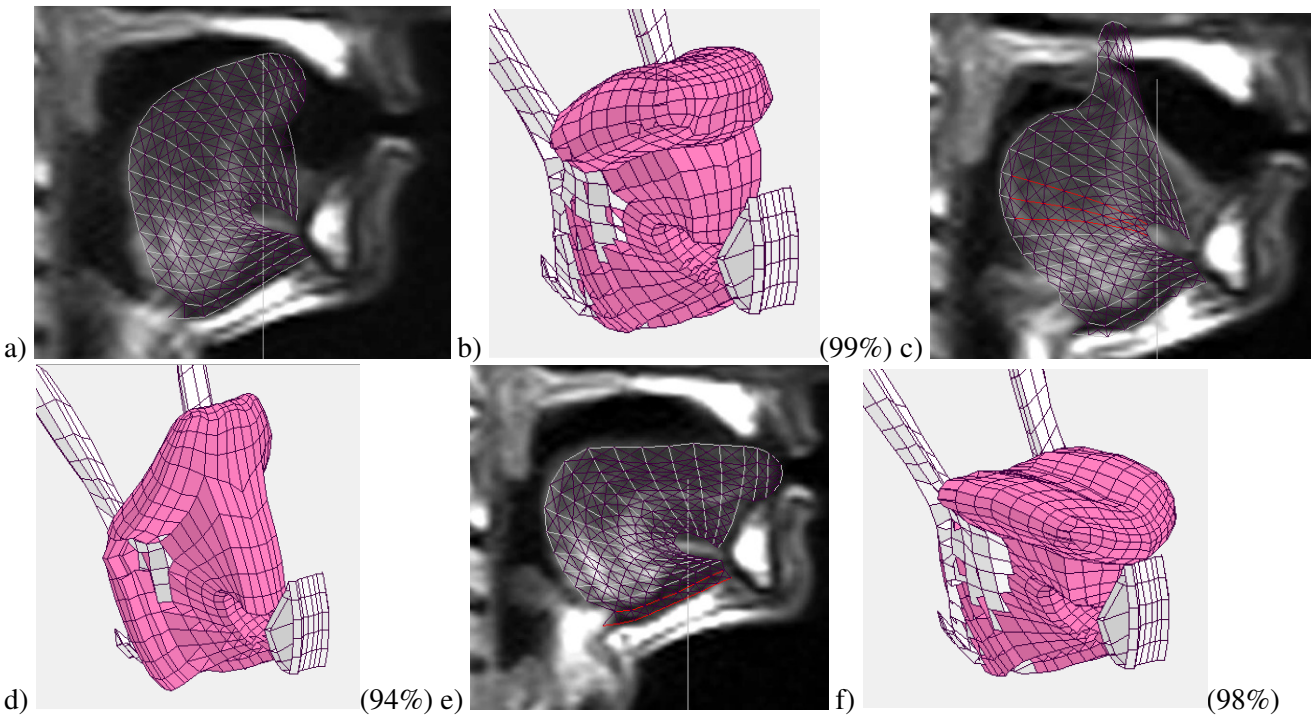
The current 3D tongue model is based on observations from the VHP female with reference to EMG and anatomical studies [2][10][15][22][23][16]. Miyawaki [17] found distinct EMG signals from 5 different regions of the genioglossus during speech. In a separate experiment they stimulated four prominent twigs of the hypoglossal nerve of a dog which inserted into the genioglossus and observed that four distinctly localised contractions of the genioglossus occurred. More recently Mu [18] dissected a number of human tongues and identified 5 or 6 twigs supplying what they term the oblique portion of the genioglossus with a further 2-3 twigs supplying the remaining horizontal portion. Each twig has at least the potential to contract a distinct region of the genioglossus. A further requirement for a functional compartment of a muscle is that it has a unique insertion point [21]. We undertook careful segmentation of the VHP female which indicated with some ambiguity between 5 and 9 muscle bundles inserting into the short tendon. There appear to be 5 bundles in the midsagittal section with possibly further bundles with origins off the midline. We assume 5 segments in our modelling of the genioglossus presented here. All other tongue/hyoid muscles are modelled in this paper with a single functional compartment. Mylohyoid and digastric muscles are not yet modelled.

3.5 Fitting the 3D model to midsagittal MRI images

It has become clear from the process of developing this model that the shape of the tongue when all muscles are relaxed is important. The shape of the VHP female cadaver is used in other models [3][12][19][25] as the starting point. There are at least three problems with this. Firstly, the visible human head was removed from the body and so the sternohyoid and omohyoid have been severed. Secondly, like all other parts of the anatomy, they vary in size and shape between people. Thirdly, a cadaver muscle elasticity and *in vivo* "at rest" muscle elasticity may differ significantly [4]. For the model described here, a number of guesses at the rest state mesh shape and position were made until all of the various tongue shapes could be modelled from the same rest position by contraction of the muscles alone.

The model was fitted to a range of midsagittal MRI vocal tract images from a single speaker by altering the contraction level of modelled muscles. Examples of the model fitting are shown in Fig 3.

Figure 3: Midsagittal slice of the 3D tongue model fitted to MRI image and corresponding full 3D model. Shows a) & b) – /r/; c) & d) – extreme retroflex; e) & f) –/s/ (/i/ context). The number in brackets is total volume of the model expressed as a percentage of the "at rest" model.



4. DISCUSSION

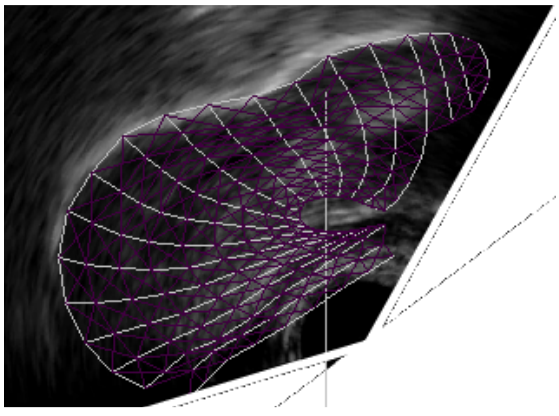
The mesh design and muscle assignment may easily be modified to investigate whether the genioglossus has more off-midline functional subcompartments and whether other muscles also have functional subcompartments by testing if there is an improvement in the fit to the MRI data.

There are a number of simulation improvements that are still required. For example, the volume preservation implementation is suboptimal and under review. Collision with the palate has been modelled but is not included in this paper as it needs further work. Only a section of anterior mandible is modelled not its entirety. The hyoglossus is particularly difficult to mesh as it does not follow the laminar structure of the intrinsic muscles and yet it wraps around the exterior of the tongue; so it will be revised.

In order to allow the model to expand and contract to fit the various extreme articulations revealed by midsagittal MRI we found by trial and error that around a 15:1 elasticity ratio between the rest length and the minimum nominal contracted length (30% rest length) worked best. This is in broad agreement with the 18:1 Young's modulus ratio measured *in vivo* by Duck [7] for human muscle tissue.

The current 3D model accounts for the change in size and shape of a range of MRI images. It is promising to see that the distortion of the shape of the geniohyoid muscle follows that in the MRI image when this is not explicitly controlled and results from the relative positions of the mandible and hyoid and the configuration required to match the upper surface of the tongue. It is also promising that tongue grooving occurs from fitting the midsagittal slice with no explicit manipulation to form the groove.

Figure 4: Model deformed to fit dark /l/ in ultrasound



When fitting a model such as this to successive frames of ultrasound (Fig 4) we should expect a smooth pattern of dynamic change in modelled muscle length. If sudden changes in muscle length are required this would indicate an error in the model. We are continuing to modify the model in this way.

5. ACKNOWLEDGEMENTS

This work is funded by Articulate Instruments Ltd. The MRI data is an output of EPSRC funded project EP/I027696/1. MRI imaging protocol (Conventional SE, variant SK\SP Option PFP, ST 8.0, RT 400, ET 52 voxel 1.72x1.72x8mm) was selected by Scott Semple and recorded using a 3T Seimens VERIO MRI system at Edinburgh CRF, Clinical Research Imaging Facility, University of Edinburgh.

6. REFERENCES

[1] Ackerman, M,J. Visible Human Project. McGraw-Hill 2004 Yearbook of Science & Technology. New York: McGraw-Hill. 2004. p. 369-72

[2] Baer, T., J. Alfonso, and K. Honda, 1988. Electromyography of the tongue muscle during vowels in /epvp/ environment. *Ann. Bull. RILP*., Univ Tokyo, 7: p. 7-18.

[3] Baker, A. 2008. A biomechanical model of the human tongue for understanding speech production and other lingual behaviors. Doctoral Dissertation, University of Arizona.

[4] Buchaillard S., Perrier P. & Payan Y. 2009. A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *JASA*, 26(4), pp. 2033–2051.

[5] Dang, J., Honda, K., 2004. Construction and control of a physiological articulatory model. *J. Acoust. Soc. Am*., 115(2): p. 853-870.

[6] Dang, J.Fujita, S.Murano, E.Stone, M. 2006. Observation and simulation of Large-scale deformation of tongue. *Proc. ISSP06*. Brazil.

[7] Duck, F. A. 1990, Physical Properties of Tissues: A Comprehensive Reference Book (Academic Press, London, UK).

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

[9] Hashimoto, K. and Suga, S., 1986. Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue. *J. Acoust. Soc. Jpn.(E)*, 7(1), pp. 39-46.

[10] Honda, K., Murano, E.Z., Takano, S., Masaki, S., Dang, J., 2013. Anatomical considerations on the extrinsic tongue muscles for articulatory modelling *JASA*, vol. 133, issue 5, p. 3607

[11] Kakita, Y., O. Fujimura, and K. Honda, 1985 Compuation of mapping from muscular contraction patterns to formant patterns in vowel space., in Phonetic linguistic, V.A. Fromkin, Editor. Academic Press: New York. p. 133-144.

[12] Kajee Y., Pelteret J.P., Reddy B.D., 2013. The biomechanics of the human tongue. *Int J Numer Method Biomed Eng.* Apr;29(4):492-514

[13] Kiritani, S.Miyawaki, K.Fujimura, O.Miller, J. 1976 A computational model of the tongue. Ann. Bull. R. I. L. P. Univ. Tokyo.;10:243–251..

[14] Levine, W.S., Torcaso, C.E. and Stone, M., 2005, Controlling the shape of a muscular hydrostat: A tongue or tentacle, in *New Directions and Applications in Control Theory*, W.P. Dayawansa, A. Lindquist, and Y. Zhou, eds., Springer-Verlag, New York, 207–223.

[15] MacNeilage, P. and G. Sholes, 1964. An electromyographic study of the tongue during vowel production. *J. Speech Hear. Res*., 7: p. 209-232.

[16] Miyawaki, K., 1974. A study on the musculartureof the human tongue. *Ann. Bull. RILP*, 8: p. 23-49

[17] Miyawaki, K., 1975. A preliminary report on the electromyographic study of the activity of lingual muscles. *Ann. Bull. RILP*, 9: p. 91-106.

[18] Mu, L., & Sanders, I. 2012. Human tongue neuroanatomy: Nerve supply and motor endplates. Clinical Anatomy, 23, 77-791

[19] Pelteret J-PV and Reddy BD, 2014 Development of a computational biomechanical model of the human upper-airway soft-tissues towards simulating obstructive sleep apnea. Clinical Anatomy 27 182--200

[20] Rohan PY, Lobos C, Nazari MA, Perrier P, Payan Y., 2014 Finite element modelling of nearly incompressible materials and volumetric locking: a case study. *Comput Methods Biomech Biomed Engin*., 17 Suppl 1:192-3

[21] Schieber M.H., Gardinier J, Liu J., 2001 Tension distribution to the five digits of the hand by neuromuscular compartments in the macaque flexor digitorum profundus. *J Neurosci*Mar 15; 21(6):2150-8.

[22] Takano, S. and K. Honda, 2007 An MRI analysis of the extrinsic tongue muscles during vowel production. *Speech Communication*, 49(1): p. 49-58.

[23] Takemoto, H., 2001. Morphological analysis of the tongue musculature for three dimensional modeling. *Journal of Speech and Hearing Research*, 44: p. 95-107.

[24] Vogt, F., Lloyd, J., Buchaillard, S., Perrier, P., Chabanas, M., Payan, Y. and Fels, S., 2006 Efficient 3D Finite Element Modeling of a Muscle-Activated Tongue, *Biomedical Simulation*, vol. 4072, pp. 19-28.

[25] Wang, Y., 2009 The Modelling of Interlacing Muscles in Tongue Movement, Masters Thesis, The University of Auckland.

[26] Wilhelms-Tricarico R., 1995 Physiological modeling of speech production: methods for modeling soft-tissue articulators *JASA*. May;97(5 Pt 1):3085-98