A BIOMECHANICAL MODEL OF THE HUMAN TONGUE FOR UNDERSTANDING SPEECH PRODUCTION AND OTHER LINGUAL BEHAVIORS

by

Todd Adam Baker

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2 0 0 8

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation
prepared by Todd Adam Baker
entitled A biomechanical model of the human tongue for understanding speech production and other lingual behaviors
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

_____ Date: 8/29/2008
Michael Hammond

_____ Date: 8/29/2008
Brad Story

_____ Date: 8/29/2008
Jonathan Vande Geest

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 8/29/2008
Dissertation Director: Michael Hammond

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Todd Adam Baker

ACKNOWLEDGEMENTS

I have been helped by the generosity and patience of many people. Here I would like to acknowledge some notable contributors.

Mike Hammond, my advisor and committee chair, has been excellent. Mikes support and patience have been exemplary. I have also had a wonderful interaction with my other two committee members, Brad Story and Jonathan Vande Geest. These gentlemen were both generous with their time and their counsel.

From the Department of Physiology, Andy Fuglevand and Fiona Bailey have both been extremely helpful. Their patience with a linguist grappling with the rudiments of neuromuscular physiology has been greatly appreciated.

Special thanks go to Diana Archangeli, my employer of four years, who provided an ideal environment in which to learn how to do science. It was unfortunate that her sabbatical year coincided with my dissertation year. I trust that our collaboration is not over.

The High Performance Computing office (particularly Lucy Caruthers, Todd Merritt, and Kathleen Bowles) has been an enormous help to me. Marvin Landis of Scientific Visualization was instrumental in connecting me with the software I needed to perform the anatomical study.

I would like to extend thanks to the donor of the female body used in the Visible Human Project. Thanks also to the donors of other cadavers I have worked with.

I would like to acknowledge, quite generally, those who have invested the time and energy to make so much valuable tutorial material online. I have benefited from professors who have posted course notes, and from those who have contributed to Wikipedia and similar resources.

Academic life has of course been a small part of my actual life over the past years; the best times have been with my wife Amy, and my sons Abraham and David. Amy has been wonderfully patient with me when stresses from work spilled over into home life, or when a burst of inspiration would take me away from the family for an evening. Her love and support has meant a lot. Seeing Abraham and David grow up, and falling deeper in love with them day by day, has been the greatest blessing of these past years. The fulfillment afforded by family life is incommensurate with that afforded by academic life.

This work would not have been possible—and if it had been possible it would not be interpretable—outside of the context of the redemptive work of Jesus Christ. I owe Him everything: my being, my accomplishments, my salvation. I do not thereby claim infallibility for this work, but rather rejoice that the Lord has shown Himself willing in all parts of life to partner with imperfect creatures. I heartily and humbly look forward to being the weaker partner in many future collaborations.

DEDICATION

This dissertation is dedicated to the people I hope to be better able to serve because of its completion.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF TABLES

LIST OF TABLES – *Continued*

ABSTRACT

A biomechanical model of the human tongue was constructed, based upon a detailed anatomical study of an actual cadaver. Data from the Visible Human Project were segmented to create a volumetric representation of the tongue and its constituent muscles. The volumetric representation was converted to a smooth NURBS-bounded solid model—for compatibility with meshing algorithms—by lofting between splines, the vertices of which were defined by the coordinates of a smoothed triangular mesh representation. Using a hyperelastic constitutive model that allowed for the addition of active stress, the model deforms in response to user-specified muscle activation patterns. A series of meshes was created to perform a mesh validation study; in the validation tests performed, a 245,223-element mesh was found to be sufficient to model tongue behavior.

Systematic samples of the behavior of the model were collected. Principal component analyses were performed on the samples to discover low-dimensional representations of tongue postures. Statistical models (linear regression models and neural networks) were fit to predict tongue posture from muscle activation, and vice versa. In all tests, it was found that a relatively small sample of tongue postures can be used to successfully generalize to larger data sets.

Finally, a variety of specific tests were performed, based on claims and predictions found in previous literature. Of these, the claims of the muscular hydrostat theory of tongue movement were best supported. Simulations were also run that simulated lingual hemiplegia. It was found that substantially different muscular activation patterns were required to achieve equivalent postures in a hemiplegic tongue, relative to a normal tongue.

The present work advances the state of tongue modeling in several respects. It is based on a detailed anatomical study of an actual specimen. It is also the first

tongue model to have been subjected to a mesh validation study. It is anticipated that the model will be of use in further research into the mechanical properties of the tongue, and the relationship between tongue muscle activation and the acoustical output of the vocal tract.

CHAPTER 1

Introduction

> All kinds of animals, birds, reptiles and creatures of the sea are being tamed and have been tamed by man, but no man can tame the tongue. It is a restless evil, full of deadly poison.
>
> —James 3:7–8 (NIV)

The tongue is unique among the anatomical structures of the human body. It has many functions, playing a central role in respiration, mastication, deglutition, and speech production. It also has a number of unique anatomical and physiological traits. Foremost among these is the mechanism of tongue movement. The purpose of most skeletal muscle in the body is to move or stabilize the body by producing forces and torques on the skeleton. Consequently, in biomechanical research skeletal muscle is frequently not even modeled, except as a force-generating element (e.g., Yamaguchi, 2001). In the tongue, however, the primary purpose of muscular contraction is to deform the tongue itself. In speech production the tongue deforms to modulate the flow and acoustic resonances of air through the vocal tract. In mastication the tongue is used to transport food boli around the appropriate occlusal surfaces, and for maintenance of the shape of the bolus. In deglutition the tongue is used to propel the bolus into the esophagus. All of these activities rely on the tongue's ability to change its own shape. The purpose of this dissertation is to further understanding of how the tongue changes shape in response to muscular contraction.

Approximately every quarter century in recent times, researchers have remarked that our knowledge of the tongue is embarrassingly limited. In a pioneering anatomical study, Abd-El-Malek remarks that the tongue had received very little attention from anatomists up to that point.

> Though it is an important organ, yet it has not, so far, been given an

> equal share in the detailed anatomical studies allotted to other organs in the body. (Abd-El-Malek, 1939, 14).

Not long after, Strong remarked that phoneticians' knowledge of the configuration of the vocal tract during different speech sounds was not coupled with an understanding of the underlying muscular activity that determined the shape of the vocal tract.

> Neither the anatomist nor the physiologist has contributed much to our knowledge of the function of the specific muscles in speech production. Linguists and phoneticians have busied themselves with one phase of speech, that of showing the configuration of the boundaries of the cavity of the buccopharynx in the production of a resonating chamber for the different *vowels*. What they have failed to do is to show which specific muscles are responsible for these configurations. (Strong, 1956, 68).

Miyawaki subsequently observed that, in spite of the enormous importance of the tongue in speech production, it had received very little attention from speech scientists.

> The tongue, in spite of always being recognized as the primary—and absolutely indispensable—organ for the articulation of speech sounds, seems to be somewhat left behind in the study of speech. One of the reasons for this may be that the anatomical structure of the tongue is far from simple and that the dynamics of lingual movement are as complex as the tongue itself. (Miyawaki, 1974, 29).

Most recently, Sokoloff has noted that the complexity of the tongue has prevented researchers from gaining much specific knowledge of tongue motor unit physiology.

> Despite the importance of the mammalian tongue in normal oro-motor behavior and the association of tongue dysfunction with many human clinical syndromes ...., we lack even a basic understanding of the physiological organization of the fundamental output elements of the tongue

> motor system, i.e., the hypoglossal (tongue) motor units (MUs). Study of hypoglossal MU organization is hampered by the complexity of tongue muscular architecture. (Sokoloff, 2000, 827)

The specific aim of this dissertation is to fill one particular gap in our knowledge of the tongue: the relationship between the motor commands the tongue receives and the posture it assumes in consequence. A biomechanical tongue model calculates the posture of the tongue as a function of applied loads. Loads can be applied from without—such as when a tongue depressor is pressed into the tongue—or from within—such as when the muscles of the tongue contract, generating stress along the direction of the muscle fiber.

1.1 Overview of the model

A biomechanical model permits quantitative studies to be performed on the relationship between the control signals the tongue receives and the posture it ultimately assumes. That is, the model presented in this dissertation is equipped to provide answers to questions of the variety, What happens when genioglossus contracts? What muscles are involved in protruding the tongue? What are the muscular contractions associated with the /i/ tongue posture? Such questions are easy to ask, but difficult to answer. Making generalizations about the function of particular muscles is difficult primarily because of the complex geometry of the tongue. Tongue muscles are all characterized by a complex geometric organization of muscle fibers. In addition, muscles of the tongue interleave and interdigitate with one another—sometimes running orthogonally to one another. These complexities make it very difficult to predict accurately the movements associated with a particular muscle or set of muscles. This has not prevented researchers from making predictions about the mechanical functions of the tongue muscles.[1] The complexity of the tongue, however, and the desire for quantitative rather than impressionistic results, motivate the

[1] Indeed, in almost all cases these predictions have turned out to be either incorrect or overly simplistic; see Section 6.2.

creation of a biomechanical tongue model.

In the present work, the tongue is modeled using the finite element method. Further detail is presented in Chapter 4, but a schematic overview is given here. In finite element analysis, complex geometry is approximated by a set of simpler geometrical elements (in this case, tetrahedra), the finite element mesh. The complex geometry is represented with simple geometric shapes, the mechanical properties of which can be determined with relative ease. Different elements of the mesh can be assigned different fiber orientations and different contractile properties. This permits the specification of the complex muscular geometry of the tongue, and a variety of mechanisms for controlling the tongue.[2]

1.2 Importance of a tongue model

Creation of a biomechanical tongue model provides for a shift away from qualitative statements about the role of particular muscles in lingual behavior (cf. Chapter 2), toward quantitative analyses of tongue behavior. Coarse generalizations about the mechanical consequence of muscular contraction can always be made: after all, muscles only contract in one direction. But such generalizations are unsatisfactory for almost any further research. They cannot be converted to quantitative results without the arbitrary intervention of a researcher, which casts doubt on the validity of any subsequently derived generalizations. A quantitative model, on the other hand, can interface with many others types of models, and has many potential applications. This section offers discussion (really, suggestions) about potential applications

[2]For further information about mechanical modeling, Özkaya and Nordin (1991) provides an excellent introduction to physics and mechanics, with special attention to biomechanics. This work is quite accessible to those who, like the author, have modest mathematical backgrounds. It also includes very clear introductory material relating to the mechanics of deformable bodies. Yamaguchi (2001) is an excellent text for understanding physics mathematically, specifically for rigid body dynamics; familiarity with linear algebra would be a benefit before approaching the text, but it is extremely readable. Lai et al. (1993) provides a more technical introduction to the finite-element method. Finally, Fung (1993) is a classic text in biomechanics. Chapters 2 and 9 are particularly relevant to the present work.

of the model in the speech sciences. The many other potential applications of the model, such deglutition or respiration research,[3] are not considered here. To illustrate the wide range of potential applications of the tongue model, three different levels of abstraction relevant to speech research are considered: low-level phonetic research, theory-oriented phonetic research (representing an intermediate level of abstraction, in that it seeks to explain general properties of language, rather than physical facts), and high-level phonological theory.[4] The purpose of this section is not to provide an exhaustive list of applications, and is certainly not to establish the scope of the dissertation. It is rather to motivate the purpose for creating a tongue model, and to suggest some potential directions for future research.

1.2.1 Phonetic research

According to the acoustic theory of speech production (Fant, 1970), the vocal tract acts as an acoustic filter for a source signal generated by the lungs and larynx. The properties of the filter depend on the shape of the vocal tract, which depends largely on the shape of the tongue. The relationship between the anatomical structures and the sounds they can generate are well understood: given a vocal tract configuration, the source can be simulated and an acoustic signal can be determined (e.g., Story et al., 1996a; Boersma, 1998). Speech production models presently rely on a geometrical model of the vocal tract as input. Detailed geometric models of the tongue or vocal tract can be created from behaving subjects using non-invasive techniques. Three-dimensional models have been created for both static (Stone and Lundberg, 1996; Lundberg and Stone, 1999) and dynamic (Yang and Stone, 2002) tongue postures. Neither are such models limited to the surface anatomy of the

[3]Other applications of the model include, but are not limited to, understanding speech disorders (Kent, 2004), understanding sleep apnea (Saboisky et al., 2007), planning and practicing surgery with computer models (Rodrigues et al., 2001), and understanding problems in tongue movement following surgery (Hirano et al., 1992; McConnel et al., 1994; Stachler et al., 1994; McConnel et al., 1998).

[4]The terms "high-level" and "low-level" are not superlative or perjorative. They refer solely to levels of abstraction in analysis.

tongue: imaging techniques incorporating internal deformation (Liu et al., 2006) and changes in tongue fiber orientation (Shinagawa et al., 2008) are also available.

Although appropriate for many purposes, these models neglect the influence of motor control on speech production. Since the relationship between motor commands and vocal tract shape is not understood, a biomechanical model of the tongue fills a significant gap in the current understanding of speech production. A biomechanical model of the tongue can produce a (potentially time-varying) geometrical model in response to a control signal, which can subsequently be used to generate an acoustic signal. The input of a speech production model can then be a sequence of motor commands. A full model of speech production would be helpful, for instance, in understanding the articulatory motivation for certain vowel space configurations. Geometrical descriptions of the vocal tract can be determined either from non-invasive imaging or, in some cases, acoustic information (e.g., Story and Titze, 1998).

A biomechanical tongue model is also important in creating a theory of articulatory ease. If one wished to calculate the articulatory distance between a set of vocal tract configurations, it would be possible to calculate the difference between two geometrical models, and thereby quantify the "cost" of moving from one configuration to another. Such measurements might not, however, be straightforwardly related to quantities that one might think relevant for planning articulation, such as required articulatory effort or motor complexity. This is because the muscular activity required to effect one tube perturbation versus another is not well understood. One might imagine that it is easy to perturb the vocal tract near the alveolar ridge, but not near the uvula, for instance, or vice versa. Such an intuition cannot be represented with a purely geometric model. With a biomechanical model, however, one could determine the muscular activities, and either quantify or invalidate the intuition.

1.2.2 Phonetic theory

Understanding the relation of motor commands to the acoustic output of the model is important for more abstract approaches to phonetics as well. For instance, Stevens (1989) presents the quantal theory of speech, in which phonetic categoricity arises as a result of the non-linear relationship between articulatory parameters and acoustic outputs. Such a relationship is schematized in Figure 1.1, with the articulatory parameter on the abscissa and the acoustic/auditory parameter on the ordinate. Variation in the articulatory parameter in regions I and III has relatively little acoustic consequence. Conversely, variation in region II has a large acoustic effect. By hypothesis, quantal distinctions in sounds arise as a result of the relative acoustic stability of certain regions of the articulatory parameter space (i.e., I and III), as compared with other relatively unstable regions (II).

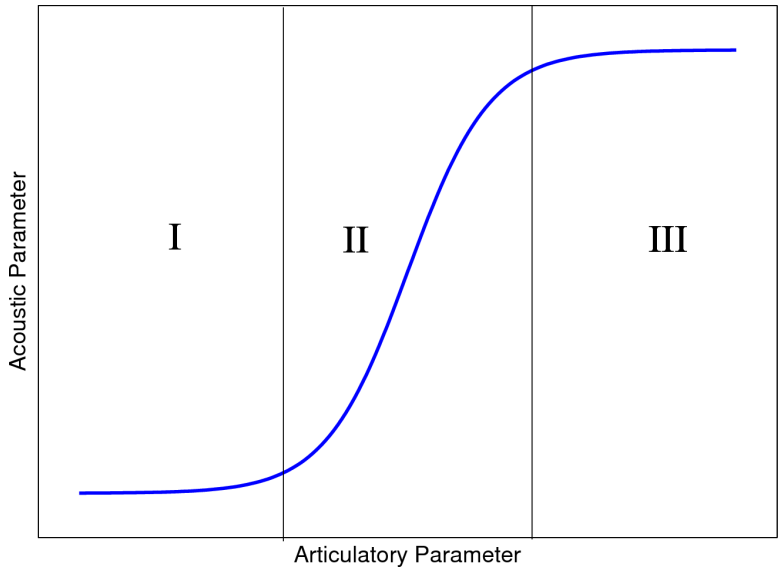

Figure 1.1: Schematization of an articulatory/acoustic relationship, after Stevens (1989)'s Figure 1.

Stevens' examples of articulatory parameters are based on simple tube models of the vocal tract: the position of a constriction relative to the source, or the diameter of the constriction. That is, the articulatory parameter is a geometric description of the vocal tract. There are a couple of ways in which a more detailed knowledge of

articulation would be helpful in realizing the full explanatory potential of the model. The structural complexity of the tongue, and the large number of motoneurons that control it (see Chapter 2), suggest that skilled tongue movement is not a trivial task. Even if the highest-level phonetic planning were done with simple articulatory representations, there is a great deal of complexity in the execution of movement. It would be surprising if these complexities had no bearing on the articulatory-acoustic relationship.

A biomechanical model also has a role in the study of speech motor control. A number of control models of speech production have been proposed in the literature (Guenther, 1992; Markey, 1994; Guenther, 1995; Bailly, 1997; Boersma, 1998; Guenther and Perkell, 2003). Each of these models is limited by the lack of a quantitative and physiologically plausible articulatory model. Guenther's DIVA model of speech production, for instance, uses stylized articulatory parameters as "Tongue body horizontal position with respect to skull" and "Tongue body height with respect to jaw." Such a simplified articulatory representation does not invalidate the results of the model, but there is clearly room for increased phonetic plausibility, which would enhance the scope of the model and permit many more quantitative tests to be performed.

1.2.3 Phonological theory

A biomechanical tongue model has applications to phonological theory as well. With the notable exception of Jakobson et al. (1951), all theories of phonological representation have made reference to articulatory parameters. Chomsky and Halle (1968) use articulatory features like [anterior] and [coronal] to describe segments. In fact, Halle (1983) goes so far as to predict muscle activation patterns that correspond to distinctive feature values for vowels.[5] More recent feature-geometric approaches to representation (including Clements (1985), and to an even greater extent Halle (1992)), have used articulatory, and particularly lingual, descriptors (Place, Coronal,

[5]His predictions are tested in Section 6.2.3.

Dorsal, Root) as non-terminal nodes of the feature tree. Browman and Goldstein (1989) have proposed that formal phonological representations should be gestural scores, representing the activity of different articulators through a word.[6]

The foregoing demonstrates that phonological theory recognizes the importance for understanding phonological patterns of articulation generally, and of the tongue particularly. This recognition has heretofore not been coupled with a systematic understanding of the articulatory mechanism, however. If the mechanical properties of the tongue were better understood, a better understanding of the relationship between the vocal tract and phonological systems would be gained.

1.3 Availability of the model

Both the anatomical data and the model input files have been made supplemental material for this dissertation for use by other researchers in subsequent research. This is intended to address two issues that become apparent in surveying previous biomechanical tongue modeling research (further discussed in Section 2.2.3).

First, tongue modeling research has not been a cumulative enterprise. Aside from cases in which a researcher improves upon his own model, most researchers have started their models from scratch—as was indeed the case in the present work. Consequently, advances in some parts of a model are usually coupled with losses in other parts. By making the entire model available, it is hoped that it can be steadily improved, so that it more closely represents the human tongue. This dissertation presents *a* model of the tongue, but certainly not the final model of the tongue. Many of the assumptions made are either empirically suspect or outright counterfactual. Further research is needed to refine the model.

Second, in previous research, the behavior of tongue models has been studied exclusively by the researchers who developed the tongue models. Ideally, a tongue

[6] As with the control models, Browman and Goldstein are limited in their gestural parameters to stylized representations of motor control commands. More detailed representations of motor commands, the effects of which can be determined with a biomechanical model, will enable more detailed predictions, which would allow better tests of the theory.

model should be available to any researcher who has use for it. All biomechanical tongue models to date have been produced by phoneticians, but a biomechanical tongue model is of equal or greater relevance to speech pathologists, respiratory physiologists, or medical doctors (see footnote 3 in this chapter). By making this model available, it is hoped that biomechanical tongue modeling will be better taken advantage of by non-phoneticians.

1.4 Outline of the dissertation

The primary focus of the dissertation is the creation of a biomechanical tongue model. As background to that, Chapter 2 contains a review of pertinent anatomical and physiological studies, and of previous biomechanical tongue models. Chapter 3 presents a detailed anatomical study of the female dataset from the Visible Human Project. This study yielded a volumetric representation of the tongue at the full precision of that data set ($0.33 \times 0.33 \times 0.33$ mm).[7]

The anatomical data are incorporated into a finite element model of the tongue in Chapter 4. The present work advances the state of biomechanical tongue modeling in a couple of key respects. First, the geometry of tongue model is based on an actual anatomical sample (Chapter 3). In addition, the orientation of muscle fibers, which varies throughout each tongue muscle, was derived from this data set. Finally, a mesh validation study was performed on the model, which ensures that the model's results are not greatly marred by the particular mesh chosen. The mesh validation study indicated that a much finer mesh is required for a valid model than has previously been included in any tongue model.

The secondary focus of the dissertation is to use the model to answer various questions about the tongue. Chapter 5 contains an analysis of the range of motion of the tongue. Samples of tongue postures are submitted to principal component

[7] Zemlin (1988) is an indispensable source of information about the anatomy and physiology of the speech mechanism. Basic information on neuromuscular physiology is available from a number of sources. Chapter 18 of Nolte (2002) is a particularly clear introduction.

analysis, resulting in low-dimensional representation of tongue postures.[8] Statistical models are fit in order to determine an efficient mapping between muscle activation level and tongue posture, and vice versa. It happens that fairly simple statistical models can be used to generate accurate mappings in both directions. Finally, an attempt is made to find a low-dimensional representation of antisymmetrical tongue postures.[9] This is made difficult by the sample size required for antisymmetrical tongue shapes, and a novel technique is used to make the problem tractable.

Chapter 6 focuses on more specific results. The consequence of contracting each tongue muscle is considered. There is a study of the range of tongue motion using just single-muscle contractions, as opposed to multi-muscle contractions, which demonstrates that muscle synergies are crucial to producing the full range of tongue movement. Tests of specific hypotheses advanced by previous researchers are also included. In general, previous researchers have not made accurate predictions about which muscle activations lead to which tongue postures. Notably however, the predictions of the muscular hydrostat theory of tongue movement (Kier and Smith, 1985, et seq.) are largely confirmed. Finally, there is a model of flaccid dysarthria, in which half of the tongue receives no motor input. The mechanical consequences of this disorder are considered. It is determined that different muscle activation patterns are required in a hemiplegic tongue than in a normal tongue, in order to produce (near) equivalent postures.

Chapter 7 concludes the dissertation. The major findings are reviewed, and suggestions for future development of the model are made. The appendices contain information that is sufficiently detailed that the model can be reproduced and modified. Appendix B contains a guide to the labeled volumetric anatomical data. Appendix A contains a careful description of the particular implementation of the finite element model. Appendix C has a description of Tongue Viewer, a piece of software included as supplemental material in the dissertation, which has been developed to

[8] Jolliffe (2002) provides a rigorous introduction to principal component analysis. There are also numerous more accessible tutorials available online.

[9] Antisymmetrical postures are those that are either symmetrical or asymmetrical.

visualize model results.

The organization of the dissertation is fairly modular, with the exception that some of the results in Chapter 6 depend on techniques described in Chapter 5. Each chapter consists of a general introduction, followed by the technical details of the implementation of the pertinent aspect of the model. Depending on the reader's particular interest, implementation details can be skimmed or skipped. Since the validity of the model depends on the details of its implementation, however, a great deal of detail has been included.

CHAPTER 2

Relation to Previous Work

> O God my God, what miseries and mockeries did I now experience, when obedience to my teachers was proposed to me, as proper in a boy, in order that in this world I might prosper, and excel in tongue-science, which should serve to the "praise of men," and to deceitful riches.
>
> —Augustine, *Confessions IX*

The purpose of this chapter is to review previous tongue research that is pertinent to the creation of a tongue model. The first half reviews studies of the tongue in itself, while the second reviews previous biomechanical modeling efforts.

2.1 Previous Human Tongue Research

The discussion of previous research is divided into two sections: anatomical and physiological studies, and theoretical approaches to tongue movement. The review of previous anatomical studies describes previous anatomical studies, but does not provide a description of the gross anatomy of the tongue. This is deferred until Section 3.1, where previous research is presented within the context of the present anatomical study; readers unfamiliar with tongue anatomy may wish to read that section before continuing in this chapter. The physiological review focuses on aspects of tongue physiology that are pertinent to speech motor planning. There are a number of neurophysiological facts that suggest that the central nervous system has very fine motor control of the tongue. In fact the present model has only a very unsophisticated control mechanism. The review of the neurophysiological research is intended to call attention to this fact, to indicate directions for future research.

Where possible, the anatomical and physiological studies of human tongue have been presented. Some types of studies, however, cannot be performed on humans; in these cases animal studies are required. The animal kingdom features a diversity of tongue types. In frogs alone, for instance, three mechanically distinct tongue protrusion mechanisms have been described (Nishikawa, 1999). Within the narrower mammalian context, Doran and Baggett (1971) have categorized mammalian tongues as Type I and Type II. Type I tongues are those which remain in the mouth (i.e., do not elongate more than 50%), while Type II tongues can operate outside of the mouth, typically for prey capture. The types are distinguished not only functionally but also anatomically. Type II tongues have a sternoglossus muscle (a muscle originating from the sternum and inserting into the tongue), while Type I tongues do not. Canines, felines, and primates all have Type I tongues, while Type II tongues are found in monotremata (echidnas, platypodes, etc.), some marsupials, and the pangolin. For comparison to human tongues, it is mammals with Type I tongues that are of interest. The rat is frequently used as an experimental animal in tongue studies.

2.1.1 Anatomical & physiological studies

Descriptions of the gross anatomy of the tongue are to be found in general anatomical works (e.g., Salter, 1852; Gray, 1918), but the seminal works in tongue anatomy were published by Shafik Abd-El-Malek. Abd-El-Malek (1938) presents a detailed study of tongue anatomy of the mammalian tongue generally, while Abd-El-Malek (1939) focuses specifically on the human tongue. Abd-El-Malek (1939) provided detailed information concerning the orientation of muscle fibers in the tongue, the tongue's connective tissue framework, and described the course of the hypoglossal nerve.

There are in addition two anatomical studies of the tongue geared specifically for speech researchers, Miyawaki (1974) and Takemoto (2001). Miyawaki studied the musculature of the tongue by investigating the orientations of fibers in the axial, coronal, and sagittal planes of section. His article includes diagrams of tongue fibers

in many of these sections. Takemoto produced a more careful study of the tongue, using eight specimens. He examined oblique sections of the tongue, which paralleled the fibers of genioglossus. Using this approach, he documented the histological strata of the tongue. An additional notable aspect of his work was that he created a three-dimensional geometric model, based on his impressions of the tongues he examined; this model is shown in Figure 3.7.

Iskander and Sanders (2003) compared adult and neonatal human tongues, noting several differences in the structure of the tongue at these two stages of development, including differences in the relative sizes of the tongue mucosa, and of the intrinsic and extrinsic muscles. Recently Saito and Itoh (2007) have studied human styloglossus using chemical maceration, a technique in which the connective tissue of a muscle is dissolved away, so that the muscle fibers can be studied in detail. This type of dissection, which has so far been applied only to the one muscle in human tongue, will undoubtedly provide substantially greater detail about the arrangement of tongue fibers than is currently available.

Several researchers have collected *in vivo* anatomical data—including fiber orientation data—using diffusion tensor imaging, a magnetic resonance imaging technique. The technique is attractive because it enables data collection from an *in vivo* tongue, whereas dissection studies are limited to cadaver tongues. At this stage in the development, the technique does not seem quite applicable to finite element modeling, because of low spatial resolution, and the time required to obtain a slice. Gilbert and Napadow (2005) imaged three human tongues. Each of their sagittal slices had a resolution of 3 mm $\times$ 3 mm $\times$ 6 mm, and it took 3 minutes to record each slice, the entire tongue being represented with 7 slices. Shinagawa et al. (2008) acquired images from single sections of *in vivo* human tongues during rest and tongue protrusion. Their data had 3.1 mm $\times$ 3.1 mm $\times$ 3.0 mm voxels, and required about three minutes for acquisition.

Two other studies have been performed on bovine tongues. Gilbert et al. (2006) obtained a spatial resolution of 2.8 mm $\times$ 3 mm $\times$ 4 mm. Kim et al. (forthcoming) obtained resolution of 2.03 mm $\times$ 2.03 mm $\times$ 2 mm resolution. The scans in the

Kim et al. study took seven and a half hours, which indicates the limited immediate potential for high resolution *in vivo* human imaging.

2.1.1.1 Neural control of the tongue

Neurophysiology is a mature field, and a great deal is understood about the organization of motor control, particularly at the level of the last-order outputs. The fundamental force-generating entity in a muscle is the motor unit. A motor unit consists of a single motor neuron (or motoneuron), and an assortment of muscle fibers onto which the neuron synapses (connects). When the motoneuron fires, an electrical potential travels down its synapse, which, through complicated chemical reactions, causes all of its associated muscle fibers to contract simultaneously. Motor units are typically not directly controlled by the central nervous system. Rather, they are organized into motor pools, which the central nervous system actually controls. Within a motor pool, motor units are activated in a systematic order—from weakest to strongest—as greater force is required (Henneman, 1979).[1]

The motoneurons that supply tongue muscles reside in the hypoglossal nucleus, a cluster of cells in the medulla, which is part of the brain stem. The hypoglossal nucleus is abbreviated CXII because it is the twelfth cranial nerve. The axons of the motoneurons that reside in the hypoglossal nucleus form the hypoglossal nerve.[2]

The number of motor units in the tongue is fantastic: anatomical studies of healthy adults place the number of neurons in the hypoglossal nucleus between 7,093 and 8,817 (Wožniak and Young, 1969; Atsumi and Miyatake, 1987; O'Kusky and Norman, 1995, cumulative $n = 14$).[3] This is an extremely large number relative to other muscles. Biceps brachii, for instance, is innervated by an average of 441.5

[1] This is the case at least for isometric contractions. For lengthening and shortening contractions, the situation appears to be more complex; see Enoka and Fuglevand (2001) for review.

[2] The nucleus is called hypoglossal not because it is beneath the tongue, but because it associated with the hypoglossal nerve, which does run beneath the tongue.

[3] Note that there are two hypoglossal nuclei in a healthy body, for the right and left sides of the tongue, bringing estimates for the total number of hypoglossal motoneurons to between 14,186 and 17,634.

motoneurons (Hamilton et al., 2004). In the hypoglossal nuclei of lower mammals, motoneurons are grouped together by muscle, a musculotopic organization. The method of investigating this organization is to inject a retrograde tracing dye into a particular tongue muscle, and then after a period to kill the animal, dissect the hypoglossal nucleus, and see where the dye went. For ethical reasons, this cannot be done to humans, so extrapolation is required from lower animals. This organization is has been shown for a variety of animals, however, so the extrapolation seems to be warranted (Rat: Guo et al. 1996; Rabbit and Cat: Uemura-Sumi et al. 1988; Macaca fuscata, Macaca nemestrina, Macaca speciosa, Macaca irus: Uemura-Sumi et al. 1981; Macaca fascicularis: Sokoloff and Deacon 1992). The musculotopic organization of the hypoglossal nucleus suggests, but does not prove, that the muscle level of organization has some role in motor planning.

2.1.1.2 Compartments

The muscle compartment is an intermediate level of organization between the motor unit and the whole muscle. The boundaries of muscles are defined by gross anatomical dissection, but more careful anatomical study has revealed that many muscles have distinct subvolumes, which are called muscle compartments. These compartments are innervated by different sets of axons (with some minor overlap), and can have distinct muscle fiber type compositions (English and Letbetter 1982a,b; see English et al. 1993 for a more accessible review). The fact that muscle compartments receive distinct subsets of axons suggests that each muscle compartment is supplied by a distinct motor pool. Muscles of the human hand can exhibit compartmental structure, and each compartment seems to be innervated by fairly distinct populations of motoneurons (Keen and Fuglevand, 2004a,b). Caution is required in generalizing across muscles: this does not prove that in the human tongue separate compartments are supplied by disjoint sets of motoneurons, but it is suggestive of that possibility.

Compartmental organization of the canine tongue has been demonstrated by

Mu and Sanders (1999, 2000).[4] They performed microdissection of several canine tongues using a stain that renders muscle transparent. If the nerve fibers are stained black, then it becomes possible to observe the course of the nerve fibers quite easily, which is helpful in identifying muscle compartments. The authors report two compartments for genioglossus. The horizontal component consists of fibers that insert onto the hyoid bone, with the remainder to the "fan" belonging to the oblique compartment. They identified about 48 first-order branches of the hypoglossal nerve, which may be diagnostic of muscle compartments. The oblique compartment of genioglossus received seven to eight first-order branches, and the horizontal three. Superior longitudinalis receives 3–4 branches, and transversus and verticalis receive 30–31 together. Inferior longitudinalis, styloglossus, and hyoglossus each received one first-order branch.[5]

Miyawaki et al. (1975) found distinct levels of electromyographic activity in five different regions of genioglossus. In tongue modeling research this is commonly taken to suggest a particular compartmental organization of genioglossus in humans, although it is not clear proof of this since EMG signals recorded from different parts of a muscle vary anyway. Unfortunately no careful anatomical data about the compartmental organization of the human tongue are not yet available.[6] The work is underway, however (Ira Sanders, p.c.; Dr. Sanders has indicated to the author that the compartmental organization of the human tongue is similar to, but more complex than, that of the dog tongue).

2.1.1.3 Muscle Spindles

Muscle spindles are specialized muscle fibers that provide proprioceptive information about the muscle in which they are embedded—namely the length of the muscle and the rate of change of the length of the muscle. There are muscle spindles in both

[4] See Section 3.1 for a description of the homologous muscles in man.

[5] One branch innervated the geniohyoid, which is not innervated by the hypoglossal nucleus, although the axons that innervate it do run in the distal portions of the hypoglossal nerve.

[6] A study by Touré et al. (2005) described the course of nerves of the tongue using Sihler's stain, but did not describe compartmental organization.

intrinsic (Cooper, 1953) and extrinsic (Forster, 1894, cited in Cooper 1953) muscles of the human tongue. Cooper reports that the tip of the tongue has very few muscle spindles, while the region proximal to the tip is richly supplied.[7] Cooper also noted an abundance of spindles in superior longitudinalis and transversus.

For a time it was thought that muscle spindles were found only in the tongues of higher mammals (see Lowe 1981 for review). This led some researchers to speculate that the presence of the muscle spindle in the tongue might be an important evolutionary development, which provided primates with unique abilities in controlling their tongues, and which therefore might be an important evolutionary landmark in the development of speech (Bowman, 1971). Further research uncovered muscle spindles in the tongue of a rat, however (Smith, 1989), which seems to have put an end to this line of speculation.

2.1.1.4 Corticobulbar connections

As discussed above, it is generally the case that the motor pool is the output of planning by the central nervous system. One exception to this general pattern is that, in some cases, there are projections from motor cortex directly on to motoneurons. Such connections have been demonstrated for motoneurons innervating the muscles of the hand, and it has been proposed these projections form part of the neural basis for primate manual dexterity (Lemon and Griffiths, 2005). There are also projections from motor cortex onto neurons in the hypoglossal nucleus in man (Kuypers, 1958). This is suggestive of some kind of unusually fine level of motor control for the tongue.[8]

Jürgens and Alipour (2002) consider corticobulbar projections from an evolutionary perspective. Noting that the projections have not been found in studies of the

[7] Cooper defines the tip of the tongue as the distal third (1953, 194). This definition is not entirely clear, however: a third as measured along the anteroposterior length of the tongue? A third of the length of the tongue measured along the curved surface from tip to root?

[8] These connections are called corticobulbar, *bulb* being a less frequently used term for the brainstem.

cat or rat, they looked for corticobulbar projections in the rhesus monkey (Macaca mulatta, an Old World primate, and therefore more closely related to man), squirrel monkeys and saddle-back tamarins (both New World primates), and the non-primate tree-shrew. Corticobulbar connections were found in the rhesus and squirrel monkeys, but not in the saddle-back tamarin or the tree shrew. The findings support the idea that corticobulbar connections are unique to primates, but do not permit a clean Old World/New World split to be defined.

Subsequent research will be required to determine whether corticobulbar connections are indeed unique to primates. Cautious generalization seems in order, given the muscle spindle results.

2.1.1.5 Tongue motor units

Fine anatomical and physiological studies of the tongue have revealed interesting facts about the tongue's motor units. Motor units in different parts of the tongue are heterogeneous with respect to a number of different properties.

Sokoloff (2000), studying the rat, identified regional (anteroposterior) variation in the magnitude of tetanic force generated by superior longitudinalis motor units, where posterior motor units generated stronger force. Localization of motor unit contractile properties seems to be a feature of the human tongue as well. Slaughter et al. (2005) documented an in-series design for human superior longitudinalis. Muscle fibers associated with motor units were between 2.8 mm and 15.7 mm, or between 3% and 17% of the length of the muscle; the average muscle fiber was about two thirds of the length of the tongue. Based on these findings, the authors speculate that the nervous system could flex different regions of the tongue, providing added flexibility in tongue deformation. Further evidence for regional specifity of the human tongue comes from Stål et al. (2003). Using immunohistochemistry they documented the type of muscle fibers in four tongue regions. They found that the tongue is primarily (60%) composed on type II, or fast contracting, fibers. Further, the anterior regions of the tongue have a relatively higher concentration of fast-twitch fibers. They comment, "the tongue apex with its high amount of type IIA [fast fatigue-

resistant] and IIAB [fast fatigable] fibres seems to be adapted for rapid positioning of the tongue tip, which probably is of special importance for speech." There were also muscle-specific patterns. Verticalis had a high percentage of fast twitch fibers compared to transversus or the longitudinal muscles. Variation in fiber type composition has also been documented in canine tongue by Mu and Sanders (2000). They found that the horizontal compartment of genioglossus contained a greater proportion of slow-twitch fibers (64%), while the oblique compartment contained more fast twitch (59%). However, posterior portions of the oblique compartment a higher concentration of fast-twitch fibers than the anterior portion.

A fascinating property of human genioglossus motoneurons has been reported by Bailey et al. (2007).[9] They found motoneurons that fired at different rates depending on the position of the tongue, quantified in their study as the extent of horizontal tongue protrusion. The firing rate of genioglossus motoneurons would plateau, and at times decrease, when the tongue was protruded beyond a certain position, which was specific to each motoneuron. Another interesting finding of the study was that genioglossus motoneurons have a faster than average firing rate when the tongue was holding its position, ∼16 Hz, whereas 6–10 Hz is characteristic of other muscles for a similar task. Finally, the authors also recorded motoneurons that had brief but very high frequency (∼80 Hz) firing spurts, which seemed to be associated with fast tongue movements.

Genioglossus motoneurons can also be differentiated on functional criteria. There are motoneurons that are active during inspiration, and others which are tonically active, but fire faster during inspiration (inspiratory/expiratory motoneurons). Both motoneurons fire faster when the head is bent backward, which suggests that they are responsible for decreasing airflow resistance in the upper airway (Tsuiki et al., 2000).

[9] See Section 3.1 for a description of genioglossus.

2.1.1.6 Intermediate summary

Anatomical and physiological studies of the tongue, in man and in lower animals, raise many questions about tongue motor control. Much work remains in delineating the primary strategies that underlie tongue movement, but it seems reasonably clear that something smaller than the whole muscle is being controlled. There is evidence for compartmental organization of the tongue (Mu and Sanders, 1999, 2000). There are also regionally specific properties of tongue motor units (Sokoloff, 2000; Stål et al., 2003; Slaughter et al., 2005), task-specific motoneurons (Tsuiki et al., 2000), motoneurons that are preferentially active at different tongue positions (Bailey et al., 2007), and motoneurons that may be specifically associated with quick movements (Bailey et al., 2007). This variety of specializations may be features of the compartmental organization of the tongue, or of some other way that tongue motor units are organized.

Unfortunately this level of detail has not been included in the present model (or any other model). Instead, control of the model is generally limited to whole muscle contraction. The two reasons for this are lack of availability of specific data, and the computational expense associated with the more complex control mechanism. Human studies are not presently available in such detail as would permit their inclusion in a quantitative model—at least in a way that would not be a caricature of the actual anatomical and physiological detail; it would be possible, but ultimately uninformative, to impose an arbitrary compartmental scheme on tongue muscles. Secondly, even if this level of detail could be included, it would pose problems for evaluating the model. Increasing the number of control parameters in a model makes it more difficult to sample the behavior of the model, a problem that is dealt with at length in Chapter 5. Nevertheless, increased computational power, and further research into the tongue, will eventually solve both of these problems. Increasing the realism of the neurophysiological control mechanism should be a priority in subsequent modeling efforts.

2.1.2 Descriptions of and theories of tongue movement

Several researchers have undertaken to describe the movement of the tongue in various behavioral contexts, and all have additionally been keen to predict the muscular contractions associated with these movements. The efforts reviewed here are fairly broad in scope, ranging from the descriptive (Abd-El-Malek, 1955; Hardcastle, 1976) to the theoretical (Kier and Smith, 1985). Descriptions of tongue movement can be quite difficult to follow and are frequently imprecise, a consequence of the fact that the vocabulary for describing parts of the tongue, and tongue deformation, is quite limited. It is evident that all of the researchers have a good deal of tacit knowledge about tongue movement, but it is equally evident that not all of this knowledge could be committed to prose.

2.1.2.1 Abd-El-Malek (1955)

Abd-El-Malek (1955) described the tongue postures and movements associated with mastication. He observed his subjects masticating nuts, gelatin, and chewing gum, while the cheeks were retracted slightly with forceps. His observations led him to the following description of the stages of mastication.

1. The preparatory stage. The tongue forms a pouch to accept the bolus. (Figures 2.1 and 2.2)

2. The throwing stage. The tongue twists to deposit the bolus onto the molars. (Figure 2.3)

3. The guarding stage. The tongue twists even more, making contact with the upper and lower teeth, in order to keep the bolus between the molars during mastication. (On the lateral side of the bolus, the buccinator holds the food in place.) (Figure 2.4)

4. The sorting-out stage. The more-masticated elements of the bolus are sorted to the lateral part of the tongue, while the less-masticated elements are sorted to the medial side. Abd-El-Malek describes these as "rapid and jerky."

5. Bolus formation. The tongue moves side to side, mixing the bolus with saliva and coating it with mucus. Deglutition begins with a trough-like tongue shape. (Figure 2.5)

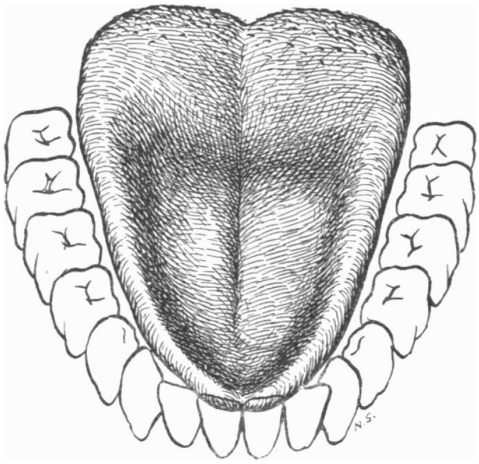

Figure 2.1: Abd-El-Malek (1955) illustration of the preparatory stage of mastication.

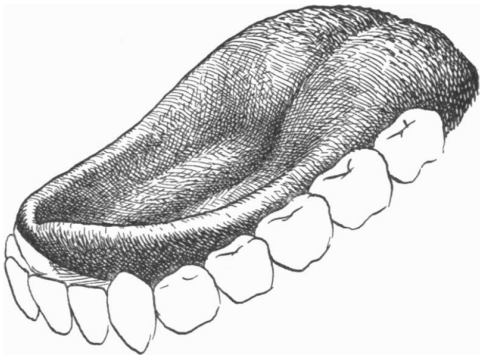

Figure 2.2: Abd-El-Malek (1955) illustration of the preparatory stage of mastication (alternate view of Figure 2.1).

In addition to describing these stages, Abd-El-Malek makes inferences about which muscles contract to produce each posture. A lowered medial surface with raised lateral surface (Figures 2.1 and 2.2[10]) is associated with bilateral contraction of genioglossus, hyoglossus, and styloglossus. Trough formation is associated

[10]Somewhat confusingly, Abd-El-Malek also includes photographs of these tongue postures, which are given the same labels, but are quite different from the illustrations. The difference

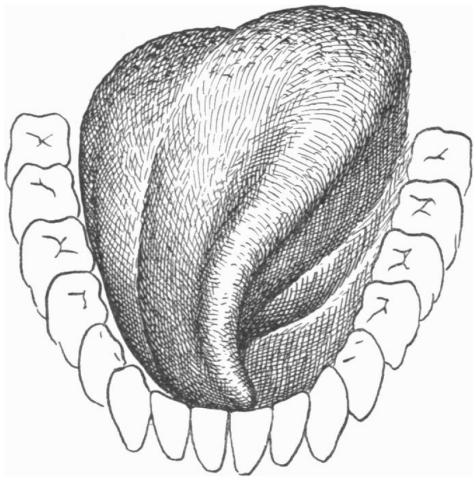

Figure 2.3: Abd-El-Malek (1955) illustration of the throwing stage of mastication.

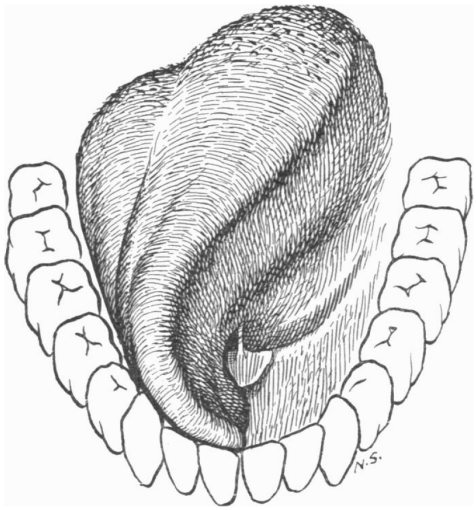

Figure 2.4: Abd-El-Malek (1955) illustration of the guarding stage of mastication.

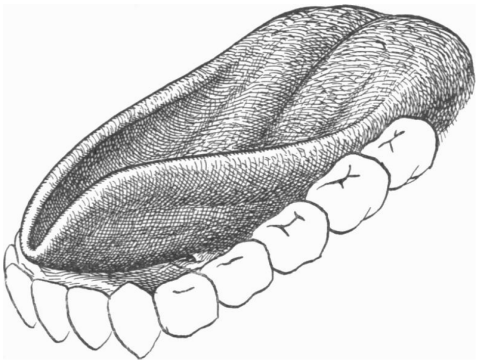

Figure 2.5: Abd-El-Malek (1955) illustration of the initial stage of deglutition.

additionally with contraction of the superior longitudinal muscles. Elongation of the tongue in this position is accomplished by contraction of transversus. He associates the throwing and guarding stages with contralateral action of styloglossus. Depression of the tongue tip comes from contracting inferior longitudinalis. These predictions are tested in Section 6.2.1.

2.1.2.2 Logeman 1983

Physiological aspects of deglutittion are reviewed in detail by Logemann (1983). She divides the degluttition into four phases: the oral preparatory phase, oral phase, pharyngeal phase, and esophageal phase. X-ray tracings of these four stages are presented in Figure 2.6. The tongue posture from the oral prepatory stage is given in Figure 2.6a. To place this posture in relation to those of Abd-El-Malek (1955), Abd-El-Malek's initial stage of degluttion corresponds to the beginning of Logeman's oral preparatory stage; the posture in Figure 2.5 occurs shortly before that in Figure 2.6a. The oral preparatory phase involves mastication and bolus formation. During mastication food is continally pushed onto the molars to be broken up

seems attributable to the fact that in the photographs the subject was making an effort to produce the postures outside of the mouth, i.e., in order to be photographed. The illustrations have been considered to be more representative of the intended postures.

(cf. Abd-El-Malek's description). The bolus is formed between the anterior of the tongue and the palate, often with the sides of the tongue pulled up to hold the food (cf. Figure 2.5). In the oral stage, the bolus is squeezed (by the tongue) toward the pharynx with peristaltic motion. Interestingly, Logeman indicates that is unknown whether the tongue moves to the palate then pulls the bolus back, or if the tongue is simply elevated sequentially toward the palate, squeezing the bolus to the pharynx. The oral prepatory phase ends when the bolus moves past the anterior faucial arch, which initiates the involuntary phase of swallowing (Figure 2.6b). As the bolus enters the pharynx (Figure 2.6c), Logeman reports that transport of the bolus is accomplished not by the tongue, but by peristaltic contraction of the pharyngeal constrictors. The bolus is moved into and through the esophagus (Figure 2.6d and Figure 2.6e) by peristaltic contraction of the esophageal muscles. The return of the tongue to rest posture following the oral phases (i.e., the change from Figure 2.6b to Figure 2.6c) presumably occurs without special muscular activity.

2.1.2.3 Hardcastle 1976

Hardcastle (1976, 100) offers a simple theory of tongue movement, in which the various movements of the tongue are captured with seven articulatory parameters. It is evident that the parameters are influenced by the articulatory phonetic taxonomy of tongue postures.

1. Horizontal forward-backwards movement of the tongue body
2. Vertical upwards-downwards movement of the tongue body
3. Horizontal forward-backwards movement of the tip-blade
4. Vertical upwards-downwards movement of the tip-blade
5. Transverse cross-sectional configuration of the tongue body: convex-concave, in relation to the palate

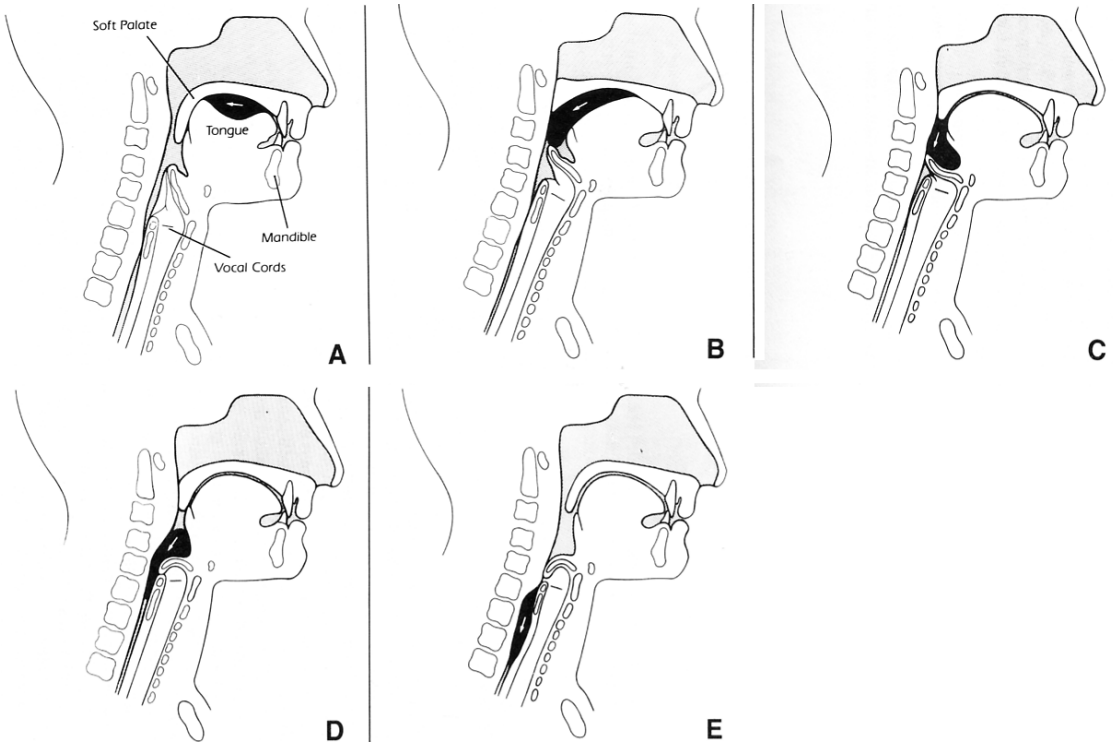

Figure 2.6: X-ray tracings of Logeman's stages of swallowing. Adapted from Logemann (1983). Voluntary portion of swallow (A); initiation of the swallowing reflex (B); passage of the bolus through the pharynx (C); initiation of passage of the bolus into the cervical esophagus (D); the entire bolus in the cervical esophagus (E).

6. Transverse cross-sectional configuration extending throughout the whole length of the tongue, particularly the tip and blade—degree of central grooving.

7. Surface plan of the tongue dorsum—spread, tapered.

Hardcastle associates these parameters with particular tongue muscles, again, drawing on his own observations of articulation, though these judgments are fairly informal.[11]

Descriptively, Hardcastle's approach is lacking in a few respects. There is no mention of tongue root movement, or indeed of any finer distinction than the tip/dorsum system; this can be viewed as a consequence of our impoverished tongue vocabulary. Second, the model assumes that the tongue is symmetrical about the sagittal plane, which is acceptable for speech production (Hardcastle's focus), but limits the general applicability of the theory.

2.1.2.4 The intrinsic/extrinsic split

One idea that has been put forward to understand the control of the tongue is that the intrinsic and extrinsic muscle of the tongue have distinct functional roles, in some sense. The idea is that a simple anatomical feature (presence or absence of bony attachment) is reflective of higher level physiological organization, which, if true, would be quite elegant. Perkell (1969) invokes this idea is explaining speech movements:

> The general differences in velocity, complexity, precision of movement, and in anatomy suggest that different types of muscles are generally responsible for consonant and vowel production. It is probable that the

[11] The following passage from the description of the muscular control of tongue grooving offers an idea of Hardcastle's confidence in his judgments: "The muscles probably most directly responsible for the central groove are the transversus (median part) and the verticalis muscle (particularly the superficial fibres). The tongue may be spread out slightly be the whole verticalis. At the same time the styloglossus and palatoglossus (from a fixed velum) may act synergistically with the transversus in keeping the sides raised." (Hardcastle, 1976, 105).

> articulation of vowels is accomplished principally by the larger, slower extrinsic musculature which controls tongue position. On the other hand, consonant articulation requires the addition of the precise, more complex, and faster function of the smaller, intrinsic musculature....Thus the speech production mechanism can be thought of as being composed of two neuromuscular systems with different behavior characteristics... (Perkell, 1969, 61,62)

Perkell predicts specifically that the extrinsic muscles are primarily responsible for positioning the tongue for vowel production, and that the intrinsic muscles are primarily responsible for positioning the tongue for consonant production.

Bailey and Fregosi (2004) note that an intrinsic/extrinsic split has become widely accepted by a number of respiratory physiologists, in spite of indications that the intrinsic/extrinsic distinction may not be functionally significant. Bailey and Fregosi found coordinated activity between the intrinsic superior longitudinalis and extrinsic hyoglossus, during respiration in the rat. Commenting on the work, Sokoloff (2004) calls into question the assumption that a function distinction exists between the intrinsic and extrinsic tongue muscles. Subsequently, Bailey et al. (2006) demonstrated that the intrinsic muscles of the tongue contribute to the volume of the velopharyngeal airway in the rat, a function known also to be associated with (extrinsic) genioglossus.

The idea of a functional division between the intrinsic and extrinsic muscles is more of a heuristic than a hypothesis. No researcher (to the author's knowledge) has proposed an experimental test that would falsify the idea. Nevertheless, if the intrinsic/extrinsic division is reflective of a substantial organizing principle of tongue movement, one might expect that the effect of contracting one set of muscles would have distinct mechanical consequences from the other. In Section 6.2.4 results are presented that indicate that this is not the case.

2.1.2.5 Hydrostatic theory

A widely accepted theory of tongue movement is the muscular hydrostat theory (Kier and Smith, 1985). This theory attempts to explain the movement of anatomical structures in many species (like tongues, trunks, and tentacles) which are able to deform themselves, while lacking a firm skeletal structure.[12] For Kier and Smith, the key observation is that muscle is composed primarily of water, and is therefore incompressible at physiological pressures. In a volume-preserving structure, the way to extend along one dimension is to contract in perpendicular directions.[13]

Biomechanical models have been proposed for a number of non-mammalian tongues, including the lizard (Chiel et al., 1992) and certain frog species (Nishikawa et al., 1999; Meyers et al., 2004), which supports the hydrostatic theory. In the case of Meyers et al. (2004), the hydrostatic model was supported by denervation studies as well.[14] Muscular hydrostat theory has not been tested thoroughly on mammalian tongues. One prediction which has been tested is that multiple muscle groups should be active simultaneously in tongue behavior. This prediction is born out in the findings of Bailey and Fregosi (2004), who found coordinated activity between the intrinsic and extrinsic tongue muscles in the rat.

Predictions made by the muscular hydrostat theory—both those of Kier and Smith (1985), and Napadow et al. (2002), who use the hydrostat theory—were tested in the present work, with the results largely corroborating the theory (Section 6.2.2).

[12]The skeletal structure could either be bony skeleton, or a rigid hydrostatic capsule.

[13]By way of illustration, imagine holding a water balloon, and wishing to expand it along the long axis. It would be possible to grasp each end of the long axis and pull. It would also be possible to extend the balloon by squeezing the balloon around the perpendicular plane. Narrowing the cross-sectional area of the balloon necessitates lengthening of the balloon.

[14]Hydrostatic theory does not account for all species, however: Nishikawa (1999) places hydrostatic elongation along side two other tongue protrusion strategies that are found in frogs, which include mechanical pulling (using genioglossus) and inertial elongation (flinging). Impressionstically, all three of these protrusion strategies seem available to humans. Tongue protrusion is possible with both a flat tongue (i.e., one without a decreased cross-sectional area) and a tube-shaped tongue. Flinging tongue protrusions are also seem possible, although studies would be required to ensure that the tongue protruded because of inertia rather than muscular contraction.

2.2 Previous Modeling Studies

Researchers have long recognized the importance of understanding the contribution of the biomechanical properties of the tongue to speech research. In consequence, there have been occasional efforts, particularly within the last fifteen years, to develop a quantitative biomechanical tongue model. Here the development of tongue models is reviewed. Two-dimensional models are considered first, and then three-dimensional models. In Section 2.2.3 general comments are made about the limitations of previous studies, and ways in which the present work addresses them.

2.2.1 Two-dimensional and "two-and-a-half"-dimensional models

Perkell (1974) presents a two-dimensional model of the tongue. The model consists of sixteen "fleshpoints," connected by active and passive tension generating elements (Figure 2.7). His model has ten control inputs: one apiece for superior longitudinalis, inferior longitudinalis, verticalis, hyoglossus, and styloglossus; three for the anterior, middle, and posterior portions of genioglossus, and one each for the pharyngeal constrictors and the mylohyoid. The hyoid and mandible also move along a single dimension.

Muscle contraction is modeled with a phenomenological spring-and-dashpot model, such as is still in use in modeling musculoskeletal biomechanics (Yamaguchi, 2001). That is, passive stress in the model was linear elastic. He also enforced incompressibility of tongue tissue by using a numerical method.

Payan and Perrier (1997) offer a finite-element model of the tongue that is very similar to Perkell's. The model, shown in Figure 2.8, consists of 48 elements. They include the same lingual muscles as Perkell (1974), but did not allow for hyoid movement. Muscle is modeled as a linear elastic material, which becomes stiffer upon contraction. They explored control of the model with the Equilibrium Point Hypothesis (Feldman, 1986), generating acoustic signals from the model that they matched to those of a speaker with a bite block, with good accuracy.

Sanguineti et al. (1997, 1998) created a model similar to that of Perkell (1974).

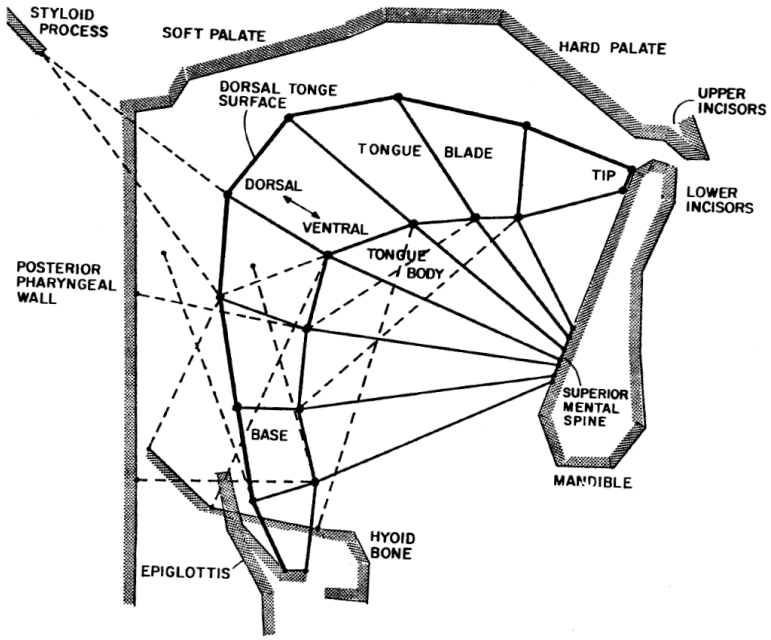

Figure 2.7: Two-dimensional tongue model from Perkell (1974).

They base the model geometry on X-ray images. They assumed a linear material response, and specified muscle geometry similarly to Perkell (1974).

Dang and Honda (1998, 2001, 2002, where latter publications focus on control and acoustic output from the model) present a model that is "two-and-a-half"-dimensional. That is, their model has thickness (2 cm) in the sagittal plane, but does not model the whole tongue. The 120-element mesh is shown in Figure 2.10. These authors collected three-dimensional anatomical data to determine the geometry of their model. The anatomical data consisted of 15 sagittal slices of MR images, each 0.7 cm thick. Each image had 30 cm $\times$ 30 cm field of view, represented with 256 $\times$ 256 pixels (0.117 cm/pixel edge). A higher resolution set of images was used to identify individual tongue muscles (25 cm $\times$ 25 cm, 0.35 cm slice thickness, with 0.05 cm overlap between adjacent sagittal slices). Of interest in interpreting the geometry of the model is the face that the anatomical images were of the vowel [e] (spoken by a Japanese man), instead of the rest position.

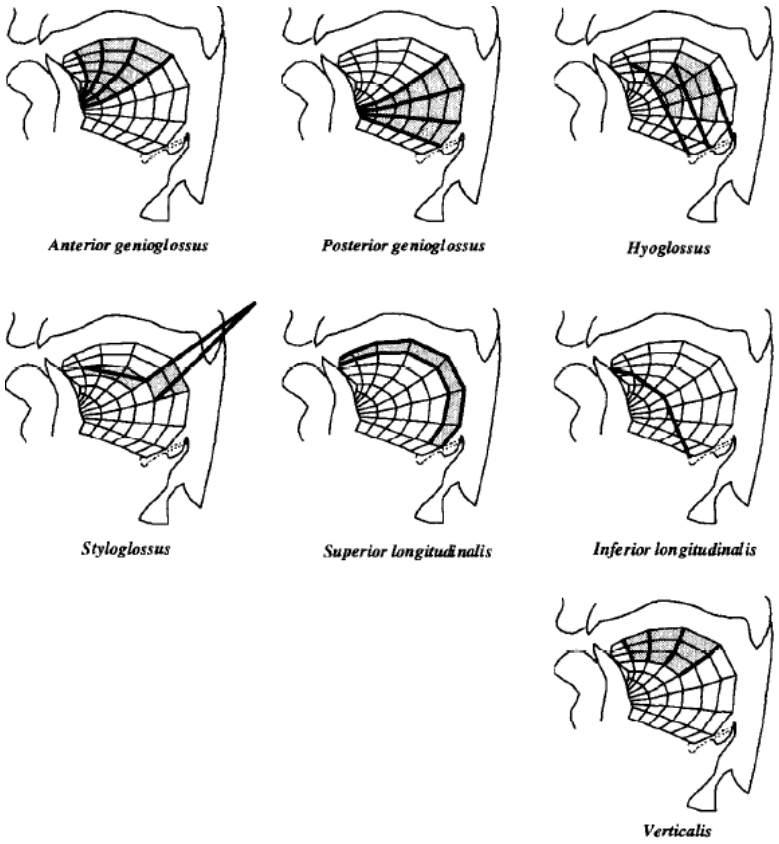

Figure 2.8: Two-dimensional model from Payan and Perrier (1997). Thick lines indicate muscles; shaded elements are those which change their elastic properties during muscle contraction.

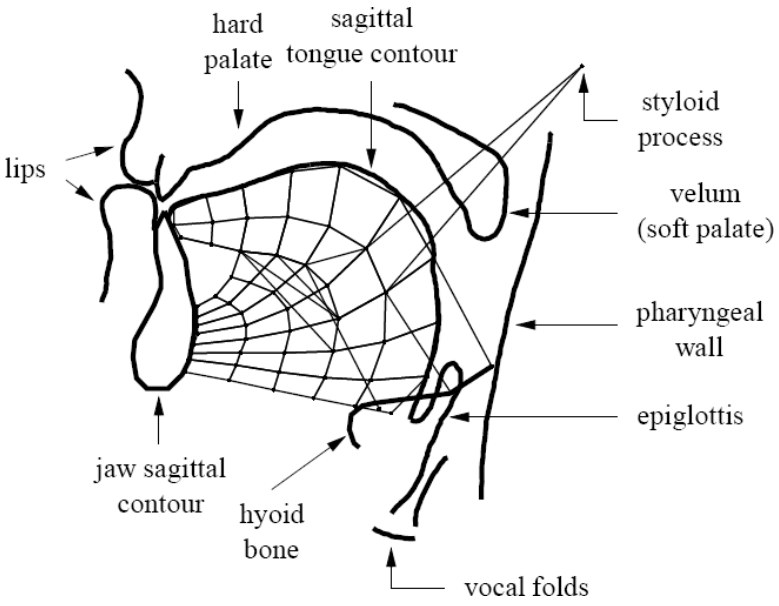

Figure 2.9: Two-dimensional tongue model from Sanguineti et al. (1998).

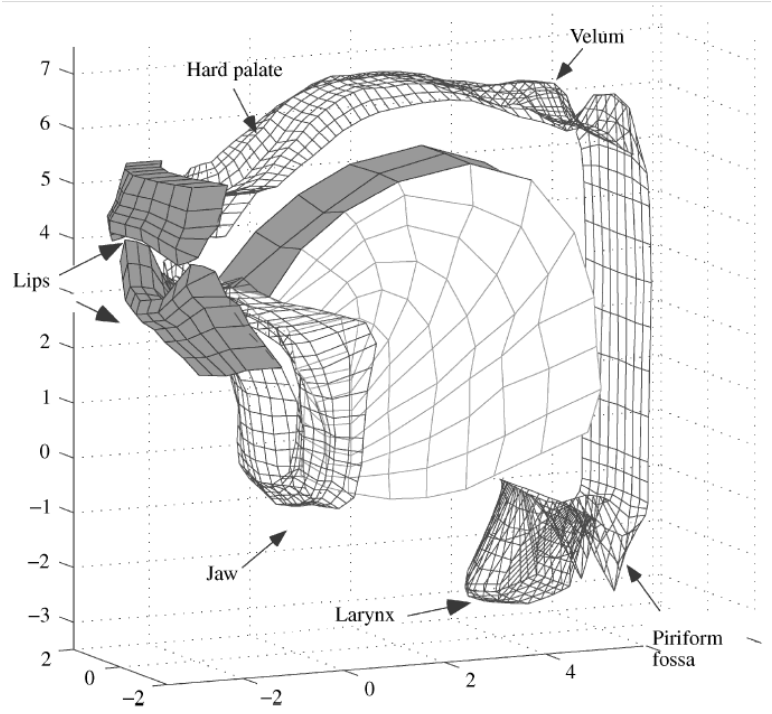

Figure 2.10: Two-and-a-half-dimensional Dang & Honda tongue model; this figure is from Dang and Honda (2002).

The Dang and Honda model does not use continuum mechanics, but is rather a network of mass points connected by spring-and-dashpot elements, similar to Perkell (1974). Their model includes eleven muscles, the seven muscles of the tongue with three separate controls for genioglossus, as well as the geniohyoid and mylohyoid. They also include jaw and hyoid motion, using the model of Laboissière et al. (1996).

Dang and Honda have made more extensive use of their model than other researchers. The use the model to quantitatively investigate speech motor control, particularly with the Equilibrium Point Hypothesis (Feldman, 1986). Dang and Honda (2002) used the model to predict the shape of the articulators, given an acoustic signal. This problem is otherwise complicated by the fact that there are infinite vocal tract tube shapes that produce a given acoustic output (Atal et al., 1978), but Dang and Honda found that the physiological model constrained the number of possible tube shapes such that the one-to-many mapping was eliminated.

2.2.2 Three-dimensional models

The modeling effort of Kiritani et al. (1976) is presented as an continuation of Perkell (1974)'s work. They produced the first model of the tongue that uses the finite element method. Geometrically their model improves upon Perkell's by adding a third dimension. They model one half of the tongue (assuming symmetry about the sagittal plane) using 14 elements, each of which is further subdivided into tetrahedra for the solution procedure (Figure 2.11). Like Perkell, they assume linearly elastic passive behavior. They do not mention the inclusion of an incompressibility constraint in their model.

The Kiritani et al. model was further developed by Kakita et al. (1985), who increased the precision of the mesh to include 30 elements, and included a complete vocal tract model. An innovation produced by the authors is the use of actual anatomical data—those reported in Miyawaki (1974)—to determine the geometry of the model. The model is shown overlaid on the anatomical data in Figure 2.12. The authors included the remaining vocal tract structures in a fairly schematic way: they approximated the rigid structures with cylinders and ellipses, included a single

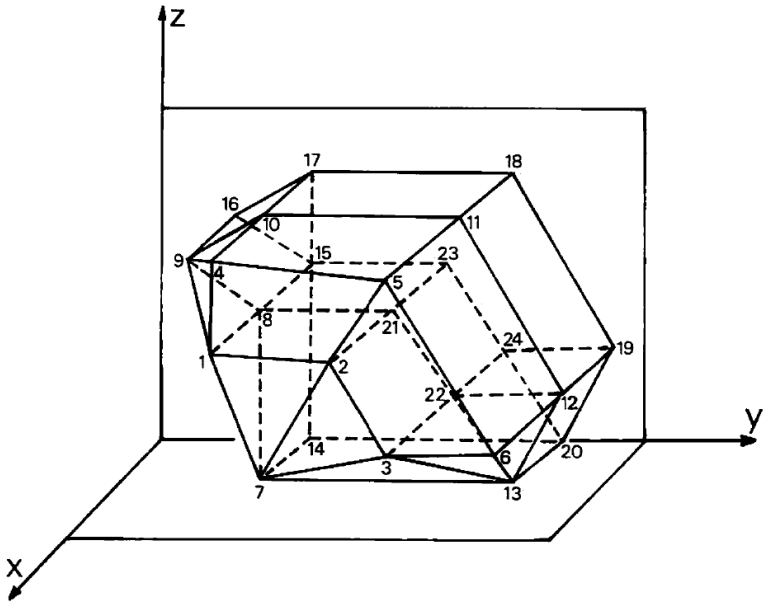

Figure 2.11: Three-dimensional tongue model from Kiritani et al. (1976).

jaw/hyoid position parameter, and a representation of the pharyngeal constrictors. With these elements in place, they were able to calculate vocal tract area functions, so that they could calculate vowel formants.

Another three-dimensional finite element model of the tongue was created by Hashimoto and Suga (1986), which included a mesh of 492 elements to represent one half of the tongue (Figure 2.13).[15] They assumed linear elasticity of tongue tissue. An innovative application of this model was to determine muscle contraction levels based on X-ray images of static images.

Wilhelms-Tricarico (1995, also reported in Wilhelms-Tricarico 1996) was the first to model passive stress using an incompressible hyperelastic material, representing an important advance over previous models, which had assumed linear elasticity. The model also included a viscous stress component. His finite element mesh also

[15]This may be the finest mesh used in a finite element model to date, in spite of many models having been published in the interim. The mesh used by Gerárd et al. (2003) might be finer; see Fn. 16.

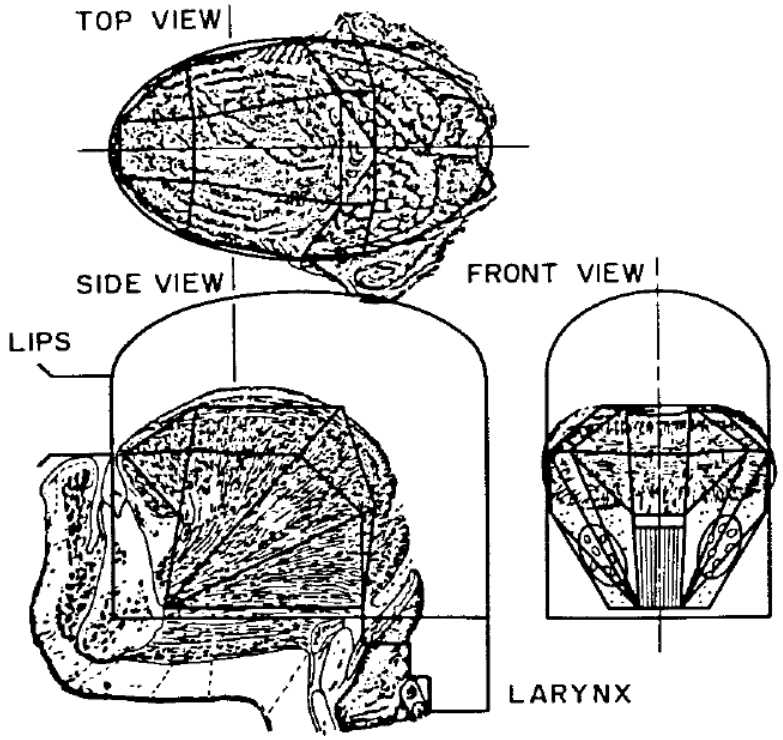

Figure 2.12: Three-dimensional tongue model from Kakita et al. (1985).

represented an increase in precision over Kiritani et al. (1976), modeling half of the tongue with 42 elements (Figure 2.14). The model included eight muscle control parameters, which corresponded to the typically defined eight muscles of the tongue, excluding palatoglossus, but with different controls for two regions of genioglossus.

Wilhelms-Tricarico subsequently refined his mesh (Wilhelms-Tricarico, 2000a), to include 740 elements (Figure 2.15).[16] His approach was to begin with a mesh of approximately the correct shape, and then to manipulate the node locations to conform to the Visible Human data set. It is not clear precisely how the dataset was used, since both the male and female data sets are mentioned in the description of

[16]The number 740 is not to be found in Wilhelms-Tricarico (2000a), but has been provided by van den Doel et al. (2006), who use the mesh. It is unclear whether this figure indicates the number of elements in one half of the tongue (making the effectice total 1480), or if it indicates the total number of elements in the model (half of which would be subject to symmetry restrictions).

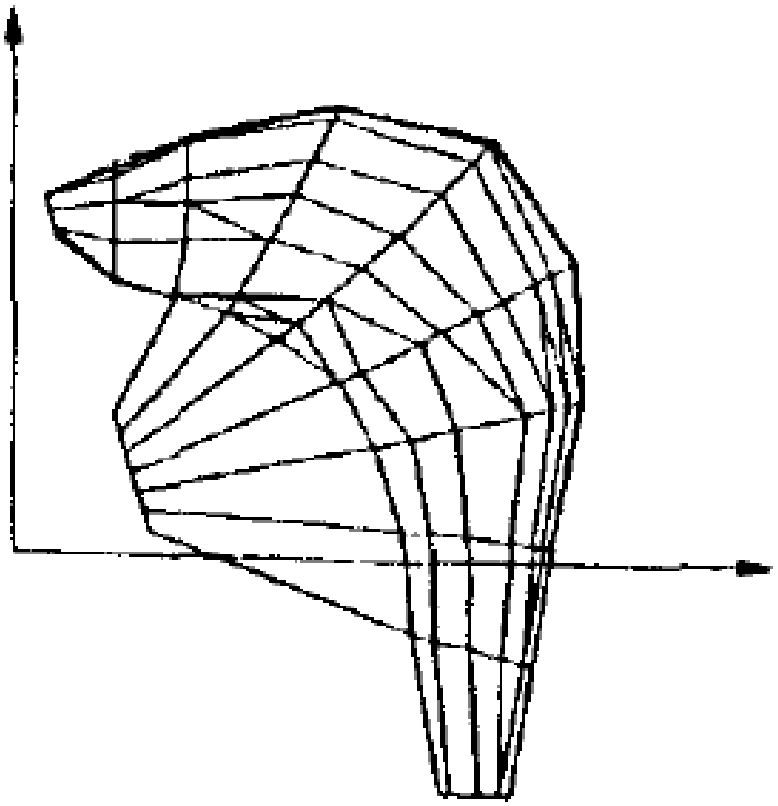

Figure 2.13: Three-dimensional tongue model from Hashimoto and Suga (1986).

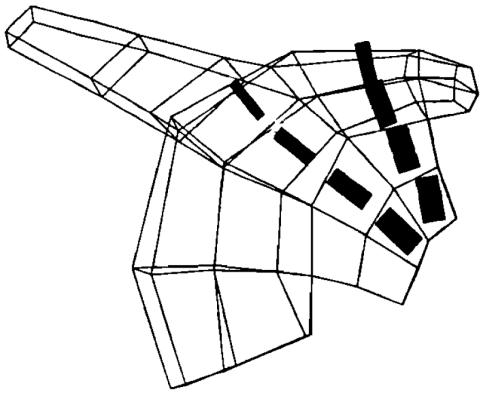

Figure 2.14: Finite element mesh from Wilhelms-Tricarico (1995). The black regions represent the contractile elements representing the anterior genioglossus.

methods; it seems likely that the model is actually an interpolation between the two data sets. Comparison of Figure 2.15 with the model in the present paper, which is shown in Figure 5.4 and is based solely on the female dataset from the Visible Human Project, reveal fairly substantial differences in overall shape.

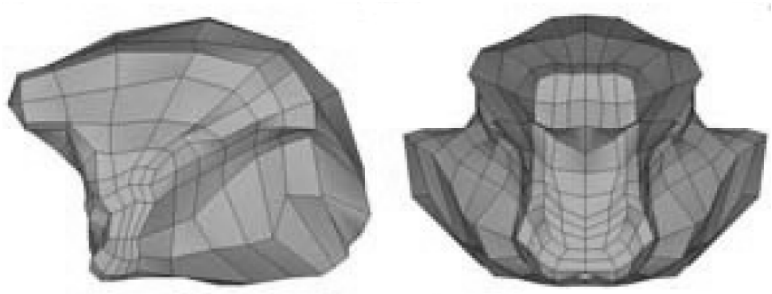

Figure 2.15: Finite element mesh from Wilhelms-Tricarico (2000a); figure from Gérard et al. (2004).

The Wilhelms-Tricarico (2000a) mesh has been used in further finite element models of the tongue. Gerárd et al. (2003) and Gérard et al. (2004) use a nearly-incompressible hyperelastic muscle description based on myocardial tissue. Gerárd et al. (2005), however, performed a tissue indentation test to determine a Yeoh form hyperelastic material description for tongue muscle specifically. This material was used in the model presented in that paper, and also in Gérard et al. (2006). In all of these models muscle contraction is approximated by applying forces directly to nodes in the model. The muscles used are the same was Wilhelms-Tricarico (1995), except that the mylohyoid is included as well. The authors also include the mandible and hard palate in the model. A further interesting aspect of the model in Gérard et al. (2006) is that the mesh was warped to conform to MR images of a particular speaker, which could enable validation studies to be performed with that speaker.

Fujita et al. (2007) expand the two-and-a-half dimensional model of Dang and Honda (1998) to three dimensions (i.e., making it thicker), as shown in Figure 2.16. Beyond this expansion, the authors indicate that other model parameters remained the same. The paper contains two notable advancements in modeling. First, theirs is the first tongue model to allow for asymmetrical tongue postures, which opens the

range of applications considerably. Second, the authors altered their model to mimic the anatomy of a tongue with a tumor, before and after partial glossectomy. Their model results were qualitatively similar to those of an actual partial glossectomy patient.

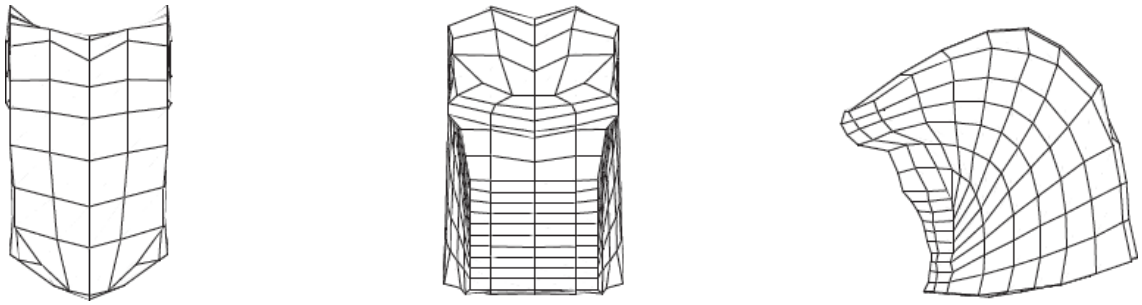

Figure 2.16: Finite element mesh from Fujita et al. (2007).

2.2.3 General Comments

In the absence of complete knowledge of the anatomy and physiology of the tongue, and given computational limitations, simplifications in modeling efforts are inevitable. This section critiques some of the simplifications that have been made in previous models, and discusses ways in which the present model overcomes some of those limitations.

Many tongue models are limited to two dimensions: a sagittal section of the tongue. This understandable limitation reflects the paucity of available three-dimensional tongue data, and the increased difficulty associated with a higher-dimensional model. Nevertheless, the problems associated with this simplification are manifold. The first is an obvious anatomical one: not all of the muscles of the tongue lie on the midsagittal plane. Styloglossus, hyoglossus, and inferior longitudinalis are not found on the midsagittal plane at all. Including these muscles in a two-dimensional model introduces geometrical distortions from the start. Most significantly, the fibers of transversus run approximately perpendicular to the midsagittal plane, so that this muscle cannot be included in the model at all. Although Perkell (1974, 105, footnote 2) indicates that is not very significant in determining

the posture of the tongue, muscular hydrostat theory and the results presented in this dissertation (cf. Chapter 6) suggest otherwise.

The second key limitation of a two-dimensional simulation is its inability to model the incompressibility of the tongue, a property thought to be key to its function (Kier and Smith, 1985). A two-dimensional model cannot capture this property of the tongue, since volume is not represented. Constraining the area occupied by the tongue in two dimensions—a common approach in two-dimensional models—is a gesture in the right direction, but is not the same restriction. The tongue is *not* constrained to have the same volume in the sagittal plane, since muscle tissue can be displaced laterally. As an illustration of this, consider Figures 6.4 and 6.5, which have quite different sagittal profiles, though their volume is nearly identical.

A third problem is that a two-dimensional model does not allow the study of lingual behaviors that are asymmetrical about the sagittal plane. Most every speech sound involves an approximately symmetrical tongue shape, but for many other lingual behaviors this is not the case. Feeding is an obvious example: the bolus must be moved to one set of molars or the other by asymmetric tongue movement (Abd-El-Malek, 1955). Two-dimensional models cannot produce this behavior.[17]

Another major limitation of previous models is the coarseness of the meshes. Models have been constructed from between 16 and 740 hexahedral elements, and this coarse level of representation has essentially been accepted acritically. Superficially, there is the question of whether the mesh actually resembles a tongue. An uninitiated researcher would be forgiven for not recognizing that the meshes shown in Figures 2.14, 2.15, and 2.16 are intended to represent the same object. The geometrical disparity between the models indicates that at least one of the models is not a faithful representation of the tongue. It would also make comparisons of results from different models difficult or impossible.

Another important limitation imposed by a coarse mesh is in the specification

[17]This problem holds for two-dimensional models, but also for three-dimensional models that constrain the tongue to be symmetrical about the sagittal plane, as have all three-dimensional models to date, except that of Fujita et al. (2007).

of muscle fiber orientations, which are specified at the level of the element. It is particularly evident from Figure 2.8 that the geometry of the model is determining to a large extent the orientation of the muscle fibers. Although the shape of the model was derived from an anatomical image (an X-ray), the orientation of the contractile elements of the model is hardly less important. If these cannot be precisely specified, then the validity of the model is called into question.

Setting aside questions of validity and faithfulness of representation, there are computational consequences to having a coarse mesh. Coarse meshes can lead to computational errors, simply because the placement of nodes in the mesh can create computational artifacts. Not incidentally, there has never been a study of whether this problem affects biomechanical tongue models. But, by way of comparison with other biomechanical research, Blemker et al. (2005) modeled biceps brachii with 20,000 elements. The fact that the number of elements in published tongue models differs from that published for other muscles by two orders of magnitude may lead one to suspect that the results obtained from previous tongue models suffer from mesh artifacts. The answer to this objection is a mesh validation study, where mesh refinement is shown after a point not to significantly change a model's output. Such a study is presented for the present model in Section 4.8.

It is relatively rare among previous models for the model's geometry to have been based on an actual anatomical study. In those cases where the model *is* based on anatomical data (Payan and Perrier, 1997; Dang and Honda, 1998; Wilhelms-Tricarico, 2000a), the orientation of the muscle fibers has not been taken from the anatomical specimen.[18] Instead fiber orientation data is borrowed from anatomical descriptions, which are inevitably less precise than actual anatomical images. For instance, it can be appreciated in almost any plane of section that the fibers of verticalis are not all vertically oriented, and that neither are the fibers of transversus all transversely oriented (cf. Figure 3.6). This is addressed in the present work by basing fiber orientations on an actual anatomical specimen (Section 3.2.3).

In considering the history of biomechanical tongue modeling, one does not sense

[18]Dang and Honda (1998) obtained *some* fiber orientation data from their MR images.

that models are becoming cumulatively more complex, or that they are converging onto a single set of parameter values. For instance, Perkell (1974, 72) notes that the output of his model depends largely on parameter values that were estimated rather than being measured experimentally. He indicates that these parameters should be revised in subsequent modeling, but it is not clear that this has been done, even in models that are the conceptual descendants of Perkell's. It is perhaps also telling that, 28 years after Kiritani et al. (1976) created the first three-dimensional tongue model, papers were being published that used two-dimensional models.

The reason for these limitations has not, of course, been malfeasance on the part of researchers, but rather the very real limitations imposed by computational costs, and the difficulty involved in implementing the finite element method. Computational limitations lessen annually. The advent of user-friendly finite element analysis software has also been significant for tongue modeling. It seems that until relatively recently, finite element models of the tongue have been hand-coded; Gerárd et al. (2003) are the first to mention use of commercial finite element software (ANSYS). Such steep technical prerequisites have limited the accessibility of the models, and also the types of papers that are written. As Dang and Honda remark, "most of the studies of physiological articulatory models, especially 3D models, focus on theory and methodologies for constructing a model" (2001, 416). This is as much to say that, in general, papers about finite element modeling of the tongue tend to be about the finite element method, not the tongue. Given recent advances in computer technology, these difficulties should be reduced. It is modestly hoped that since the present model is being made publicly available, future modeling efforts can build on it, rather than starting from scratch.

CHAPTER 3

Anatomy of the Tongue

> For our tongue, as being carnal, is not sufficient to minister to the rapidity of the human mind, inasmuch as that is of a spiritual nature, for which reason our word is restrained within us, and is not at once expressed as it has been conceived by the mind, but is uttered by successive efforts, just as the tongue is able to serve it.
>
> —Irenaeus, Against Heresies, III.XXVIII.4

This chapter begins with a description of the anatomy of the tongue, drawing from previous research. Then, the anatomical data collection performed in the current study is presented; this involved segmentation of the tongue into various regions, and determination of the orientation of the fibers for the muscles of the tongue.

3.1 Tongue anatomy

The human tongue is primarily composed of skeletal muscle, with a covering of mucosa.[1] Muscle fibers attach either to bone via tendons, to internal septa, to other muscles, or to the mucosal covering of the tongue.

The tongue's external bony attachments are to the mandible, the hyoid bone, and the styloid process of the skull. These attachments are illustrated in Figure 3.1. Anteriorly, the tongue connects to the posterior aspect of the mandibular symphysis (the place where the two halves of the mandible are joined). Posteriorly, the tongue connects to the hyoid bone, a bone which is suspended in the neck by muscles, and

[1] Apart from those places where an external resource is cited, the facts in this section may be found in any anatomical reference that includes a description of the tongue. The author has particularly relied on Zemlin (1988).

which forms the upper boundary of the larynx. The hyoid is depicted in Figure 3.2; the location of the styloid process on the skull is shown in Figure 3.3.

The tongue also has soft tissue interfaces. It is attached to the velum by palatoglossus, a muscle that originates in the tongue and inserts into the velum. Finally, the tongue's base is connected by fascia to the supralaryngeal muscle that lies immediately inferior to the tongue and forms the muscular floor of the mouth, the mylohyoid. The mucosa of the tongue is also continuous with the mucosal covering of the floor of the mouth.

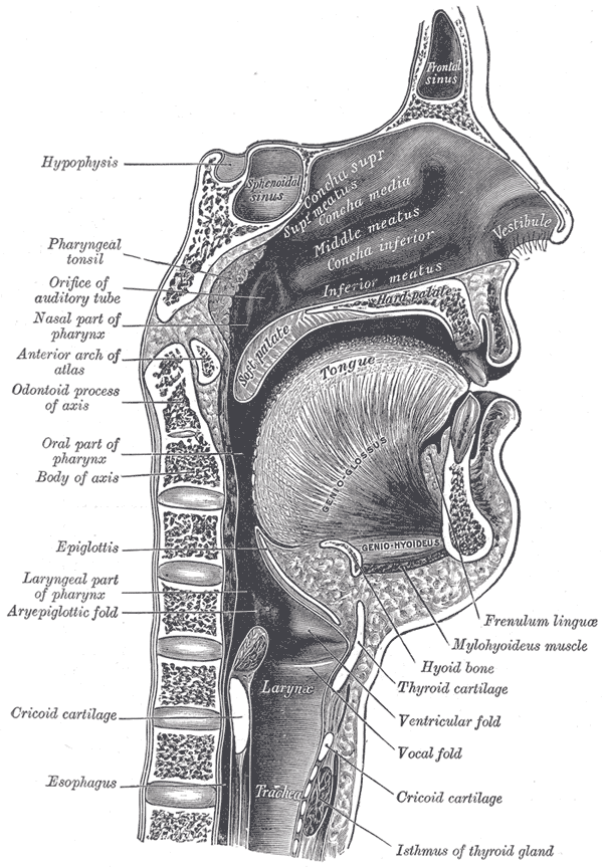

Figure 3.1: Bony attachment of the tongue. The hyoid and mandible are cut in cross-section. From Gray (1918).

The tongue has several fibrous septa, as described by Abd-El-Malek (1939). The most prominent of these is the median septum, which lies on the sagittal plane, dividing the tongue in half (Figure 3.4). The paramedian septum conforms to the

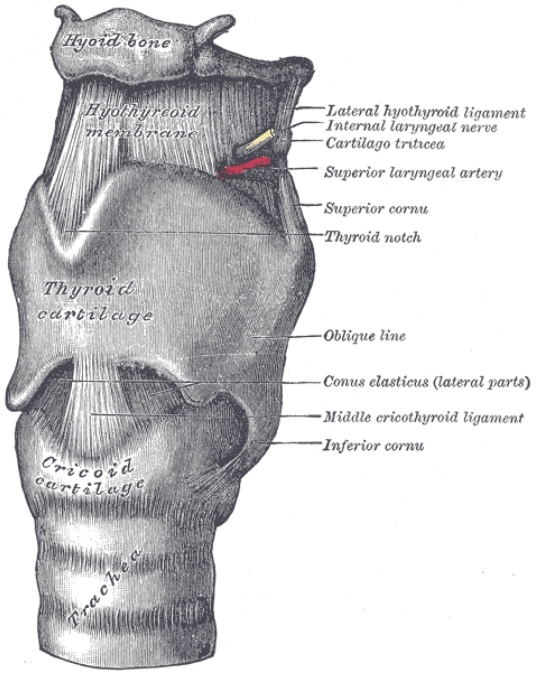

Figure 3.2: Perspective view of larynx. The hyoid bone also supports the major cartilaginous structures of the larynx. From Gray (1918).

lateral margins of genioglossus, attaching posteriorly to the hyoid bone (Figure 3.5). The lateral septum is interposed between inferior longitudinalis, hyoglossus, and styloglossus; the presence of the lingual artery in this septum makes it a convenient anatomical landmark (Figure 3.6).

The musculature of the tongue has been described as being composed of eight paired muscles. These are illustrated in Figure 3.7. The muscles of the tongue are classically divided into the extrinsic and intrinsic muscles (following Salter, 1852). The extrinsic muscles of the tongue are those that originate from bone; the intrinsic muscles have no bony attachment (and are thus intrinsic to the tongue). The extrinsic muscles are genioglossus, hyoglossus, styloglossus, and palatoglossus. The remaining muscles—transversus, verticalis, superior longitudinalis, and inferior longitudinalis—are intrinsic to the tongue.

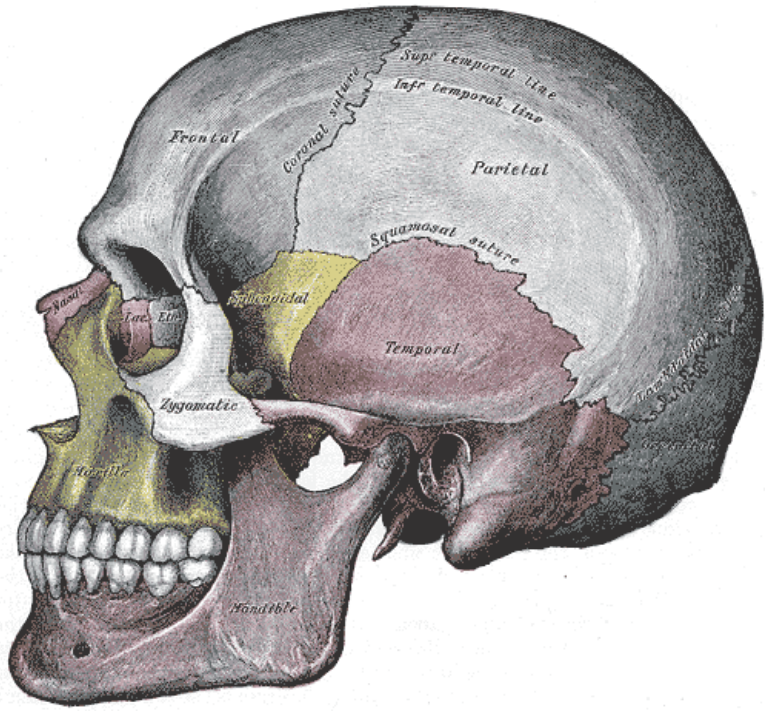

Figure 3.3: Side view of the skull. The styloid process is just posterior (in the image, to the right of) the mandible. From Gray (1918).

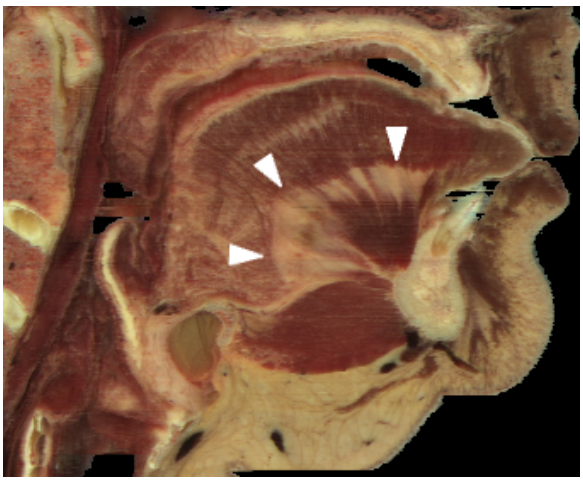

Figure 3.4: The median septum (white) cut along the sagittal plane.

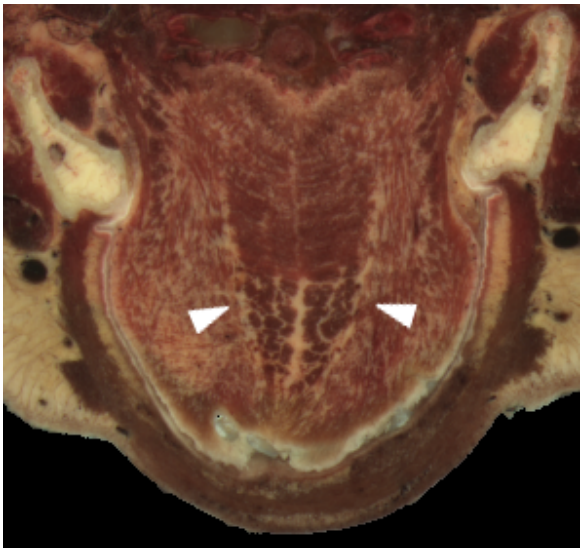

Figure 3.5: The paramedian septum cut on the axial plane.

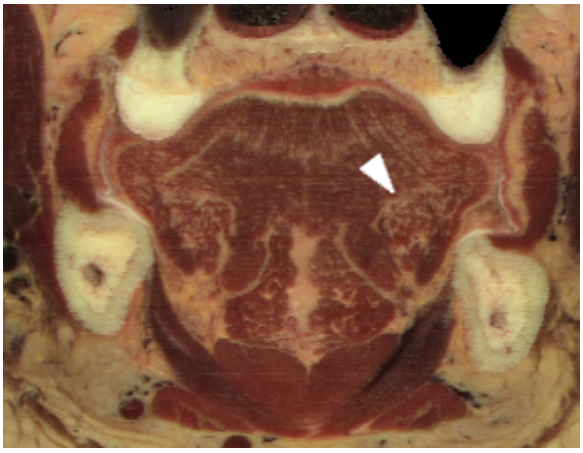

Figure 3.6: The lateral septum cut on the axial plane. Note the presence of the lingual artery.

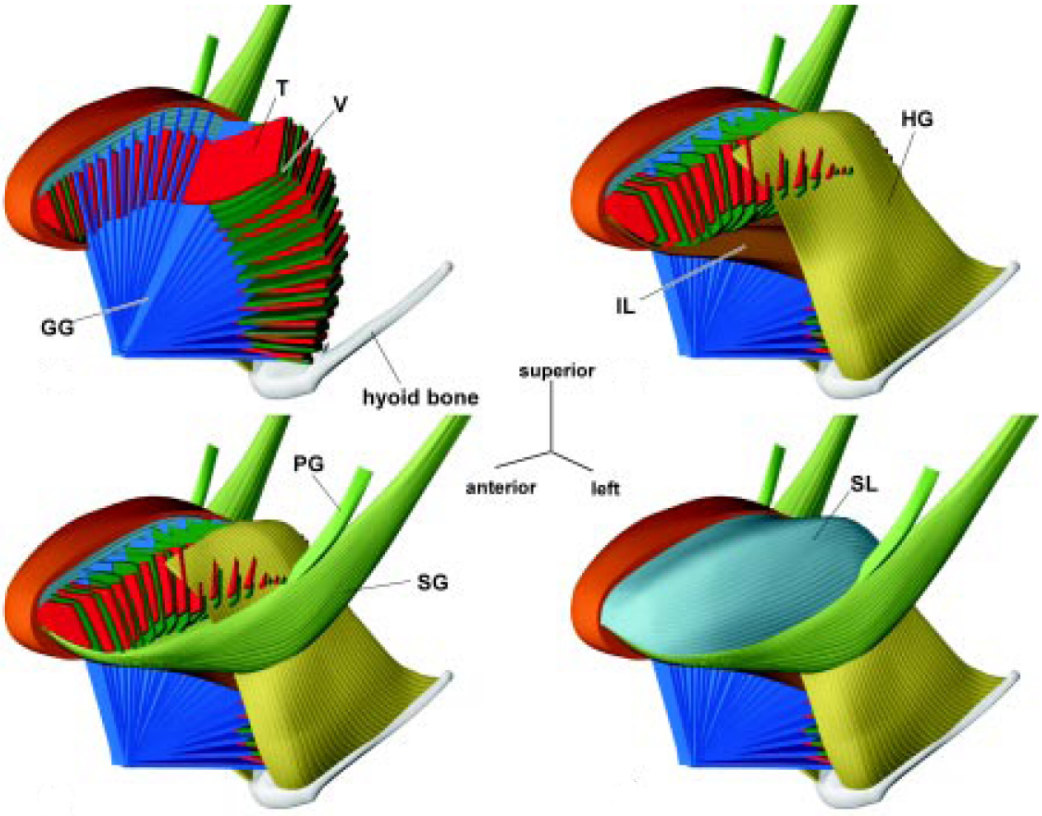

Figure 3.7: Muscles of the tongue, from Takemoto (2001). GG genioglossus, T transversus, V verticalis, HG hyoglossus, IL inferior longitudinalis, S superior longitudinalis, PG palatoglossus, SG styloglossus.

3.1.0.1 Genioglossus

Genioglossus is a fan- or wedge-shaped muscle that originates from the posterior aspect of the mandibular symphysis and radiates anteriorly and posteriorly. The majority of its fibers insert into the mucosa that covers the tongue, while the inferior fibers insert onto the hyoid bone. As discussed below, the more distal portion of genioglossus interdigitates with transversus and verticalis.

3.1.0.2 Hyoglossus

Hyoglossus originates from the hyoid bone, coursing superiorly and slightly anteriorly to insert into the body of the tongue.

3.1.0.3 Styloglossus

Styloglossus originates from the styloid process of the skull, and descends into the tongue. Its fibers insert into the body of the tongue. After inserting into the tongue, the fibers divide into two bundles. The anterior bundle continues anteriorly along the inferior surface of inferior longitudinalis, lateral to hyoglossus, to the tip of the tongue (Abd-El-Malek, 1939; Takemoto, 2001). The posterior bundle of fibers penetrates hyoglossus and courses medially to insert into the lingual septum; these bundles of fibers also coursed slightly anteriorly and inferiorly (Abd-El-Malek, 1939; Takemoto, 2001; Saito and Itoh, 2007). A diagrammatic representation of the two bundles is given in Figure 3.8. Salter (1852, cited by Saito and Itoh (2007)) claimed that fibers from the posterior bundle of styloglossus fuse with those of transversus, which Saito and Itoh deny. Both Takemoto (2001) and Iskander and Sanders (2003) also indicate that the fibers run parallel to the inferior portion of transversus in their specimens. One way or the other, this indicates that the course of the fibers are at least very similar in orientation to those of transversus.

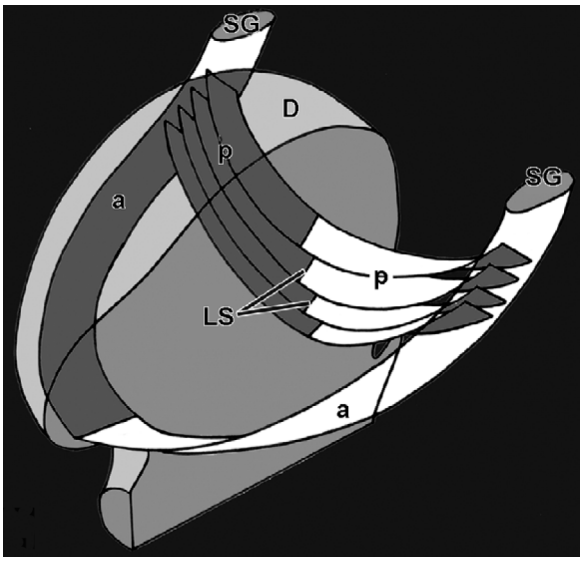

Figure 3.8: Schematic of the course of styloglossus from Saito and Itoh (2007). Styloglossus (SG) descends from the styloid process, dividing into anterior (a) and posterior (p) fiber bundle after entering the tongue dorsum (d). Fibers from the posterior bundle insert into the lingual [medial] septum (LS).

3.1.0.4 Palatoglossus

Palatoglossus originates from the soft palate, and inserts into the body of the tongue. In previous anatomical investigations, palatoglossus has variously been considered a muscle of the tongue or a muscle of the palate. Its contraction might either lower the velum, or raise the tongue body. Miyawaki (1974) does not consider palatoglossus to be a muscle of the tongue and does not report on it, while Abd-El-Malek (1939) and Takemoto (2001) both consider it a muscle of the tongue (though Takemoto was unable to observe all of the fibers of the muscle because of its small size).

In the present work palatoglossus will be considered a muscle of the palate—and thus ignored—for three reasons. First, the tongue is large relative to the velum and the muscles of the velum; it seems unlikely that the velum could be stabilized sufficiently to move the tongue relative to the skull, rather than being moved itself, upon contraction of palatoglossus. Second, palatoglossus is innervated by the Cranial

Nerve XI (the accessory nerve), and not by Cranial Nerve XII (the hypoglossal nerve), as is the case with all other tongue muscles. Conversely, all muscles of the soft palate are innervated by either the accessory (XI) or trigeminal (V) cranial nerves. This does not prove that the muscle is used primarily to move the palate, but it is suggestive of it. The third reason, which is purely pragmatic, is that palatoglossus was not visible in the anatomical data set employed. Of these reasons for ignoring palatoglossus, the first would be most compelling, if it were shown to be true.[2]

An indirect validation of the decision comes from Section 6.1, which shows that styloglossus—a larger muscle that one would expect to have a greater effect on tongue posture than palatoglossus—contributes very little to tongue posture.

3.1.0.5 Transversus and Verticalis

Transversus and verticalis form the bulk of the tongue. They fill the width of the tongue, lying vertically between the superior and inferior longitudinalis muscles. Fibers from transversus and verticalis are arranged in alternating sheets of muscle, perpendicular to the main curve of the tongue (see Figure 3.7). Fibers of transversus generally course horizontally, though they fan out slightly toward the right and left edges of the tongue. Verticalis fibers are likewise generally vertical, but fan at the superior and inferior margins.

Genioglossus, transversus, and verticalis partially overlap with one another. Takemoto (2001) describes the relation in detail. He states that the fibers of genioglossus and verticalis, where they are coextensive, run together in the same lamina, and are difficult to separate from one another. Transversus fibers run in the alternating lamina. He indicates approximately one hundred repetitions of these lamina.

There is some disagreement in the literature over the extent to which transversus and verticalis overlap. Abd-El-Malek (1939) indicates that transversus extends

[2]One way to test the validity of the assumption would be to create biomechanical models of both the tongue and the soft palate, to determine the respective effects of contracting palatoglossus on the tongue and the soft palate.

further anteriorly and posteriorly than does verticalis. Miyawaki also notes a more limited range for verticalis, saying that "[t]he muscle is found only in the superior half of the tongue" (1974, 25). Takemoto (2001), however, indicates that the muscles are coextensive even at the base of the tongue, and specifically claims to have observed verticalis fibers at the base of the tongue. Miyawaki cites coronal sections from the posterior of the tongue to support his claim. Verticalis fibers, however, if present, would be parallel to genioglossus, and a coronal section taken from the posterior region of the tongue would cut them in cross-section (i.e., perpendicular to the long axis of the fiber). It is possible that for this reason Miyawaki did not observe the full extent of verticalis. Conversely, Takemoto's observations were based on oblique sections of the tongue that followed the "curve" of the tongue from tip to root. It seems that this allowed him to see the full extent of verticalis.

3.1.0.6 Superior Longitudinalis

Superior longitudinalis is a sheet-like muscle that covers the top of the tongue; its fibers course longitudinally along the surface of the tongue. Takemoto (2001) notes the absence of consensus on the extent of fibers of superior longitudinalis. Abd-El-Malek (1939) indicates that most of the fibers terminate in the posterior third of the tongue, while Miyawaki (1974), Takemoto (2001), Iskander and Sanders (2003), and Slaughter et al. (2005) indicate that the fibers are present from the tip to the root, where the tongue attaches to the hyoid bone.

3.1.0.7 Inferior Longitudinalis

Inferior longitudinalis lies beneath transversus and verticalis; its fibers course front-to-back as well. Some researchers have identified external attachments of inferior longitudinalis. Abd-El-Malek (1939) indicates that posteriorly inferior longitudinalis attaches to the anterior surface of the hyoid, and also to the body and root of the greater horn of the hyoid. He also notes a minor attachment to the stylohyoid ligament. Salter (1852, 1126), however, calls this muscle "entirely intrinsic." Miyawaki

(1974) also does not note any external attachment. Takemoto (2001) observed only the attachment to the body and root of the greater horn of the hyoid; none of his figures display this connection, however.[3]

3.2 Collection of Anatomical Data

A three-dimensional finite element model of any structure requires a three-dimensional geometric description of the structure. The anatomical descriptions reviewed in the previous section, and discussed in Chapter 2, provide two-dimensional data, but none provide sufficient detail to reconstruct the three-dimensional geometry of the tongue with precision. Takemoto (2001) produced a three-dimensional tongue model based on impressions from his tongue dissections, which is shown in Figure 3.7. The two reasons that this three-dimensional model were not used that it is not publicly available, and that it represents a somewhat stylized tongue, since it was based on the author's understanding of the tongue, and not on a particular anatomical specimen.

Given the complexity of the tongue, it would be reasonable (if impractical) to maintain that nothing short of an individual representation of each muscle fiber would be a fair representation of the tongue's musculature. Scanning electron microscopy studies by Saito and Itoh (2003, 2007) in particular seem to justify that level of detail in representation. While such an approach would be ideal, it is not possible to develop such a representation with presently-available technology, nor would it be computationally practical to incorporate such a level of detail into a biomechanical model. The present work aims for a middle ground, producing a precise-yet-tractable description of the architecture of the tongue. Future biomechanical tongue models will benefit from more detailed anatomical studies of the architecture of human tongue muscles, particularly those that offer sufficient detail to make a 3D render-

[3]In the secondary literature, inferior longitudinalis is always described as lacking external attachments. Although the presence of external attachments could have been omitted purely for simplicity, this probably also reflects the beliefs of competent anatomists (e.g., Zemlin, 1988) that the attachments are insignificant.

ing possible. A new anatomical description of the tongue was generated using data from the Visible Human Project (VHP National Library of Medicine, 1987).[4] This section describes how these data were collected and analyzed.

The Visible Human Project is an initiative of the National Libraries of Medicine, which provides digital anatomical data to researchers. There are three datasets that are applicable to the study of the tongue. For each dataset, a donated cadaver was frozen, and sliced at short intervals along the long axis of the body. Photographs were made and digitized for each plane of section. The Male Dataset consists of images of axial slices of a male cadaver, taken at 1 mm intervals; the side of each pixel of the anatomical data is 0.33 mm long. The Female Dataset is similar, but the axial slices were just 0.33 mm thick. This is a slightly more than a threefold increase in axial precision over the Visible Male, since voxels are 0.33 mm $\times$ 0.33 mm $\times$ 0.33 mm, rather than 0.33 mm $\times$ 0.33 mm $\times$ 1 mm. Another dataset has been made available recently, which consists exclusively of sections of a male's head, with 0.174 mm precision in all directions (nearly twice the precision of the Female Dataset). All other things being equal, the dataset with the greatest precision is most desirable. An issue arises with the 0.174 mm dataset, however, in that the cadaver's tongue is notably distorted, as shown in Figure 3.9. The Visible Human Project web site does not provide details about the preparation of this cadaver, but the distortion of the tongue is consistent with what is observed in a cadaver that has desiccated with the mouth in a wide gape. Therefore this dataset was judged to be unacceptable, and the Female Dataset was used instead. A sample parasagittal image from the head of the Female Dataset is provided in Figure 3.10. Researchers affiliated with the Visible Human Project are not aware of any pathological anatomy in the head or neck region of the Female Dataset (Dr. Charles Sneiderman, p.c.).

[4]Use of data from the VHP was suggested by the work of Wilhelms-Tricarico (2000b, 2005). Iskander and Sanders (2003) have use the dataset to investigate lingual anatomy. The dataset has also be used for musculoskeletal modeling by Teran et al. (2005).

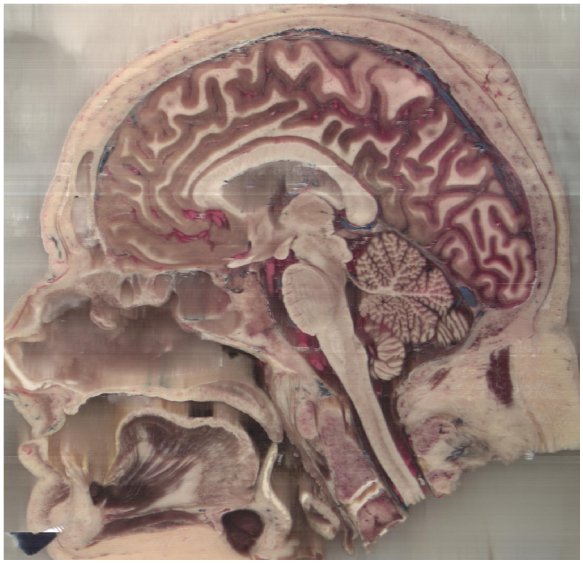

Figure 3.9: Parasagittal slice of the head of a 72-year-old male. The tongue is noticeably distorted.

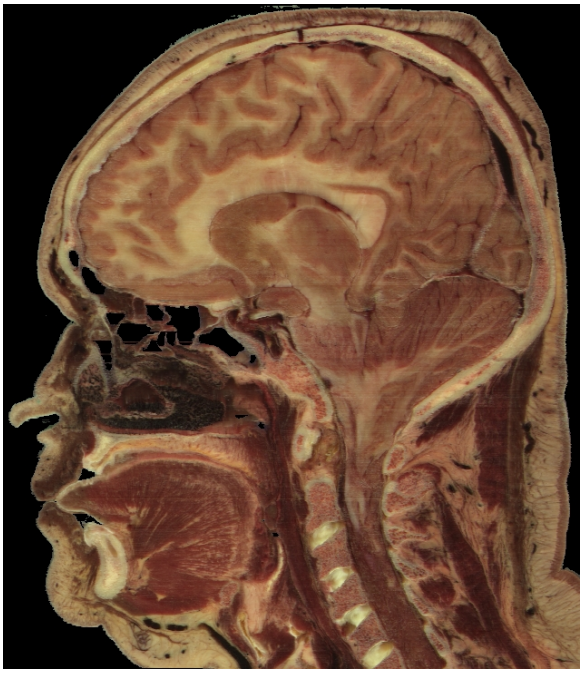

Figure 3.10: Parasagittal slice of the head of the Visible Woman.

3.2.1 Volumetric Data

Volumetric imaging techniques can be used to reconstruct three-dimensional structures from serial two-dimensional images. This section provides a conceptual overview of those techniques by illustrating the reconstruction of the skull.

The Female Dataset consists of serial photographs of a three-dimensional volume. These images are two dimensional, but their relationship in three-dimensional space is known: each image was taken after removing a 0.33 mm-thick slice of the body. As such, each image can be oriented in three-dimensional space. Figure 3.11 shows such a configuration.[5] One can determine which part of the images correspond to bone; Figure 3.12 updates Figure 3.11 by emphasizing those parts of the image.

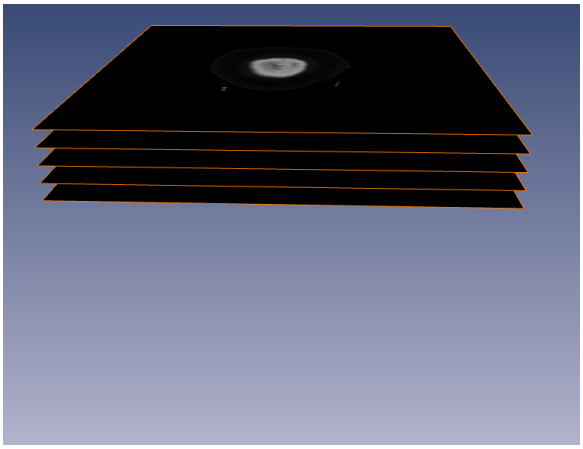

Figure 3.11: CT images from the Visible Woman Dataset, arranged in three dimensions.

As two-dimensional representations, the images have no thickness. Since each image is 0.33 mm apart from each other image, though, each image could be considered to have 0.33 mm thickness. That is, each pixel can be considered a voxel. A voxel is a rectangular prism (here, a cube) that can be used to represent three-dimensional shapes just as a pixel is used to represent two-dimensional shapes. Under this interpretation, the data can be considered to be three-dimensional. When the

[5]For ease of illustration, CT scans of the head have been used. This is purely a convenience; it is identical conceptually to using the full color images taken of the sliced, frozen cadaver.

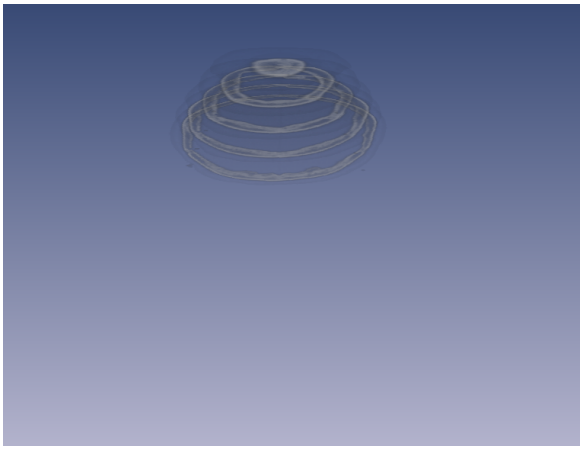

Figure 3.12: The images from Figure 3.11, with non-bone image portions deemphasized.

"thickened" images are arranged, a three-dimensional volume results. The outcome of this process, applied to the entire skull, is shown in Figure 3.13.

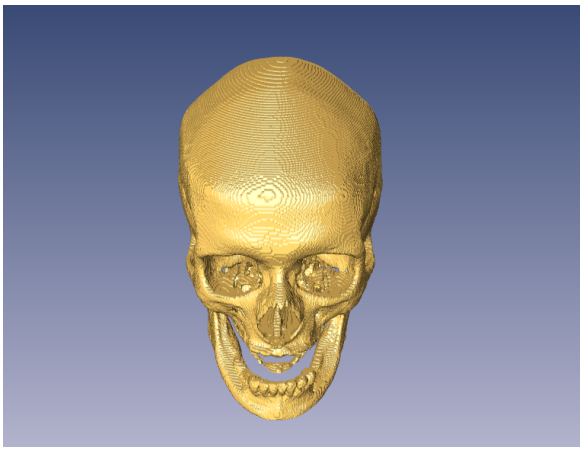

Figure 3.13: A three-dimensional reconstruction of the Visible Woman's skull from axial images.

3.2.2 Image segmentation

Images from the Visible Woman dataset were segmented manually using the Amira software package (Visage Imaging, Carlsbad, CA). This software allows a user to

view a volumetric dataset in axial, coronal, or sagittal planes of section. Image labels (in this case, corresponding to particular muscles) are indicated by shading regions of the image, with an interface similar to standard image processing packages. Although the dataset was in full color, Amira is only able to show grayscale images. The grayscale image corresponded to the luminance of the full color image.

For computational efficiency, the entire Visible Woman was not loaded into Amira, but rather a subportion of the dataset was extracted. The boundaries where chosen to include the pertinent anatomical data. The superior and posterior bounds were selected to include the styloid process. The anterior bound included the lips. The inferior bound included the hyoid bone. The left and right boundaries included the widest extent of the tongue on the axial plane. See Appendix B for further detail.

Labels were assigned to the images for mucosa of the tongue, the mandible and hyoid bones, superior longitudinalis, inferior longitudinalis, genioglossus, hyoglossus, styloglossus, and transversus/verticalis. The resolution of the dataset did not permit identification of the individual lamina of transversus and verticalis/genioglossus. Therefore Takemoto's claim about the extent of verticalis was accepted, and transversus and verticalis were labeled as a single muscle.

The anatomical images were segmented using the aforementioned anatomical works to assist in identifying the regions that correspond to particular muscles (particularly Abd-El-Malek, 1939; Miyawaki, 1974; Takemoto, 2001). Additionally, the full color anatomical images could be viewed for reference using another computer program. For each muscle, the images were considered in at least two planes of section, and frequently a third. After an initial segmentation, the image labels were examined and revised at least twice (once in a different plane of section, and once in the original plane of section after modifications had been made).

Below, notes are provided on the segmentation of each region, along with segmentation of bone and the identification of attachment points of the tongue. An isosurface reconstruction of each region, shown in relation to the mucosal covering of the tongue, is provided in each section. (The colors of the isosurface follow those employed by Takemoto (2001); see Figure 3.7.)

3.2.2.1 Genioglossus

Genioglossus (Figure 3.14) was the most straightforward muscle to segment. In parasagittal sections the radiation of its fibers is quite clear. The segmentation was confirmed by examining the data in coronal and axial sections.

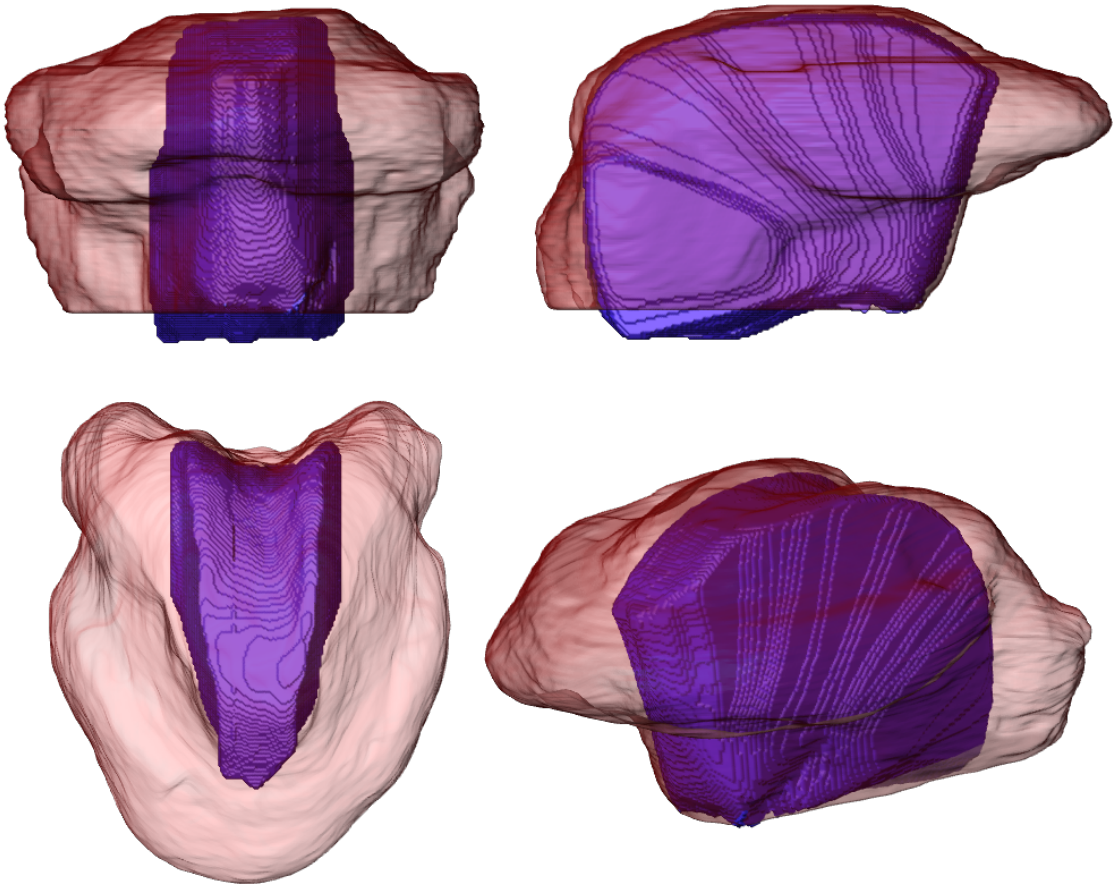

Figure 3.14: Isosurface reconstruction of genioglossus. In this and subsequent figures, the muscle is shown in the context of the mucosal covering of the tongue. The color of the region follows the color scheme of Figure 3.7, from Takemoto (2001)

3.2.2.2 Hyoglossus

Hyoglossus (Figure 3.15) was a relatively simple muscle to identify. The only ambiguous area was found in the superior portion of the muscle, where its fibers intersect with styloglossus.

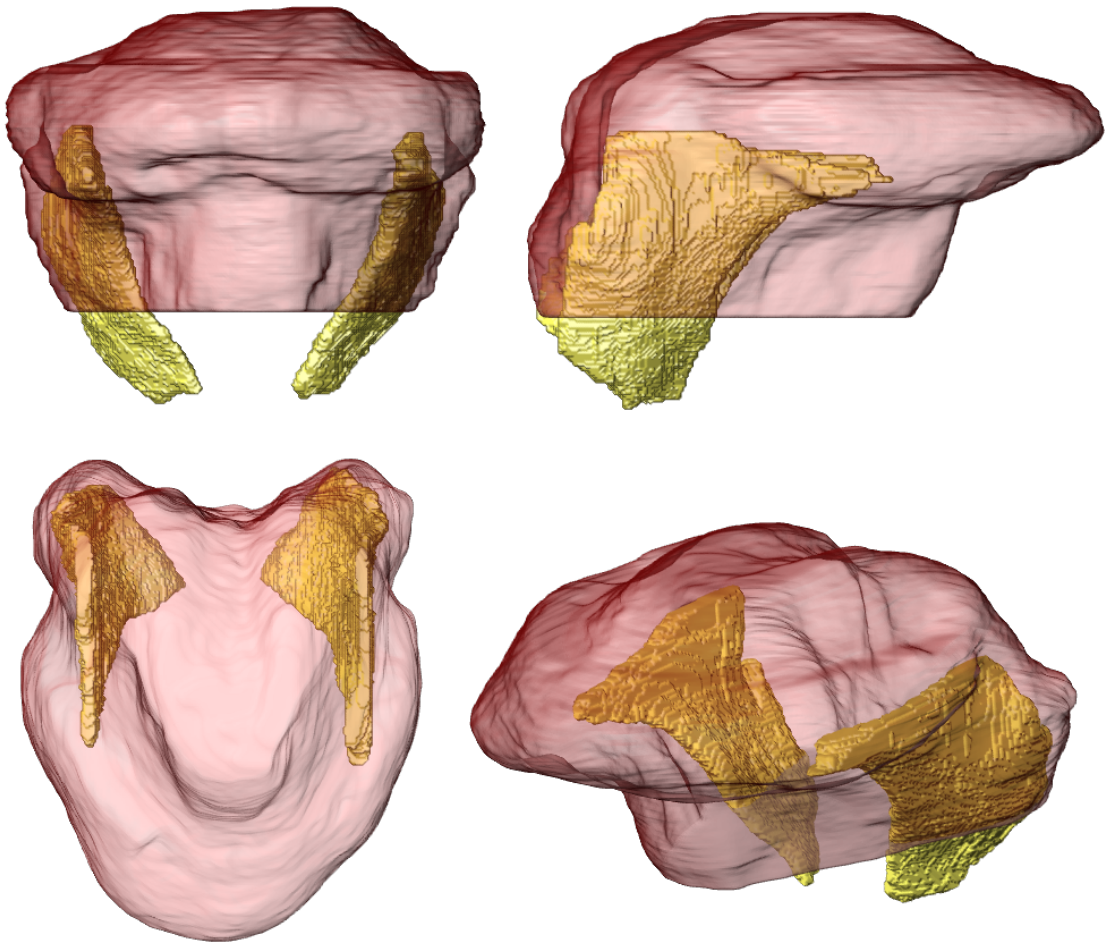

Figure 3.15: Isosurface reconstruction of hyoglossus.

3.2.2.3 Styloglossus

Styloglossus (Figure 3.16) was quite easy to segment. The portion of the muscle that is extrinsic to the body of the tongue is easily identifiable in any plane of

section. Within the tongue, however, the complex fiber directions observed above were not observed. Specifically, the anterior bundle was not traced all the way to the tongue tip, and the posterior bundle was not visible at all. This probably indicates that some of the image regions attributed to inferior longitudinalis rather belong to styloglossus, and that some transversus/verticalis fibers belong rather to styloglossus.

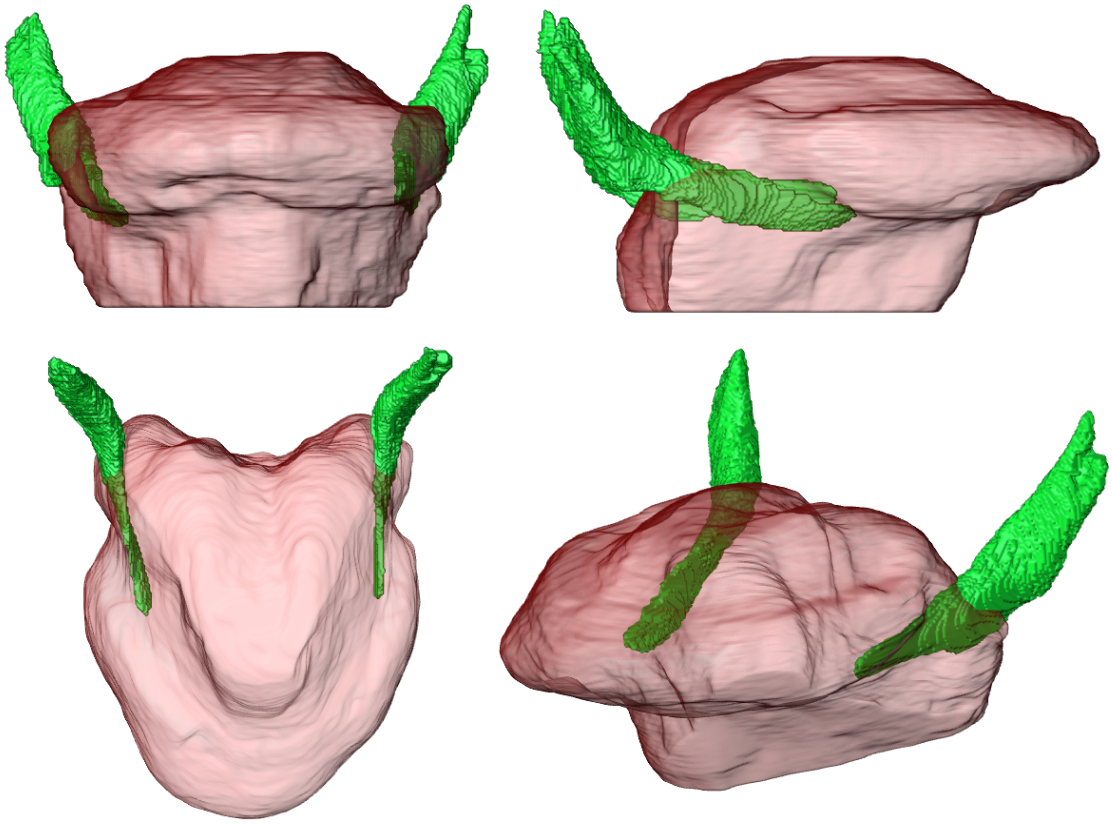

Figure 3.16: Isosurface reconstruction of styloglossus.

3.2.2.4 Transversus and Verticalis

Segmentation of transversus and verticalis (Figure 3.17) was straightforward. As noted above, no effort was made to differentiate between the lamina of these muscles.

Rather, both were segmented as a single region.[6]

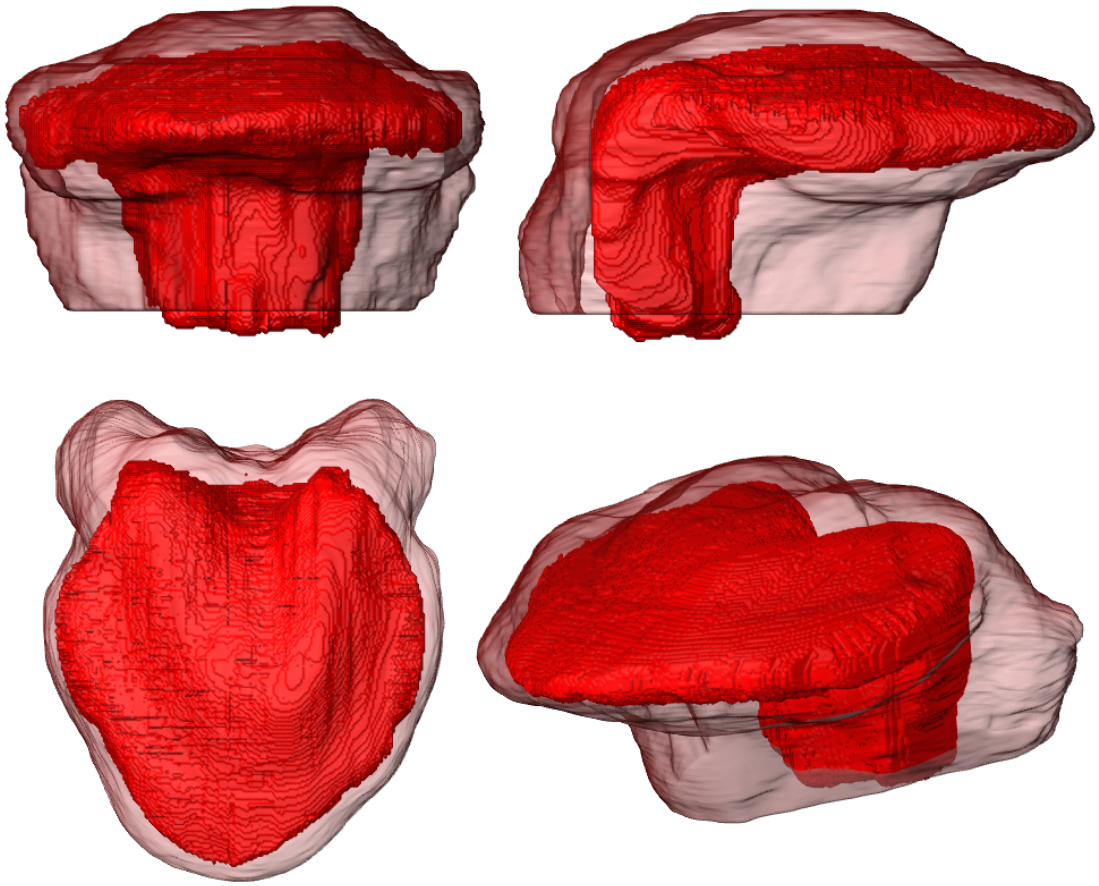

Figure 3.17: Isosurface reconstruction of transversus/verticalis.

3.2.2.5 Superior Longitudinalis

Superior longitudinalis (Figure 3.18) was easy to segment. Fibers were observed from the tip of the tongue to the root, in agreement with the majority of previous studies (Miyawaki, 1974; Takemoto, 2001; Iskander and Sanders, 2003).

[6]See Chapter 4 for description of how the different fiber orientations were determined in the finite element model.

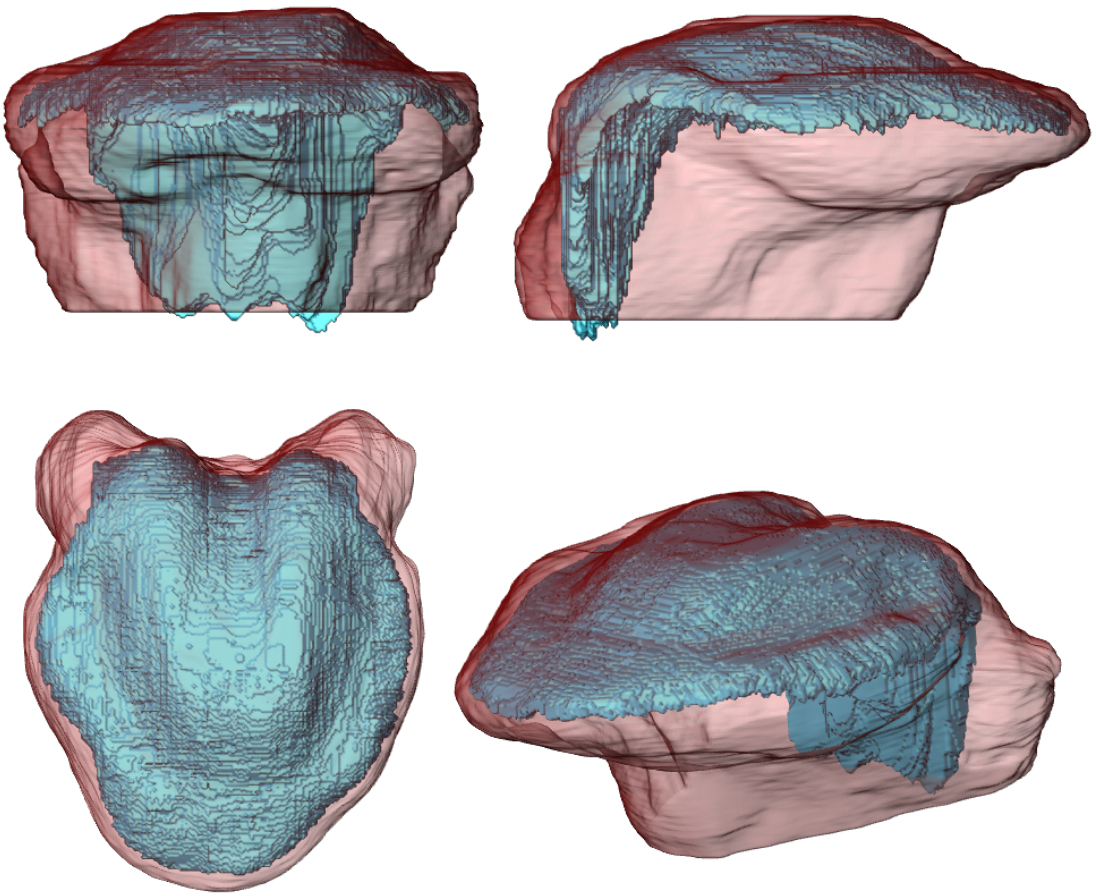

Figure 3.18: Isosurface reconstruction of superior longitudinalis.

3.2.2.6 Inferior Longitudinalis

Inferior longitudinalis (Figure 3.19) was the most difficult muscle to identify. The extent of the muscle was not always clear, particularly at the anterior and posterior limits. Consequently, the muscle was repeatedly revised with use of all three planes of section. No external attachments of inferior longitudinalis were observed.

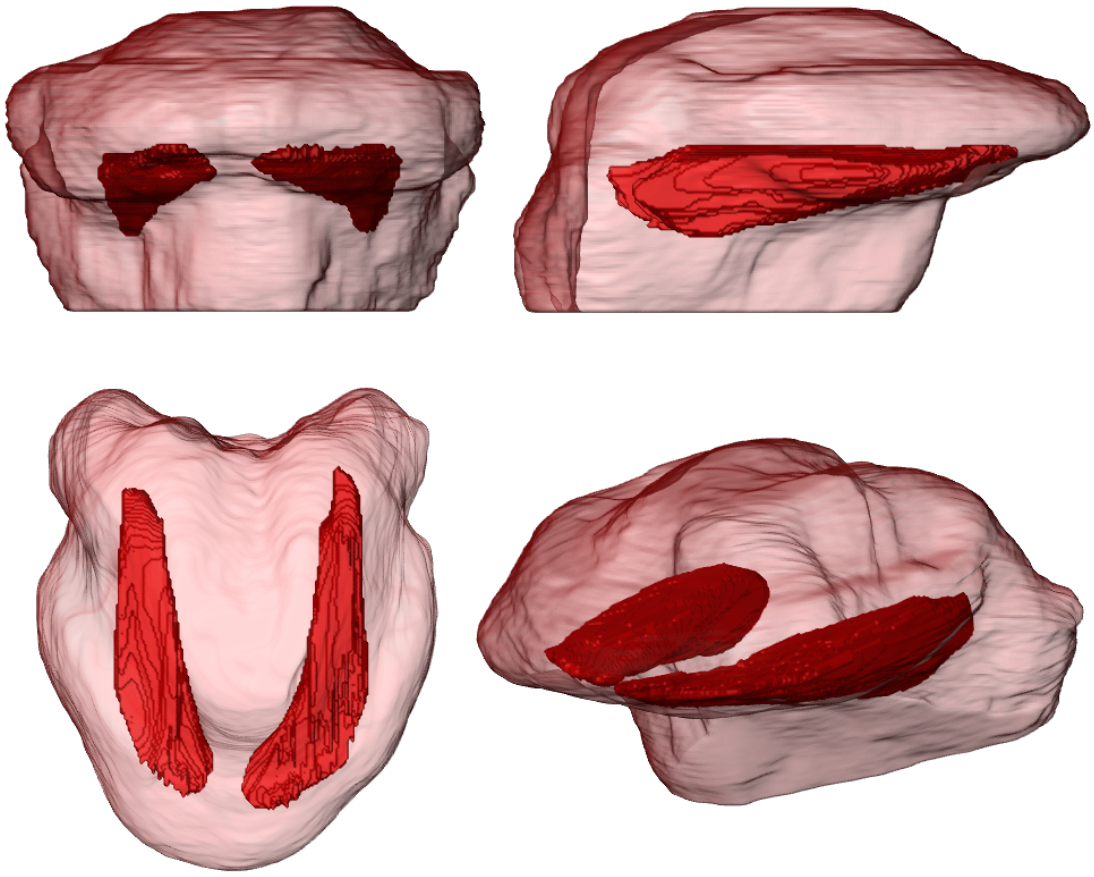

Figure 3.19: Isosurface reconstruction of inferior longitudinalis.

3.2.2.7 Mucosa

Identifying the mucosal surface of the tongue (e.g., Figure 3.19) was generally simple. There were ambiguous regions in the posterior-most portions of the tongue (near

the epiglottis); it was also difficult to identify the lateral boundary of mucosa at the base of the tongue. As a compromise, the lowest level of mucosa represented was the lowest level that the mucosa could be identified in the axial sections. Consequently the segmentation of the mucosa has an abrupt lower boundary.

3.2.2.8 Bone

Both the mandible and the hyoid (Figure 3.20) were easy to identify in the anatomical data. These regions were *not* as carefully analyzed as the muscles, as their detailed shape was not crucial to the model. Note that the superior portion of the mandible was not fully identified, as the region of interest in the anatomical data was defined only to include the superior-most extent of styloglossus.

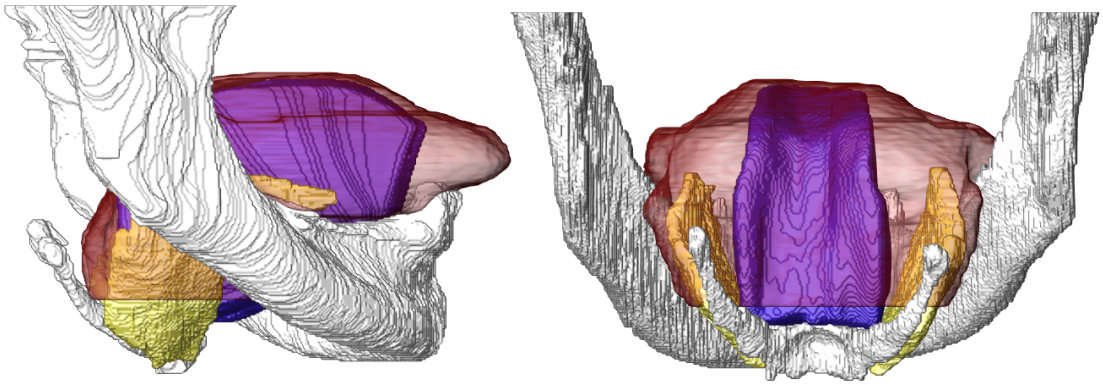

Figure 3.20: Isosurface of the bone, with attachments of hyoglossus and genioglossus onto the hyoid visible.

3.2.2.9 External attachment points

Attachment sites were identified in the volumetric data set. The sites, shown in Figure 3.21, were the anterior and posterior attachments of genioglossus (to the mandible and hyoid bone, respectively), the inferior attachment of hyoglossus (to the hyoid bone), and the superior attachments of styloglossus (to the styloid process). The attachment sites were easy to identify, as they corresponded to those regions of tongue muscle that are adjacent to bone. The particular regions were not very

carefully identified. This is because the purpose of identifying the attachments was to indicate a region within which the finite element nodes should be constrained. Since nodes within a certain distance of a voxel were constrained, there was not a need for great care in identifying the attachment sites. (This is described more fully in Chapter 4.)

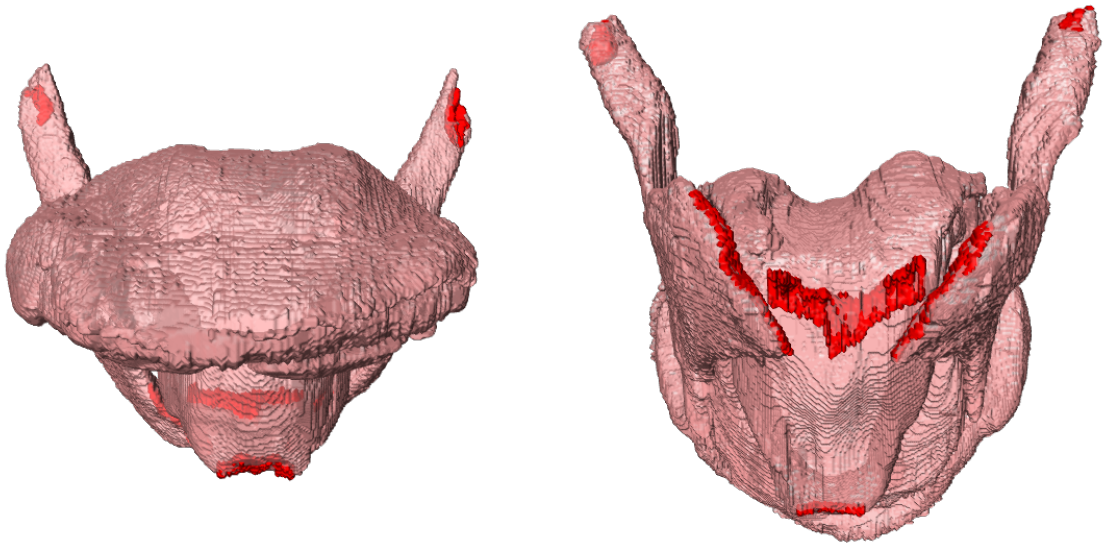

Figure 3.21: External attachments of the tongue (red), superimposed on a transparent volumetric representation of the muscles of the tongue. The views of from the front (left) and from a posterior and inferior position (right).

3.2.3 Muscle Fiber Orientations

Modeling the results of muscular contraction in the tongue requires data pertaining not just to the position of the muscle, but also to the orientation of the muscle fibers. This section describes how this information was obtained and represented.

For every location within the volume of a muscle, a vector is required to indicate the orientation of the muscle fibers. Mathematically, this is a $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ mapping: a 3D vector is defined for every point in a 3D volume. One way to obtain this mapping would be to manually define the muscle fiber orientation at every point of a muscle's volume at some level of precision—say, for each voxel. A more efficient strategy is

to define a small number of vectors for each muscle, and to obtain fiber orientations at intermediate points by interpolating. This approach was used by Legrice et al. (1997), for example, to describe the fiber orientations in the heart. To perform the interpolation, a generic vector interpolation algorithm was selected (Verestóy et al., 1999), which is described below. The selection of the interpolation algorithm was arbitrary. No claim is made that the method selected represents a more biologically correct algorithm than any other. Rather, the validity of the interpolation derives from the visual inspection of the interpolated vector field by the researcher.

The algorithm begins with a set of k ordered pairs, W, where the first member of the ordered pair is a point in the muscle volume, and the second is the orientation of the muscle fiber at that point.

$$W = \{(\vec{p}_1, \vec{v}_1), (\vec{p}_2, \vec{v}_2), (\vec{p}_3, \vec{v}_3), \ldots\}$$

To determine the vector $\vec{v}_n$ associated with a new point $\vec{p}_n$, we identify the subset of points in W, W', that is within 100 pixels of $\vec{p}_n$ (the choice of 100 pixels is arbitrary). We denote members of the subset $w_i = (\vec{p}_i, \vec{v}_i)$, $i = 1, 2, \ldots k$. We define two more variables: d_i^2 is the squared Euclidean distance from $\vec{p}_i$ to $\vec{p}_n$. d^2 is the mean of d_i^2 for all i.

We define a variable α_i to represent the closeness of each $\vec{p}_i$ in W' to $\vec{p}_n$. As the distance from $\vec{p}_i$ to $\vec{p}_n$ grows (relative to the mean distance of each $\vec{p}_i$ to $\vec{p}_n$), α_i decreases exponentially.

$$\alpha_i = \exp\left(-\frac{d_i^2}{d^2}\right) \tag{3.1}$$

The variable β_i represents the closeness of each $\vec{p}_i$ to $\vec{p}_n$ (α_i) as a proportion of the closeness of all $\vec{p}_i$ to $\vec{p}_n$. Hence $\sum_{i=1}^{k} \beta_i = 1$.

$$\beta_i = \frac{\alpha_i}{\sum_{i=1}^{k} \alpha_i} \tag{3.2}$$

The interpolated vector direction, $\vec{o}_n$, the sum of all $\vec{v}_i$, weighted by β_i.

$$\vec{v}_n = \sum_{i=1}^{k} \beta_i \vec{v}_i \tag{3.3}$$

As an example of an interpolation, consider the vectors plotted in Figure 3.22. Using the interpolation algorithm just described, a more densely populated vector field can be produced by interpolating between the pre-defined vectors. The interpolated vector field is shown in Figure 3.23.

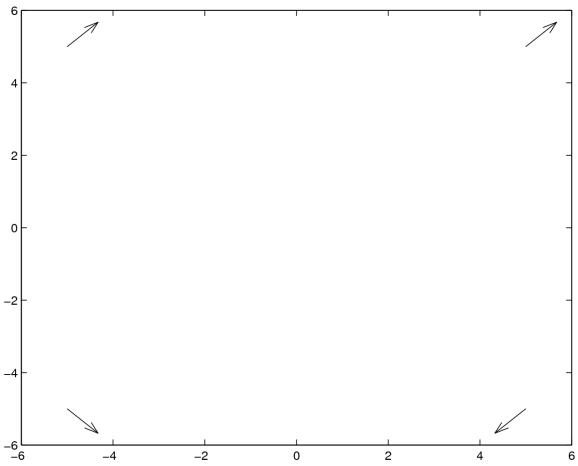

Figure 3.22: Some arbitrary vectors. (Values are dimensionless.)

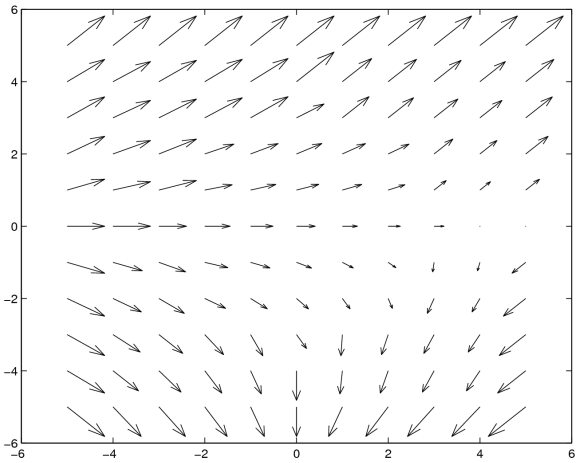

Figure 3.23: A vector field created by interpolating between the vectors in Figure 3.22. (Values are dimensionless.)

A Matlab interface was created to facilitate the task of defining the set of vectors for each muscle. A screenshot is shown in Figure 3.24. The interface updated the interpolated vector field every time a new vector was defined, so that the vector

field could be validated by visual inspection. It was possible to define and modify vectors in the axial, sagittal, and coronal planes of section. This was necessary since only two dimensions of a vector can be specified in a two-dimensional anatomical section.

The resolution of the Visible Human data set made it simple to identify the orientation of the muscle fibers for each muscle. The only muscle for which fiber orientation was not defined by clear striations in the muscle was styloglossus. The fiber orientation of this muscle was easy to determine based on its gross shape and reports from previous anatomical studies, however.

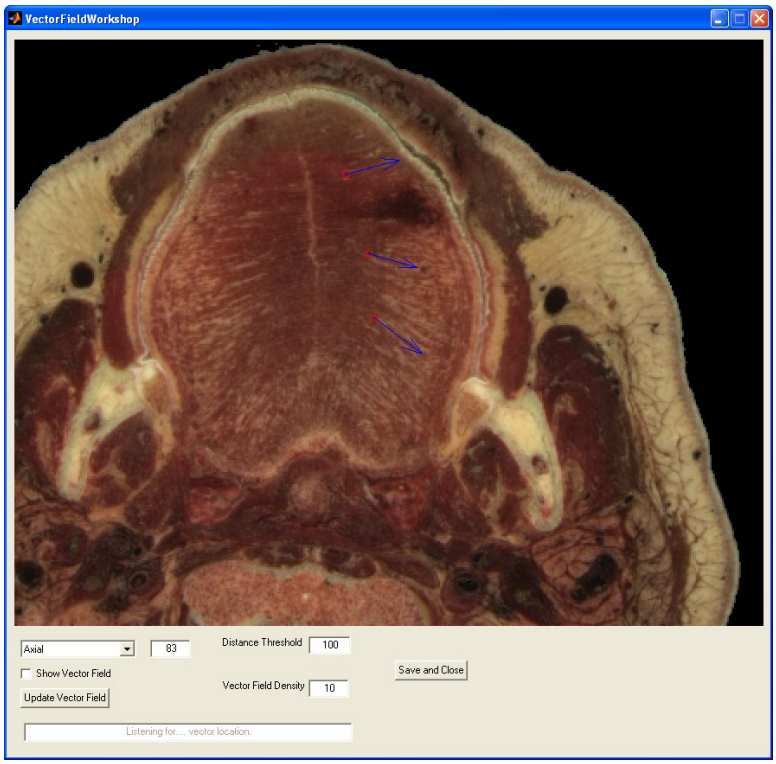

Figure 3.24: Screenshot of the custom Matlab interface used to defined vectors.

CHAPTER 4

The Finite Element Model

God is in the details.

—American proverb

The devil is in the details.

—American proverb

4.1 Introduction

A model of tongue deformation requires both continuum mechanics and the finite element method. This chapter describes the application of these mechanical methods to the creation of a tongue model. A brief introduction to continuum mechanics and the finite element method is offered. Following this, the particular details of the tongue model are presented.

In elementary physics, objects are treated as rigid bodies: bodies that do not change shape when loads are applied to them. In real life, all materials deform to some extent when external loads are applied. For certain materials and loading conditions, the assumption of rigidity is sensible, but for the tongue it is clearly inappropriate. The study of *how* materials deform is continuum mechanics.[1] With continuum mechanics, the deformation of an object can be calculated when the applied loads are known. The particulars of this calculation for muscle are described below in Section 4.5.

[1]The term "continuum" refers to another counterfactual (though often acceptable) assumption that matter, which is discrete at the atomic level, can be represented as a continuum. This assumption enables the use of calculus.

Use of the finite element method is motivated by the complex geometry of the tongue. No analytical equations are available for modeling the deformation of tongue-shaped objects: if equations could be derived, it would certainly be prohibitively time-consuming to do so. But the complex geometry of the tongue can be *approximated* with simpler three-dimensional shapes, like tetrahedra, and the mechanical properties of these shapes can be determined with relative ease. This is the kernel of the finite element method. Just as complex structures can be built from geometrically simple wooden blocks, so complex structures like a tongue can be represented with discrete elements. The finite element method was developed primarily in the 1950's and 1960's in order to model aircraft (see Barron 1993 for a full history), and has subsequently been applied to many other fields as well, including biomechanics.

A finite element model of a deformable object consists of a few basic parts. The first part is the geometric description of the structure to be modeled. The second part is the discretization of the part into simple elements—the finite element mesh. The third part is a description of a behavior of the material that the model is made of: its deformation in response to applied external loads, and, in the case of an active material like muscle, the stresses generated by the material itself. The fourth part is a set of boundary conditions on the model. The final part of the model is a description of the loads applied to the model. In the case of the tongue, the loads are provided mainly by muscle contraction (in this dissertation, solely by muscle contraction). With these parts of the model defined, it is possible to calculate how the model will deform. Each part of the model is described below.

A summary of the model, along with a review of the crucial assumptions made by the model, is given in Section 4.9.

4.2 Geometric description

The geometric description of the model is derived from the anatomical study reported in Chapter 3. From the segmentation of images, a volumetric representation

of the exterior of the tongue is produced, which is shown in Figure 4.1a. The representation is discrete: the volume of the tongue is represented by a collection of voxels.[2] Software programs that generate finite element meshes typically require that three-dimensional objects be represented as solid entities with continuously defined boundaries, rather than as volumetric objects. In the case of complex curved geometries, such as the tongue, solids are modeled as regions bounded by trimmed nonuniform rational b-spline surfaces (NURBS surfaces; Rogers 2001).

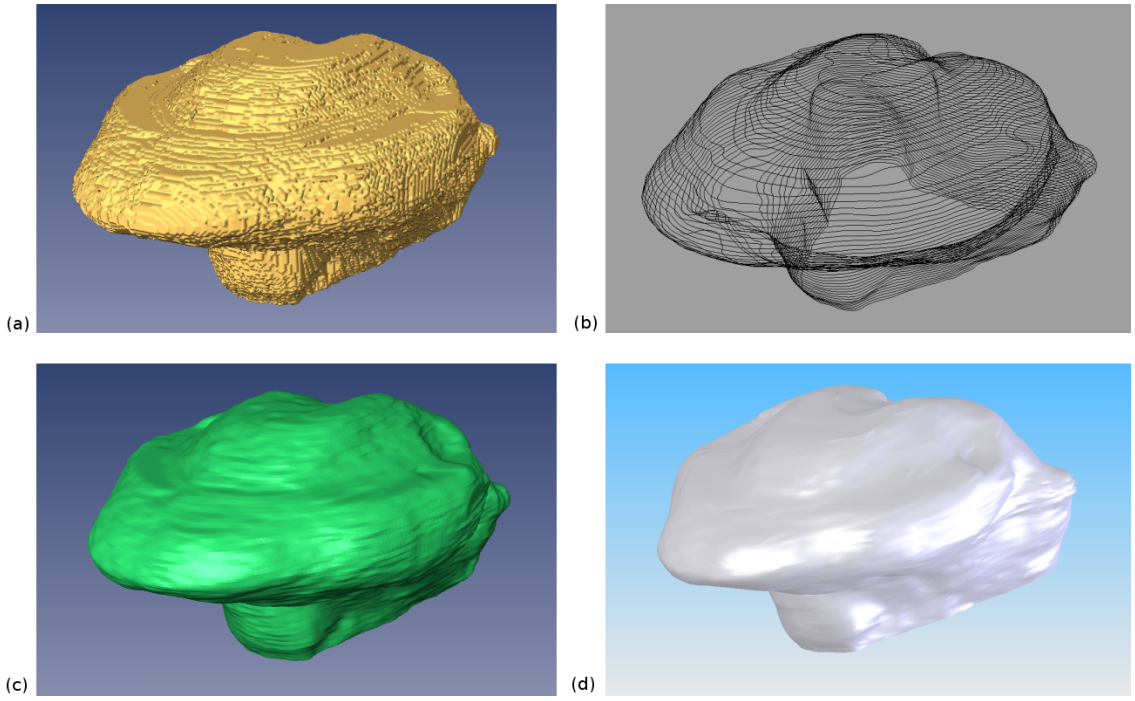

Figure 4.1: A conceptual overview of the process of creating a solid from a volumetric representation. (a) volumetric representation (b) surface representation (c) spline representation (d) solid representation.

Splines are mathematical functions used to interpolate between *control points*. A spline that interpolates between a set of points is given in Figure 4.2.[3] Splines are used for convenience to define continuous functions. A NURBS surface is conceptually an equivalent tool, used to define a surface in three-dimensional space. A

[2]See Section 3.2.1 for an introduction to voxels.

[3]In the present context, it can be assumed that the spline always passes through the control points, though this is not always the case.

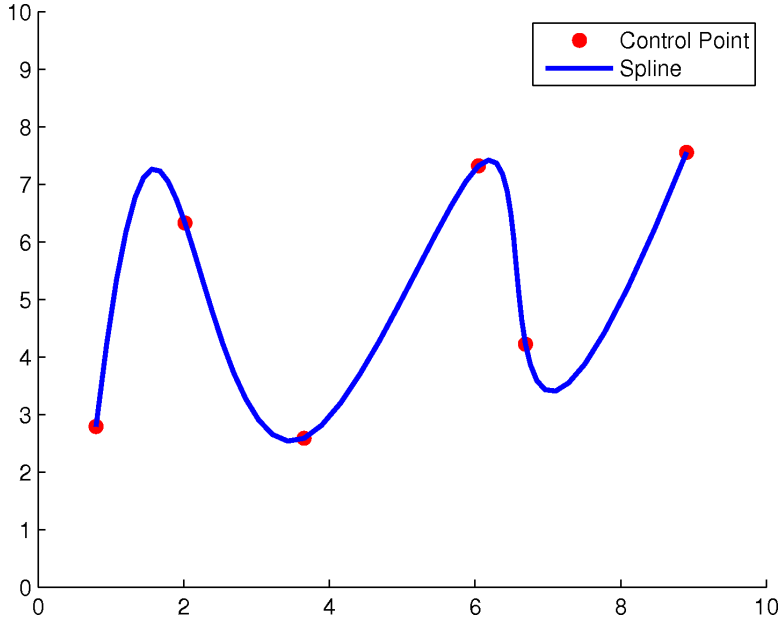

Figure 4.2: Illustration of control points and a cubic spline. (Values are arbitrary and dimensionless.)

single three-dimensional object is represented with multiple (at least four) NURBS surfaces.[4]

As it happens, there is not a standard method for generating a smooth NURBS surface from a discrete data set. Therefore the method described in this section is *ad hoc*. Its particulars are determined largely by the capabilities of and constraints imposed by various pieces of software. The general approach is to smooth the volumetric representation, determine a set of splines that lie on the surface of the body, and then create a three-dimensional form by lofting between the splines.

[4]To appreciate how one can represent a solid object with a NURBS surface, one can imagine covering a statue with small squares of infinitely stretchy material, each of which represents a NURBS surface. In covering the statue, it would not be possible to make each square line up exactly. The squares of material would have to be trimmed to fit. NURBS surfaces are likewise trimmed, so that a single object is represented by several trimmed NURBS surfaces.

4.2.1 Making splines

The first step is to smooth the volumetric representation. The Amira software package (Visage Imaging, Carlsbad, CA) was used for this purpose. It has the capability to create a smoothed triangular surface mesh from a volumetric object.[5] Applied to particular regions of the tongue, the algorithm produced a surface that was satisfactorily smooth, and substantively similar to the volumetric representation. The validation of the smoothing procedure is the researcher's judgment that the resulting triangular mesh resembles closely the original volumetric representation. The volumetric representation of the mucosal surface of the tongue and the corresponding smoothed polygonal surface representation are shown respectively in Figure 4.1a and Figure 4.1b.

The surface generated by Amira consists of thousands of triangles. For the most part, the vertices of the triangles were located on pixel boundaries, i.e., integer coordinates. Consequently, for any integer axial, coronal, or sagittal slice, a set of points was available to define the shape of the region, cut in that plane of section. These points were used as control points for cubic b-splines. The representation of the mucosa as a set of splines on axial planes is shown in Figure 4.1c.

For each region, a plane of section was selected such that the intersection of the region with the plane was a single continuous region (i.e., instead of two separate regions). This was a pragmatic decision, which simplified the lofting process. The planes of section used for each region are given in Table 4.1.

4.2.2 Creating solids

A solid was defined by lofting between the splines to create a surface. A loft creates a surface by interpolating between two curves (in this case, enclosed curves). Lofting between multiple splines is a simple extension of this process. The transition

[5]This was done with the SurfaceGen module, with the smoothing parameter set to "unconstrained smoothing." Amira's algorithm is proprietary, but is widely used in modeling biological structures.

Region	Plane on which splines lay
Genioglossus	Sagittal
Hyoglossus	Axial
Styloglossus	Axial
Transversus/Verticalis	Sagittal
Superior Longitudinalis	Sagittal
Inferior Longitudinalis	Coronal
Mucosa	Sagittal

Table 4.1: Plane of section used to define each region.

from a set of splines to a continuously defined surface is shown in Figure 4.1c and Figure 4.1d.

Lofting created surfaces open at either end. Each end of the surface was enclosed by "capping" the splines on either end of the model. This created an enclosed three-dimensional region. Since the "cap" is flat, each model has smooth artifacts on either end. One such artifact is shown in Figure 4.3. While undesirable, these artifacts are small relative to the size of the model. A solid representation of the tongue mucosa is shown in Figure 4.1d.

Lofting was performed with the proprietary lofting algorithm in the Rhinoceros solid modeling software (McNeel North America, Seattle, WA). This software was selected for its felicity in performing lofts. Lofting produced individual solid models for all regions of interest in the tongue (mucosa, genioglossus, hyoglossus, styloglossus, transversus/verticalis, superior longitudinalis, and inferior longitudinalis).

Not all of the parts were used to define the geometry used in the model. The mucosa, genioglossus, hyoglossus, and styloglossus were merged into a single solid model using Abaqus CAE (Simulia, Providence, RI). The resulting solid model is shown in Figure 4.5 (left).

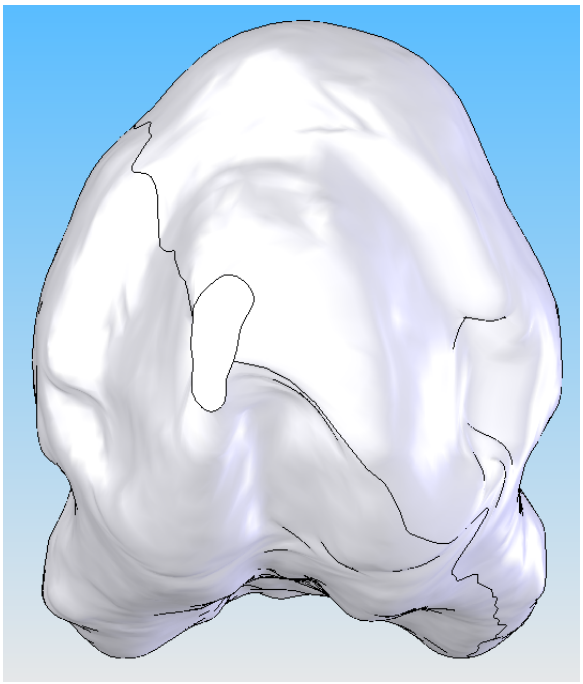

Figure 4.3: Capping artifact. Image generated with SolidWorks (SolidWorks Corporation, Concord, MA).

4.2.3 Sub-lingual structures

The lower boundary of the tongue is not smooth. On the midsagittal plane, the lowest fibers of genioglossus can be considered the tongue boundary. Laterally, however, hyoglossus has portions that are lower than genioglossus. These muscles are shown in Figure 4.4.

Aside from its attachments to bone, the tongue rests on extrinsic laryngeal muscles—the mylohyoid, geniohyoid, and the anterior bellies of the digastric. These muscles are themselves surrounded inferiorly by adipose tissue. To the extent that the base of the tongue protrudes below the its rest position, these structures provide resistance to movement—not a rigid boundary, but a soft barrier—which might influence the posture of the tongue. Such movements are not expected to be very great, however. Including additional anatomy is also computationally expensive.

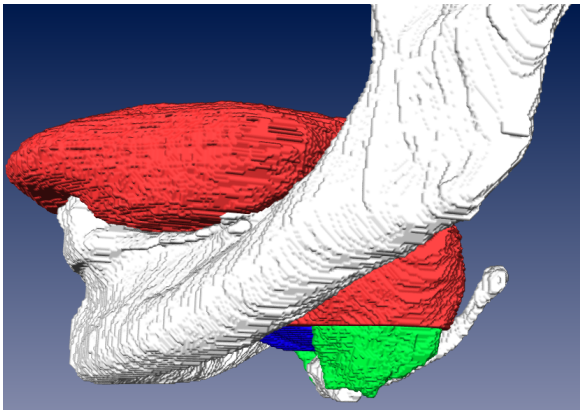

Figure 4.4: Hyoglossus and genioglossus extend beneath what would otherwise be called the base of the tongue. The mandible and hyoid are white. The mucosal surface of the tongue is red. Genioglossus is blue. Hyoglossus is green.

Given these considerations, these sublingual structures were not included in the model.

4.3 Mesh

The finite element mesh determines the geometry of the model. Here the mesh consists of tetrahedral elements. The exterior of the mesh coincides with the exterior of the tongue. Muscle properties, such as the identification of the muscle and the fiber orientation, are assigned at the level of the mesh element.

Meshes were generated using Abaqus CAE (Simulia, Providence, RI). The solid part and a representative mesh is shown in Figure 4.5. It should be noted that when the model was imported into Abaqus, the units of the model were converted to millimeters by uniformly scaling the model by 0.33 (i.e., 0.33 millimeters per Visible Human Project voxel-edges).

Different elements of the mesh were assigned different fiber orientations and contraction controls; these assignments are described in Section 4.6.

A number of meshes were generated in order to determine whether the model results were independent of the particular mesh used. The results of this mesh

validation study are described in Section 4.8.

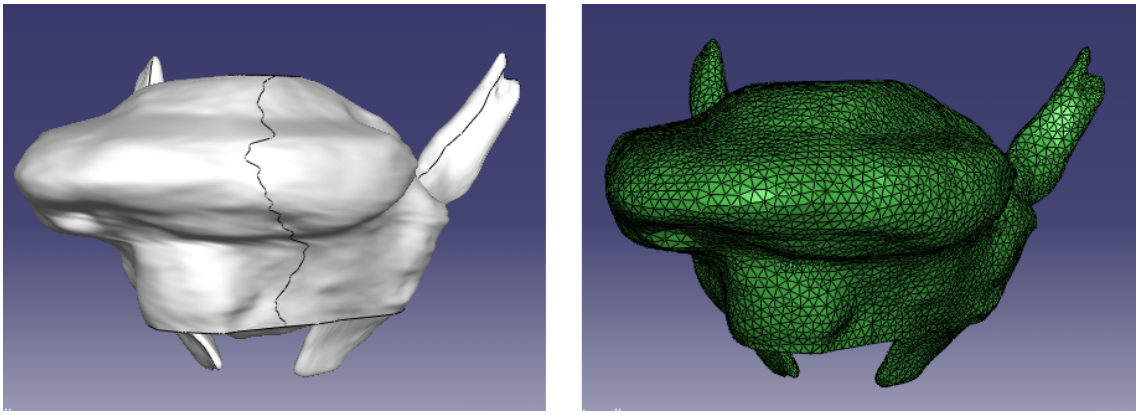

Figure 4.5: Solid part (left) and tetrahedral mesh (right). The mesh consists of 28,304 nodes and 147,635 tetrahedral elements. Image generated with Abaqus CAE (Simulia, Providence, RI).

4.4 Boundary conditions

Neither the jaw, hyoid, nor styloid process of the skull were included in the finite element model. Instead, nodes in positions that corresponded to external attachment sites of the tongue were fixed in position. Since different finite element meshes were used in different simulations (in the mesh validation simulations, for instance), it was necessary to be able to identify nodes based on anatomical criteria. From the anatomical study in Chapter 3, a list of voxels was obtained that corresponded to attachments sites.[6] A node was fixed if it was within 2 mm of one of these voxels.

In all simulations the skull was assumed to be stationary. This is an appropriate assumption because, although the skull is not necessarily immobile in speech productions, no speech sounds depend on skull movement. The skull in the cadaveric position therefore makes a convenient reference frame. Neither the mandible nor the

[6]The attachment sites are: the origination of genioglossus on the posterior aspect of the mandibular symphysis; the insertion of genioglossus onto the hyoid bone; the origination of hyoglossus (bilaterally) from the hyoid bone; the origination of styloglossus (bilaterally) from the styloid process of the skull.

hyoid are actually stationary relative to the skull during speech production, however. A more complete model would include the possibility of specifying the position of the jaw and hyoid. In almost all of the simulations in this dissertation, the attachment sites of the jaw and hyoid bone remained stationary.[7] The reason for this is simply that the model was already of such complexity that sampling its behavior was difficult (see particularly Section 5.1; two more degrees of freedom would minimally require a four-fold increase in the number of samples, and even that would produce only a very clumsy model of hyoid movement). Hyoid movement was included in a small number of simulations described in Section 6.2.3. The methods for specifying hyoid movement are described in that section, since they are specific to the focus of that section.

4.5 Material description

This section describes the material model used in the finite element model, which is considerably simplified from the anatomical reality. Although the tongue is composed of mucosal, muscular, tendinous, adipose, nervous, and vascular tissues—and although the composition of the tongue varies by region with respect to connective tissue composition (Miller et al., 2002)—in the present study the entire tongue model was given identical passive properties, approximately those of skeletal muscle. That is, the tongue is assumed to be homogenous. Changes in the material properties that would result from changes in temperature are not modeled, the assumption of isochoricity. Changes in the material properties of the tongue as a result of the flow of fluids in or out of the tongue (such as blood) were not modeled. These simplifications are commonly made in modeling muscle (e.g., Blemker, 2004). Skeletal muscle was approximated with a transversely-isotropic hyperelastic consitutive model. The only isotropy in the model, however, was in active muscle contraction. The passive properties where assumed to be isotopic, i.e., identical no matter the direction of

[7] The exceptions to this are simulations reported in Section 6.2.3, in which the hyoid is moved relative to the jaw (and styloid process), and those in Section 6.1.1, in which the jaw and hyoid are moved relative to the styloid process.

the applied load. The work done on a hyperelastic material is dependent only on the undeformed state of the material and its deformed state; this is discussed further in Section 4.5.3. Note that this consitutive relation ignores the viscoelastic properties of muscle.

Following a brief introduction to the notation used, the description of the material model used is provided.

4.5.1 Introduction to notation

This section introduces the notation used for continuum mechanics, generally following that of Bonet and Wood (1997). The set of points $\vec{X}_1, \vec{X}_2, \ldots$ defines the shape of the object in its undeformed state. After the body is deformed these material points become $\vec{x}_1, \vec{x}_2, \ldots$. The coordinates of the deformed material are described by the function $\Phi(\vec{X}_i)$, which maps material points from their undeformed position to their deformed position:

$$\vec{x}_i = \Phi(\vec{X}_i)$$

Φ is a function of $\vec{X}_i$ since the coordinate of the material point in the deformed configuration depends on the location of the point in the undeformed configuration. One could imagine deformations in which all material points are affected uniformly (such as a scaling deformation); generally, however, different points of a body deform in different ways. For this reason it is helpful to consider the second-order tensor F, the derivative of Φ with respect to the original spatial coordinates.

$$F = \frac{\partial \phi}{\partial \vec{X}}$$

It happens that the determinant of F represents the proportional change in volume, J.

$$J = \det F$$

Another important tensor is C, the right Cauchy-Green deformation tensor, which is calculated from F. This tensor is used below in defining the strain energy function of the material.

$$C = F^T F$$

4.5.2 Isotropy

In modeling of soft tissues, and particularly in muscle modeling, an important distinction is to be made between isotropic and anisotropic materials. Isotropic materials exhibit the same stress/strain relationship no matter the direction of loading.[8] For an anisotropic material, the stress/strain response can be different for different loading directions. The simplest kind of anisotropic material is a transversely isotropic material (Weiss et al., 1996; Holzapfel and Gasser, 2000; Blemker et al., 2005). A description of such a material includes a vector, $\vec{a}$, that indicates the characteristic orientation of the material. In the plane perpendicular to $\vec{a}$, the material is isotropic (hence *transverse* isotropy). The material can have a different stress-strain relationship, however, for loads applied along $\vec{a}$.[9]

Muscle is a transversely isotropic material (Fung, 1993). Along with a ground substance (the non-contractile parts of the muscle), it has a characteristic direction—the orientation of the contractile fibers—along which its material properties are different. In particular, it is along this direction that the muscle generates stress during contraction. For contractile regions, following van Leeuwen (1992), the Cauchy stress in an element is defined to be the sum of the passive stress and active (i.e., contractile) stress:

$$\sigma = \sigma_{\text{active}} + \sigma_{\text{passive}} \tag{4.1}$$

The active stress is determined by the magnitude of the contraction, T_a, and the orientation of the muscle fiber, $\vec{a}$.

$$\sigma_{\text{active}} = T_a \cdot \vec{a} \otimes \vec{a} \tag{4.2}$$

The stress generated by the muscle depends on a user-specified time-dependent

[8] Examples of approximately isotropic materials are gelatin, pure rubber, or steel.

[9] Wood could be considered approximately anisotropic, since its deformation differs depending on whether the load is applied along or against the grain. Well-cooked beef or pork is transversely anisotropic: it is hard to pull apart along the fiber direction, but easy to pull apart perpendicular to that direction. (Note however that, in this work, the passive properties of muscle are assumed to be isotropic.)

activation curve $C(t)$. The actual stress generated by the muscle is given length-dependence with an elastance model (Guccione et al., 1993).

$$T_a = \frac{Ca_0^2}{Ca_0^2 + ECa_{50}^2} C(t) \tag{4.3}$$

Where Ca_0^2 is a model parameter, and ECa_{50} is given by the following:

$$ECa_{50} = \frac{(Ca_0)_{\max}}{\sqrt{\exp[B(l - l_0)] - 1}} \tag{4.4}$$

$(Ca_0)_{\max}$ and B are model parameters that determine the length-dependence. l_0 is the sarcomere length at which no stress is generated (a model parameter). $l = l_r \cdot \lambda$, where l_r is the initial sarcomere length, and λ is the stretch of the material in the direction of the muscle fiber. The parameter values, which are given in Table 4.2, were identical to those used by Guccione et al. (1993). Guccione et al. take the parameter values from sources cited in Guccione and McCulloch (1993), all of which are studies of cardiac muscle. There may be some inconsistency in the dependence of stress on muscle length between cardiac and skeletal muscle.

The muscle activation curve $C(t)$ was specified for each muscle (and occasionally for a subvolume of a muscle). A muscle fiber orientation was defined for each individual element, as discussed in Section 4.6.

There are not published data about the *passive* transverse isotropic properties of human muscle. Therefore the passive material behavior was modeled with an isotropic material model. Many researchers who model muscle, and all who model the tongue, have made this assumption (Wilhelms-Tricarico, 1995; Meier and Blickhan, 2000; Johansson et al., 2000; Dang and Honda, 2001; Gerárd et al., 2005; Felton et al., 2007).

4.5.3 Hyperelasticity

Muscle can be further specified as a transversely isotropic *hyperelastic* material (Fung, 1993). A hyperelastic material is one for which the work done by stress depends only on the initial configuration of the material and its current state. For such a material, one can define a strain-energy function W, which is written in terms

of the right Cauchy-Green deformation tensor C. Following Bonet and Wood (1997), the second Piola-Kirchoff stress tensor is then:

$$S_{ij} = 2\frac{\partial W}{\partial C_{ij}} \tag{4.5}$$

The second Piola-Kirchoff stress tensor is a material (or Lagrangian) tensor. For the solution procedure the Cauchy stress tensor (a spatial or Eulerian tensor) is required (cf. Section 4.7). The Cauchy stress tensor can be calculated using the second Piola-Kirchoff stress tensor. The derivation and result are not provided here, however (see Bonet and Wood, 1997, 122–123).

In modeling incompressible and nearly-incompressible materials, a common strategy is to separate the volumetric and deviatoric components of deformation. This resolves some computation difficulties in modeling incompressible materials. The volumetric component of deformation is that which results in a change in volume, while the deviatoric is component of deformation that does not change the volume. The deviatoric components of the tensors defined above are:

$$\tilde{F} = J^{-1/3}F \tag{4.6}$$

$$\tilde{C} = \tilde{F}^T\tilde{F} \tag{4.7}$$

Then the second Piola-Kirchoff stress tensor becomes

$$S_{ij} = 2\frac{\partial \tilde{W}}{\partial C_{ij}} + U(J) \tag{4.8}$$

where $U(J)$ is the strain energy associated with changes in volume, which is discussed below.

The only anisotropic material behavior was in contraction; the passive properties were purely isotropic. Maas and Weiss (2008) provide the following form for the strain energy function of this material:

$$W = F_1(\tilde{I}_1, \tilde{I}_2) + F_2(\tilde{\lambda}) + \frac{k}{2}[ln(J)]^2 \tag{4.9}$$

Here F_1 is the material response of the ground substance. Note that F_1 is a function of the scalar invariants of $\tilde{C}$. The invariants of a tensor are properties that do not

depend on the particular set of axes chosen as a basis. Hence, the contribution of F_1 to the strain-energy function does not take the fiber direction of the material into account: it is isotropic. The definitions of I_1 and I_2 are:

$$\tilde{I}_1 = \sum_{i=1}^{3} \tilde{C}_{ii} = \tilde{C}_{11} + \tilde{C}_{22} + \tilde{C}_{33} \tag{4.10}$$

$$\tilde{I}_2 = \sum_{i=1}^{3} \sum_{j=1}^{3} \tilde{C}_{ij} \tilde{C}_{ij} \tag{4.11}$$

The passive isotropic properties of muscle were approximated with a two-term Mooney-Rivlin model.

$$F_1 = C_1(\tilde{I}_1 - 3) + C_2(\tilde{I}_2 - 3) \tag{4.12}$$

C_1 and C_2 are empirically derived scalar quantities. C_1 is the shear modulus.

F_2 is a function of the stretch λ in the material's direction, i.e., the anisotropic behavior of the material. Since the passive properties of the material were isotropic, the material parameters were set so that the contribution of F_2 would be zero.[10]

Finally, $\frac{k}{2}[ln(J)]^2$ has to do with enforcing near-incompressibility of the material. This term penalizes changes in volume by requiring more energy for changes in volume. Recall that J is the volume change of the material. Plotting $\frac{k}{2}[ln(J)]^2$ against J with an arbitrary value of k illustrates that the term is zero for $J = 1$ (no change in volume), but becomes very large for values of J less than 1 (indicating compression of the material). The severity of the penalty depends on the value of k, which is the bulk modulus.

Taking into account that F_2 was set to zero, the final strain energy function for the material is:

$$W = C_1(\tilde{I}_1 - 3) + C_2(\tilde{I}_2 - 3) + \frac{k}{2}[ln(J)]^2 \tag{4.13}$$

Values for C_1 and C_2 were taken from Teran et al. (2005), as 0.03 MPa and 0.01 MPa, respectively. Teran et al. used a value 0.06 MPa for the bulk modulus (k), which is not large enough to approximate incompressibility. Based on the guideline from Fung (1993) and Bonet and Wood (1997) that k must be three or four orders of

[10]This was done in FEBio, the software package used, by setting C_3, C_4, and C_5 to 0.

Parameter	Setting
C_1	0.03 MPa
C_2	0.01 MPa
C_3	0 MPa
C_4	0 MPa
C_5	0 MPa
k	30 MPa
T_{max}	-0.35 MPa
$(Ca_0)_{\max}$	4.35 μM
B	4.75 μm^{-1}
l_0	1.58 μm
l_r	2.04 μm

Table 4.2: Material parameters

magnitude greater than the sheer modulus (C_1) for near-incompressibility, k was set to 30 MPa. To complete the description of the material parameters, the maximum stress generated by muscle was 0.350 MPa (Zajac, 1989, 374). The full set of parameters used in the model is given in Table 4.2.

4.6 Assigning mesh properties

As discussed above (Section 4.5), the entire tongue model was given identical passive hyperelastic material properties. Regions of the tongue differed, however, in the orientation of the muscle fibers. Contraction magnitude was assigned variously under different control schemata, though generally at the level of the whole muscle.

Each element of the mesh had a different muscle fiber orientation, which was determined based on the anatomical study reported in Chapter 3. The volumetric data from Chapter 3 were condensed into a single grid, in which each region of the tongue was represented by a single number. Each voxel was assigned a value $0 \ldots 8$, with the respective interpretations of: empty space, mucosa, genioglossus, hyoglossus, styloglossus, transversus/verticalis, superior longitudinalis, inferior longitudinalis, and the region where transversus/verticalis and genioglossus overlap.

Fiber orientation was assigned to each element based on the element's center, the mean of its four nodes.[11] In general, each element was assigned a fiber orientation by the following procedure. The first step was to determine the region of the tongue to which each element belonged. This was accomplished by calculating the vowel-valued coordinate of the element center, i.e., converting back to voxels by multiplying by 1/0.33 voxel/mm and rounding. Then the aforementioned data structure was used as a lookup table, to determine the material associated with each element. Once the element had been determined to belong to a particular material, the fiber orientation was calculated using the control vectors and vector interpolation algorithm described in Section 3.2.3. (No fiber orientation was defined when the center of the element was identified as mucosa.)

It happened occasionally that the center of an element was not assigned a label, i.e., it corresponded to a voxel of value 0, an impossible situation. This is because the location of the element was based on a solid-part representation of the tongue, which had been derived from the volumetric data set, and various smoothings had occurred in the conversion process. To fix this case, the 26 voxels surrounding the original voxel were examined (i.e., one voxel in each direction). The element was then assigned the mode material for these voxels.

There were two exceptions to the above procedure, both having to do with cases where the volumetric representation was insufficient to represent the full anatomical detail of the tongue. This was the case for the interdigitation of transversus and verticalis, and the interdigitation of verticalis and genioglossus. Recall that transversus and verticalis were identified as a single muscular region, in spite of the fact that they are distinct muscles, with alternating lamina of transversely- and vertically-oriented fibers. Recall also that genioglossus and verticalis interdigitate where they intersect (near the midline).

A laminar structure was recreated for transversus and verticalis. First, the cen-

[11] In geometry and computer graphics, one speaks of "vertices" of a tetrahedron, rather than nodes, which is the preferred term in the finite element method. The term "node" is used throughout the dissertation, however, for the sake of consistency.

ter of the transversus/verticalis region along the left-right axis was determined (the arithmetic mean of the left-right coordinates of the voxels with the transversus/verticalis label). The center slice was the 173rd (or 0-indexed, the 172nd). The corresponding sagittal slice of the transversus/verticalis label was extracted. Using this slice, a centerline for the region was defined by manually identifying control points to define a cubic b-spline (Figure 4.6, red circles). Next, N equidistant points, including the endpoints, were calculated along the length of the spline. At these points, lines perpendicular to the spline were calculated. These lines were extruded along the x-axis to define planes that intersected the transversus/verticalis region. These planes defined intervals corresponding to alternating lamina of transversus and verticalis fibers. A transversus/verticalis element was assigned to transversus if it fell in an even interval, and to verticalis if it fell in an odd interval.

Takemoto (2001) estimated that there were 100 lamina of transversus fibers and 100 lamina of verticalis fibers, or 200 lamina in total. Although ideally this figure should be represented in the tongue model, computationally practical finite element meshes do not permit such fine detail to be specified. As a compromise, various values of N were tested: 20, 100, and 200 (total lamina). At practical mesh densities (see Section 4.8), it was possible to specify only 100 lamina (50 transversus, 50 verticalis). Figure 4.7 shows a sagittal view of the partition of the transversus/verticalis label into transversus and verticalis, where each point is the center of an element (rounded to the nearest integer coordinate).

In places where verticalis overlapped with genioglossus, it was necessary to represent the interdigitation of the muscle fibers. This interdigitation was represented similarly to what was described for transversus and verticalis fibers. Starting from a "top-down" view of the genioglossus region,[12] two lines were defined as boundaries. A total of M lines were defined by interpolating between these two. The lines were extruded along the long axis in order to define planes, which in turn defined intervals of genioglossus and verticalis fibers. If an element lay in an odd interval it was

[12]Formally, this was the boolean union (in the image processing sense) of all axial slices of the genioglossus region.

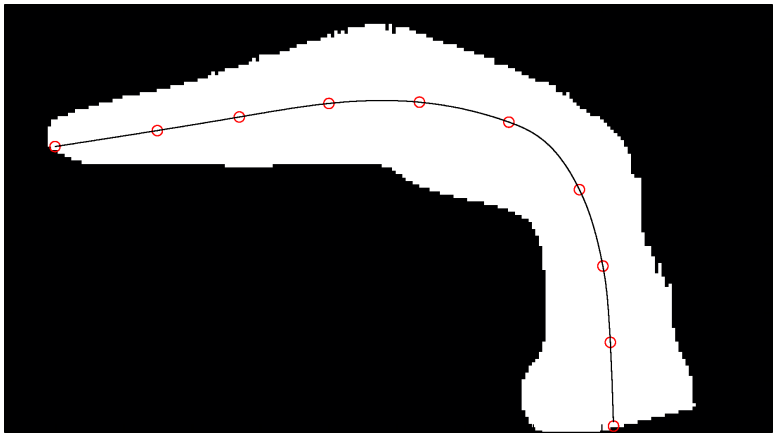

Figure 4.6: The manually defined centerline of the transversus/verticalis region. It is shown over the center (173rd) sagittal slice of the region.

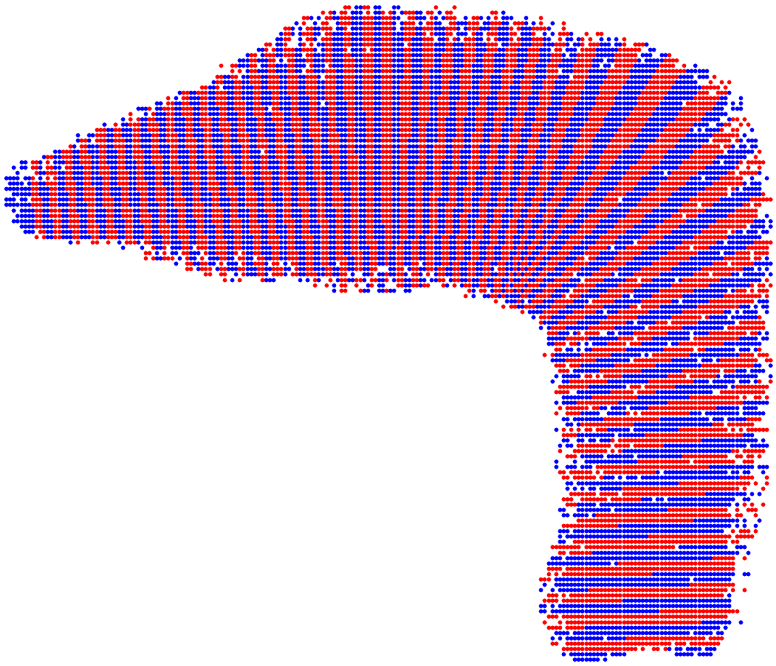

Figure 4.7: Alternating lamina of transversus (blue) and verticalis (red) elements.

assigned to verticalis; it if lay in an even interval it was assigned to genioglossus. No anatomical study has provided an estimate for M. A value of 20 was found to be reasonable given computationally practical meshes, though the true number is doubtless much higher.

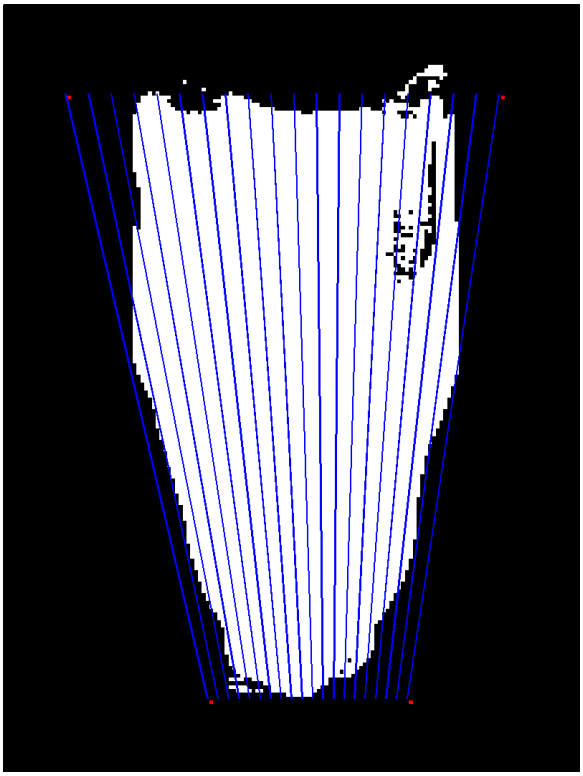

Figure 4.8: Division of the genioglossus region into alternating genioglossus and verticalis fibers.

In summary, an element was assigned to a specific muscle based on the volumetric anatomical data. Typically this was a straightforward assignment, with each voxel of the volumetric data corresponding to a single muscle. The situation was more complicated for the transversus/verticalis region, and particularly along the midline. The transversus/verticalis region was divided into 100 distinct lamina (50 verticalis, 50 transversus). Anatomical criteria were used to define these lamina, but the lamina themselves did not correspond to the lamina of the anatomical specimen, since these could not be individually identified. In the place where genioglossus and the verticalis labels overlap, elements were assigned to one muscle or the other based on another

such division.

4.7 Solution procedure

This section describes the procedure used to find the tongue deformation that corresponded to a particular muscular activation schema.[13] Muscle activation is specified by the user in pressure units (MPa, here 0.35 MPa was typical). This stress was added to each element that was associated with the muscle. The task is then to find the equilibrium position of the nodes of the finite element mesh, given the stresses imposed by muscular contraction.

For a deformable body with area a and volume v, equilibrium implies that the sum of the traction forces $\vec{t}$ per unit area and the body forces $\vec{f}$ per unit volume is zero.[14]

$$\int_a \vec{t}\, da + \int_v \vec{f}\, dv = 0 \tag{4.14}$$

It can be shown (Bonet and Wood, 1997, 102) that this is condition is equivalent to the following:

$$\operatorname{div} \sigma + \vec{f} = \vec{0} \tag{4.15}$$

Since a computational solution involves *minimizing* the force residual rather than actually reducing it to zero, the above equation can be written to allow for a non-zero residual, $\vec{r}$.

$$\operatorname{div} \sigma + \vec{f} = \vec{r} \tag{4.16}$$

In the finite element method the displacements of the nodes at equilibrium are typically solved using the principle of virtual work (or power). This states that, at equilibrium, the total work done by an imaginary (virtual) displacement δv consistent with the constraints on the system is zero. The virtual work per unit volume, δw, done by a residual force $\vec{r}$ is $\vec{r} \cdot \delta\vec{v}$. If the body is in equilibrium the work done by the residual force is zero, so

$$\delta w = \vec{r} \cdot \delta v = 0 \tag{4.17}$$

[13] The discussion in this section follows Bonet and Wood (1997) and Maas and Weiss (2008).

[14] Note that the effect of inertia is not included in this formula.

Summing over the volume of the body, the equilibrium condition becomes

$$\delta W = \int_v (\operatorname{div} \sigma + \vec{f}) \cdot \delta \vec{v} \, dv = 0 \tag{4.18}$$

For the present model, in which an active material is used, the Cauchy stress is defined as the sum of the active and passive stress of the material. Substituting from (4.1),[15]

$$\delta W = \int_v (\operatorname{div}(\sigma_{\text{active}} + \sigma_{\text{passive}}) + \vec{f}) \cdot \delta \vec{v} \, dv = 0 \tag{4.19}$$

Since body forces such as gravity were not included in the model, the solution procedure is essentially to balance the active stress with the passive stress. The user specifies σ_{active}, and σ_{passive} depends on the displacement of the nodes and the consitutive model (Section 4.5.3). Solution consists in finding the nodal deformations to balance the active stress. Since the problem is non-linear, Newton's method is used.

Newton's method is applicable when one has an equation and its derivative. Figure 4.9 illustrates Newton's method graphically for finding the zeros of $y = f(x)$. That is, the solution is an x-coordinate such that $f(x)$ is zero, or satisfactorily close to zero. Both $f(x)$ and $\partial y/\partial x$ can be calculated. The first step is to make a guess about the solution, x_1. The value of y is calculated at this x-coordinate (y_1). The slope of the tangent of y at x_1 is calculated (m_1). Then the intersection of the tangent line with the zero axis is calculated; this x-intercept is x_2, the second guess. By following the slope of the tangent, a new x-coordinate is determined that is closer to the zero of the function. The procedure is iterated until the y-coordinate for some x_n is either zero or satisfactorily close to zero. In Figure 4.9, this occurs at x_4.

In finding a solution to a finite element problem, the stresses generated by muscle contraction are known, while the mesh deformation that produces the appropriate internal stresses are not known. The solution begins with a guess of the deforma-

[15]Here σ_{active} is not defined. Active stress is specified in the model on an element-by-element basis. (Rather, active stress is specified for a material, which is in turn specified on an element-by-element basis, cf. Section 4.6.) Properly then, σ_{active}, should be added to the discretized equilibrium equations, which are not given here.

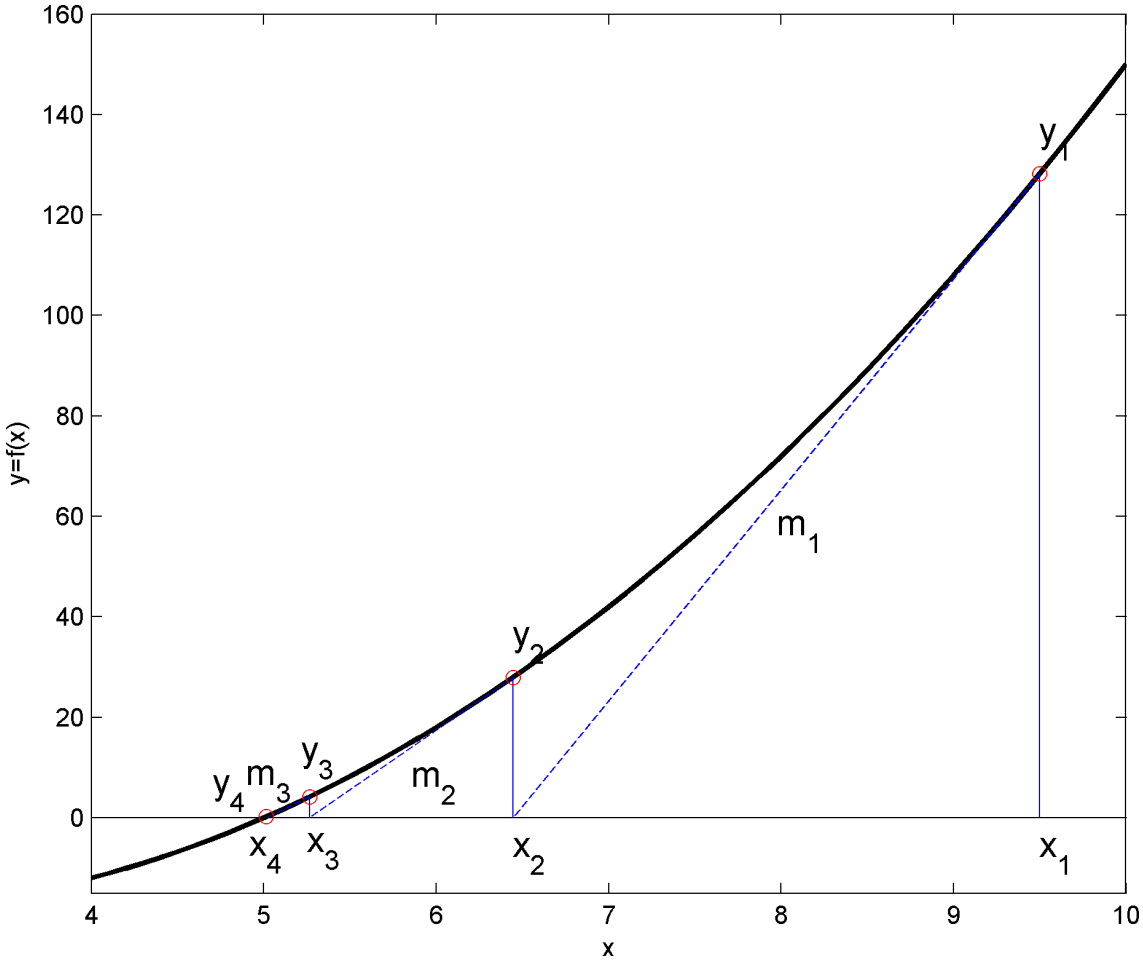

Figure 4.9: Illustration of Newton's method for finding the zero of a non-linear equation.

tion that will produce the appropriate internal stresses. Then the force residual R_f calculated, and the estimate is revised iteratively.

4.8 Mesh validation

The purpose of a finite element model is to approximate a continuous structure with a discrete representation. In the present work a continuous tongue model is approximated with discrete tetrahedra. One may imagine that approximating the tongue with a very small number of tetrahedra would introduce error into the simulation, since the geometric representation would be quite coarse. Conversely, using an infinite number of tetrahedra would admit no geometric error, but would be computationally intractable. The purpose of a mesh validation study is to determine a point at which increasing the precision of a tetrahedral mesh fails to produce substantial increases in accuracy.

To validate the mesh, a series of tetrahedral meshes were generated using Abaqus, with varying numbers of tetrahedra in each mesh. Node and element counts for the meshes are given in Table 4.3.[16]

Since, for a particular mesh, each muscle was represented by a set of tetrahedra (Section 4.6), it is necessary to check whether a mesh was sufficiently fine to represent the muscle groups. To assess this, the volume of superior longitudinalis was computed for each mesh. Figure 4.10 shows the results. It is evident that the volume of superior longitudinalis does not change substantially after the 46,490-node mesh (having 245,223 elements).

For applications in speech science, the displacement of surface nodes is of primary interest. Validation of node displacement involves tracking the deformation of a single point during deformation, for meshes of varying density. For this purpose, the displacement of the front-most node of each mesh—the tongue tip—was

[16]For the first four meshes in Table 4.3, Abaqus was unable to mesh the portions of genioglossus and hyoglossus that protruded beneath the base of the tongue; consequently these regions were not included in the meshes. Since the validation study simulations involved only contraction of superior longitudinalis, this is not thought to have made much of a difference, if it made any.

Number of Nodes	**Number of elements**
988	3,952
1,243	5,107
1,673	7,086
2,628	11,764
14,830	72,766
15,788	75,752
24,237	123,484
29,072	149,188
46,490	245,223
106,449	579,383
181,758	995,351
211,814	1,162,878

Table 4.3: Properties of generated meshes

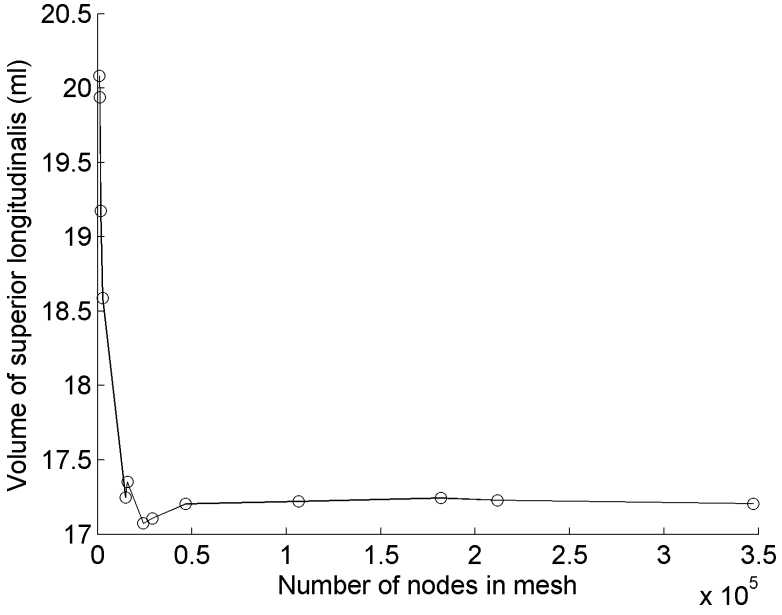

Figure 4.10: Volume of superior longitudinalis as a function of mesh density.

measured during full contraction (0.35 MPa) and half contraction (0.175 MPa) of superior longitudinalis. Contraction of superior longitudinalis produces a fairly large displacement of the tongue tip (see Figure 6.6). Because the meshes had different numbers of nodes, and therefore different nodal coordinates, the front-most node of each mesh was in a slightly different position. Table 4.4 gives the coordinates for the front-most nodes of each mesh.

The results of the displacement study are given in Figure 4.11.[17] Aside from minor fluctuations, displacement of the front-most node does not change substantially, in either contraction scheme, for meshes finer than the 46,490-node mesh. Starting from the 46,490-node mesh, the x and y displacements are within 0.029 mm, and the z displacement is within 0.79 mm, of the finest mesh, for the full contraction. For the half contraction, the x and y displacements are within 0.18 mm, and the z displacement within 0.95 mm.

Number of Nodes	**Tip X**	**Tip Y**	**Tip Z**
988	56.3476	102.27	26.9194
1243	57.8806	102.52	28.2275
1,673	57.3365	102.409	27.3817
2,628	56.2146	102.526	27.8359
14,830	58.0992	102.41	27.6394
15,788	57.1848	102.561	28.6808
24,237	56.7764	102.564	28.0357
29,072	57.1848	102.555	28.0583
46,490	56.867	102.582	28.6478
106,449	56.5081	102.6	28.3759
181,758	56.6155	102.583	28.162
211,814	56.3621	102.602	28.4927
347,298	56.0807	102.596	28.636

Table 4.4: Coordinate of the front-most node

A final validation was of stress. The stress was measured in a small region near the tongue tip. The average Von Mises stress was calculated. Stress is a tensor

[17]Note that there is displacement of about 2 mm on the horizontal (x) axis. This would typically not be expected from a symmetric muscle contraction, but since the data are based on an actual anatomical specimen, the structures modeled are not quite symmetrical.

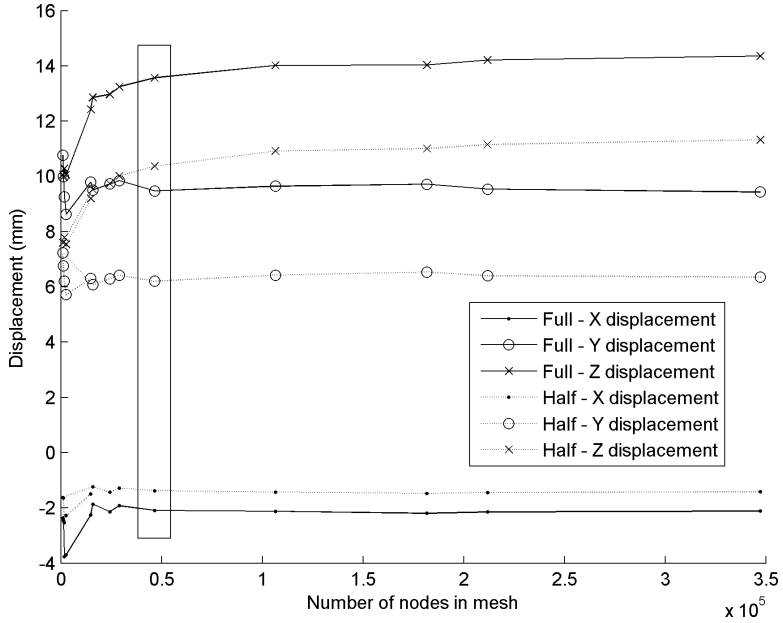

Figure 4.11: Displacement of the frontmost node with full (0.350 MPa) and half (0.175 MPa) contraction of superior longitudinalis, as a function of mesh density.

quantity, but the Von Mises stress (σ_{vm}) is a scalar quantity.

$$\sigma_{vm} = \sqrt{\frac{(\sigma_x - \sigma_y)^2 + (\sigma_x - \sigma_z)^2 + (\sigma_y - \sigma_z)^2}{2}} \tag{4.20}$$

Von Mises stress can be calculated for each element. For convergence testing, the average σ_{vm} for the elements coming from a small region near the tongue tip was computed. A geometric region was defined (as opposed, for example, to calculating σ_{vm} at the frontmost element) so that the stress would come from a comparably-sized region in all simulations. The size of elements varied greatly, of course, among the validation meshes. An element was included in the average if it was anterior of 100.98 mm from the origin, superior to 28.38 mm from the origin, and between 54.45 mm and 59.4 mm on the left to right axis. Figure 4.12 illustrates the region. The convergence of stress with increasing mesh density is shown in Figure 4.13. Stress does not appear to converge as quickly as did displacement. It might be only at the greatest mesh densities that stress converges. There is probably also error

associated with the measurement technique, however, and this may call into question the validity of the stress convergence test. In spite of the consistent anatomical criterion for identifying the test region, the region was actually different for each mesh, since different tetrahedra fit within the region for each mesh. Figure 4.14 shows the volume of the small region as a function of mesh density, demonstrating that the volume of the region does not converge except for the three or four finest meshes.

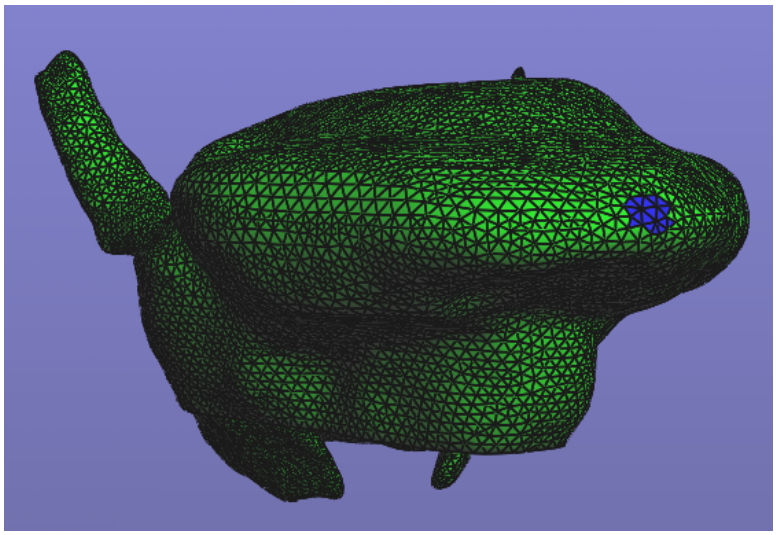

Figure 4.12: Illustration of the small region near the tongue tip used for stress validation. (The 46,490-node mesh is shown).

Since the variable of primary interest in this dissertation was nodal displacement, the 46,490-node mesh was selected for further study. This mesh produced stress values reasonably close to those of the finest mesh, and most importantly, produced nodal displacement values that did not differ greatly from the finest mesh.

4.9 Summary and review of model assumptions

This section has described the creation of the finite element model and the validation of the finite element mesh. The geometry of the model was determined from the anatomical study reported in Chapter 3. The volumetric data were converted into a trimmed NURBS surface representation, which served as input for the mesh

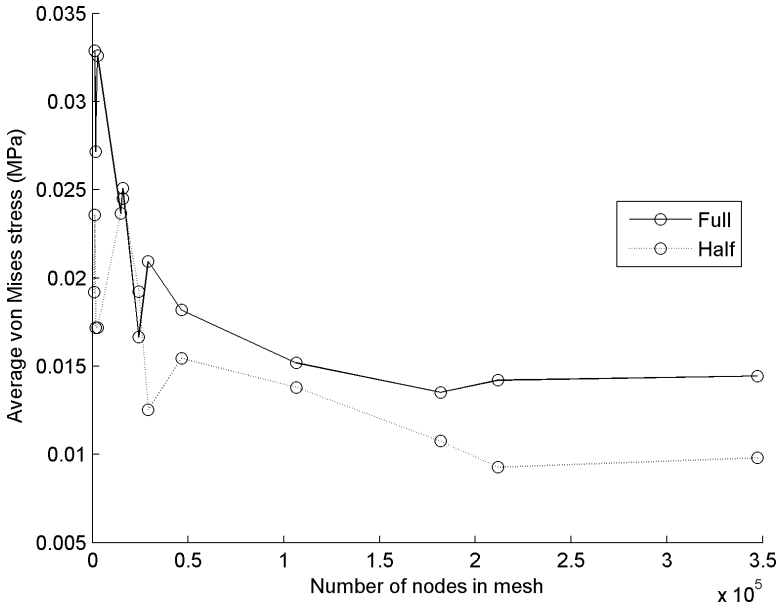

Figure 4.13: Average Von Mises stress in a small region near the tongue tip with full (0.350 MPa) and half (0.325 MPa) contraction of superior longitudinalis, as a function of mesh density.

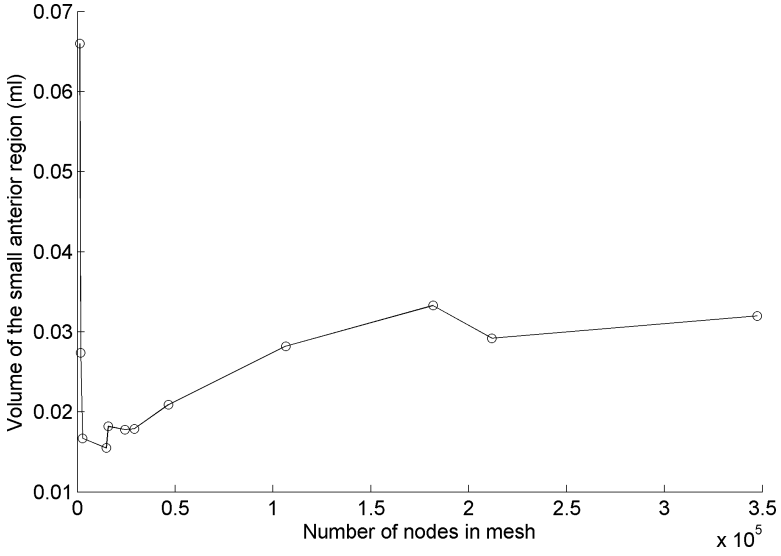

Figure 4.14: Volume of the small region near the tongue tip, as a function of mesh density.

generator. Muscles were represented in the model by assigning elements of the mesh to particular muscles, which determined both fiber orientation and the control of the muscle. The model was validated by comparing the results from a number of meshes of varying density.

The validity and generality of the results derived from the finite element model rely upon the assumptions made in the model, which are reviewed below.

The geometry of the model was based on the anatomical study described in Chapter 3. This study was based upon volumetric data from a frozen cadaver. Consequently, the undeformed (rest) position in the present model is the position of the tongue in that cadaver, the tacit assumption being that this represents the tongue shape that results from the absence of muscle contraction. This is not necessarily a significant posture for studying, e.g., speech or mastication, which have their own "rest" postures. In addition, physiological processes associated with dying (e.g., rigor mortis) may have altered the "true" rest geometry of the tongue. It is also possible that freezing the cadaver distorted the shape of the tongue in significant ways.

The entire tongue was modeled as a homogenous material. This is counterfactual, since the tongue is composed not only of muscle, but of mucosa, adipose, glands, and nervous, vascular, and connective tissue. Although the assumption of homogeneity is commonly made in modeling muscle, this assumption affects the results of the model, since different kinds of tissue have different rheological properties.

The material of the tongue was modeled with a transversely-isotropic hyperelastic constitutive model. This consitutive model assumes that the work done on the material can be expressed by a strain-energy function, which depends *only* on the undeformed state of the material and the current state of deformation. Therefore, the viscoelastic properties of muscle cannot be captured with a hyperelastic model. The passive material behavior was further assumed to be fully isotropic; the only anisotropy in the model resulted from the generation of active stress in the material (i.e., muscle contraction). In reality one would expect some anisotropy in the passive properties of muscle, given the fibrous structure of muscle. A two-term Mooney-Rivlin model was used to model these passive properties. The model param-

eters were taken from previously published data for skeletal muscle. It is possible, however, that tongue muscle has distinct rheological properties from skeletal muscle as a whole.

The constitutive model did not include effects of temperature, or of the flow of fluids in or out of the tongue.

Active stress in the muscle is specified by the user. It has been assumed that tongue muscle generates 0.35 MPa of active stress, a typical value from the literature. This value is not drawn from experimental data from the tongue, however, which may turn out to have significantly different active stress generation properties. The stress generated by the muscle was given length-dependence with an elastance model developed for cardiac muscle. Discrepancy between the relationship between muscle length and active stress generated by skeletal and cardiac muscle would introduce error here. The elastance model also did not take into account the rate of change of the length of the muscle, which is known to affect the magnitude of stress generated (Hill, 1938).

CHAPTER 5

Taming the tongue: sampling, dimension reduction, and prediction

> Through patience a ruler can be persuaded, and a gentle tongue can break a bone.
>
> —Proverbs 25:15 (NIV)

This chapter deals with three interrelated topics: sampling, dimension reduction, and prediction. These topics bear on two quite general questions about the tongue.

- What tongue postures can the tongue model produce?[1]
- What is the muscular activation pattern associated with a given posture? Or conversely, what posture is associated with a given activation?

A possible answer to the first question would be a very long list of the tongue postures associated with each possible muscle activation pattern.[2] The problem with such an answer is that it would be very difficult to form generalizations from such a large data set. One could consider variation along of the dimensions of the data set, but there are thousands of parameters, since the shape of the tongue is defined by thousands of nodal coordinates. None of these coordinates is really interesting in itself, which would make the interpretation of the data more difficult.

[1] The tongue model has numerous outputs of potential interest: stress, strain, displacement of nodes or elements, and measures of interest derived from these quantities. The variables that a research attends to will be determined by his/her research focus. For speech production, the primary focus of this dissertation, the focus is the shape of the air in the vocal tract. Insofar as the tongue contributes to this shape, the critical variables are the displacements of the nodes on the surface of the tongue.

[2] The list would not be infinite, since matter is discrete and the tongue has a finite number of atoms. But it would be a very, very long list.

Making a data set comprehensible to humans must in this case involve dimension reduction. In the present work, this is done with principal component analysis. The exterior of the tongue is defined by 11,331 exterior nodes in the tongue model, and since each node is located in three-dimensional space, the posture of the model is described by 33,993 numbers. That is, each tongue posture can be considered a point in 33,993-dimensional space, in which the basis vectors for the space have ready interpretations as nodal coordinates. For a given set of tongue postures, most of the dimensions will be very highly coordinated with one another, since the positions of the nodes are not actually independent—there is a mechanical linkage between the nodes. A principal component analysis finds a new basis, in which correlation between the dimensions is minimized. In fact, a few basis vectors can be identified that account for most of the variance in the data set; these basis vectors can be retained, and the others discarded, at the researcher's discretion.

The second question—of what muscle activation pattern is associated with a given posture—can be addressed with a statistical model. Here both artificial neural networks and simple linear regression models are used. A statistical model can also be used to produce a tongue shape from an activation pattern, which is advantageous since calculating the solution from a linear model is much more efficient than calculating the solution to a finite element problem. The statistical model is a more efficient alternative to the finite element model.

In addressing both of these questions the researcher is confronted by a difficult problem, that of sampling the space of possible tongue postures. Both a principal component analysis and a statistical model must be based on a non-exhaustive sample of the data. Since obtaining a sample is computationally expensive, it is best to collect the smallest sample that permits generalization. For a structure as complex as the tongue, however, it is not at all clear how large a sample is necessary. In this work, this is assessed by determining the adequacy of a 127-sample data set for predicting values from a 2,059-sample data set. If the smaller data set generalizes well to the larger, there are grounds to make generalizations to even larger data sets, which it would not be practical to validate independently.

These issues of sampling, dimension reduction, and regression are interesting from at least two different perspectives. A speech scientist may have his goal a complete model of the vocal tract. He may then desire an efficient way to represent the posture of the tongue, and an efficient way to compute the tongue postures for a given muscle activation, or muscle activation from a certain posture (depending on the research project). The issues are also of interest to a developmental physiologist. The central nervous system must have some reasonably efficient representation of tongue posture, and of what posture will result from a given activation pattern (or, what activation pattern is required for a certain posture, depending on one's theory of motor planning.) Whether these problems have tractable solutions is of interest for understanding how the central nervous system controls the tongue.[3]

The results of the studies presented in this chapter are, in short, all good news. It is not necessary to represent tongue posture with thousands of dimensions. Upwards of 98% of the variance in a given data set can be described with just seven parameters. Additionally, these parameters can be reliably calculated from a relatively small sample. The data from a 2,059-member set can be quite adequately represented in a vector space derived from a 127-member set (Section 5.2.5). Furthermore, using fairly simple statistical models it is possible to make quite good predictions about tongue posture from muscle activation levels (Section 5.3.1) and about muscle activation levels from tongue posture (Section 5.3.2).

The chapter is divided into four parts. Since commitments about sample size must be made before analysis can proceed, Section 5.1 describes two samples of data obtained from the model. Section 5.2 provides an introduction to principal component analysis, and describes the analysis performed on the two data sets. The issue of sampling is addressed by determining whether a basis computed for the smaller data set is sufficient to faithfully represent the larger. In Section 5.3, the results of the

[3]Importantly, it is not claimed that the CNS is performing a principal component analysis on its proprioceptive data, and then performs a linear regression. Rather, using terminology from Marr (1982), this work aims to elucidate the nature of the computational level, offering a particular algorithmic solution that might provide some hypotheses to be tested with neurophysiological data.

statistical analysis are presented. Again, the validity of the sample size is assessed by testing the ability of a model fit to a smaller data set to predict the larger one. Section 5.4 offers a way to calculate the range of the sample, using a convex hull to determine whether a given point is within the sample; this is important for ensuring validity of the output of the statistical models. Finally, in Section 5.5 the scope of the investigation is broadened by considering antisymmetrical tongue postures.[4] A sampling technique is used that overcomes the (relative) intractability of sampling antisymmetrical tongue movement. The results are discussed in Section 5.6.

5.1 Sampling

The purpose of selecting a sample of tongue postures to study is to be able to make generalizations about a much larger set of tongue postures, the set of *possible* tongue postures. This complete set includes as subsets the postures relevant to particular lingual behaviors (speech production, mastication, etc.). Aiming to make generalizations about all possible tongue postures, as opposed to those relevant to a particular behavior, is advantageous for two reasons. The first is that results obtained are relevant to all lingual behaviors, and are therefore maximally interesting to researchers. The second is that an exhaustive sample requires little to no prior knowledge of the tongue. No research exists that would indicate how to generate tongue postures specific to particular a particular kind of lingual behavior (say, by emphasizing certain muscles). Therefore a sample of all possible tongue postures is necessary.

Sampling the space of possible tongue postures—sampling the activation space—is difficult because the tongue is controlled by a fairly large number of parameters. As such, sampling the activation space runs into "the curse of dimensionality," a problem with a name that nobody finds overly dramatic who has worked with high-dimensional data. Figure 5.1 illustrates the curse of dimensionality. Suppose that

[4]The term "antisymmetrical" is used to refer to postures that are either symmetrical or asymmetrical.

in a one-dimensional space a researcher judges that ten samples are required to adequately sample a particular interval (left panel). When a second dimension is added, achieving the same sampling precision requires ten times as many samples (middle panel). If a third dimension is added, the requisite number of samples again rises tenfold, to one thousand (right panel). The number of samples require rises exponentially with each added dimension, making it very difficult to sample a high-dimensional space.

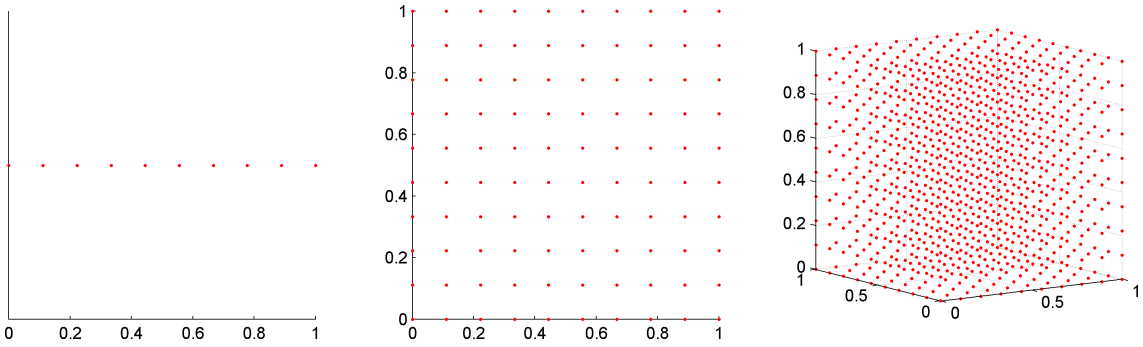

Figure 5.1: Illustration of the curse of dimensionality. Sampling a space becomes more difficult with more dimensions. (Values are arbitrary and dimensionless.)

Considering now the particular case of the tongue, for P activation parameters sampled at L activation levels, there are L^P possible parameter settings, i.e., L^P simulations to run. In consequence, sampling the parameter space at either many activation levels (large L) or for a large number of control parameters (large P) becomes computationally very expensive.

The number of samples required increases faster as L rises, but the number of samples also increases quickly with greater numbers of control parameters. The simplest activation scheme is symmetrical contraction of whole muscles, which results in seven control parameters. Sampled at two levels, there are $2^7 = 128$ combinations, or at three levels, $3^7 = 2,187$ combinations. Modeling antisymmetrical whole muscle contractions at two levels produces $2^{14} = 16,384$ combinations—almost eight times as many as the symmetrical case. Sampling antisymmetrical whole muscle contractions at three levels produces $3^{14} = 4,782,969$ combinations.

To put these figures in context, a typical simulation requires 2–6 hours at present-

day computing speeds. Simulating 2,187 tongue postures requires approximately one year of processor time.[5] In practice a year of calendar time is not required, since the simulations can be run simultaneously with many processors. Nevertheless, time imposes limits on what simulations can be run. Sampling the model behavior where muscle compartments are concerned, with 184 control parameters, would be extremely expensive. Sampling this space at two activation levels would require more than a septendecillion simulations ($> 2.45 \times 10^{55}$). At this level storing the results of the simulations would be impossible, as of this writing, to store the results of so many simulations.

The smallest non-trivial value for L is two, corresponding to no activation and full activation. If antisymmetrical whole muscle contractions were modeled, $2^{14} = 16,384$ simulations would necessary, which is not feasible. It was therefore decided to limit the data to symmetric muscle contractions, which have just seven control parameters. At seven control parameters, sampling at two activation levels and three activation levels is computationally tractable (128 and 2,186 samples, respectively; at four levels, 16,384 simulations are required).

To test whether a small sample of tongue postures can be successfully used to make generalizations about a larger sample, analyses and predictions based on the two-level data set were used to make testable predictions about the three-level data set.

The two-level data set, D_1, is the set of postures associated with A_1, the set of possible activations at two activation levels, for seven activation parameters. The two levels were full activation (0.35 MPa active contraction) and no activation (0 MPa active contraction).

$$A_1 = \{(a_1, a_2, a_3, a_4, a_5, a_6, a_7) | a_1 \ldots a_7 \in \{0, 0.35\}\} \tag{5.1}$$

In the second set, D_2, was derived from a set of activation parameters, A_2, in which three levels were used: no activation (0 MPa), half activation (0.175 MPa), and full

[5] 4 hours/run $\times$ 2187 runs $\times$ 1/8760 year/hours = 0.9986 year.

activation (0.35 MPa).

$$A_2 = \{(a_1, a_2, a_3, a_4, a_5, a_6, a_7) | a_1 \ldots a_7 \in \{0, 0.175, 0.35\}\} \tag{5.2}$$

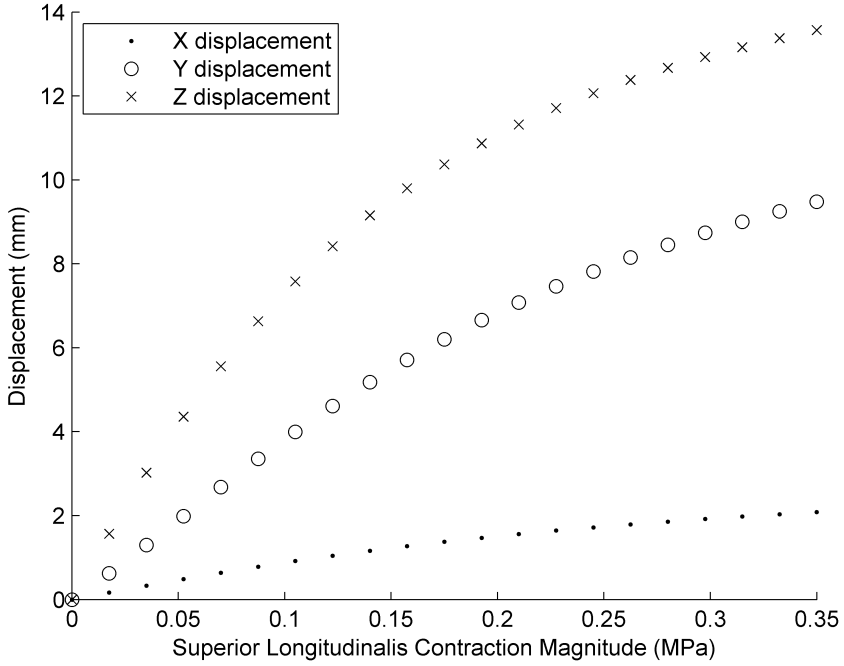

Figure 5.2: Displacement of the tongue tip as a function of magnitude of contraction of superior longitudinalis. The absolute value of displacement is shown.

Note that $A_1 \subset A_2$ and therefore that $D_1 \subset D_2$. The data sets are identified separately for expository purposes, since one of the goals of the chapter is to determine the extent to which the smaller data set can be used to generalize about the larger.

The evenly spaced activation values in D_2 are intended to capture some of the non-linearity in the stress-displacement relationship. Given the non-linearity that stems from the geometry of the tongue and the properties of muscle, it would not be surprising if there were a sampling scheme that would more effectively sample tongue postures. There are currently no data that would suggest what that sampling schema would be, however. The choice of 0, 0.175, and 0.35 MPa as sampling

points is heuristic, being based on a very small subset of data. Figure 5.2 shows the displacement of the front-most node as a function of the magnitude of contraction of a single tongue muscle, superior longitudinalis. If displacement of this node were estimated by interpolating linearly between 0 MPa and 0.35 MPa activation, the estimated value at 0.175 MPa would have nearly the greatest error. Therefore sampling at that point should remove a great deal of error in the three-level data set, when generalizations are made to other activation levels. Although this sampling committment is based on the displacement of a single node of the mesh in a single stimulation schema, there is no *a priori* reason to assume that a substantially different result would obtain for a different node in the mesh, or under other stimulation conditions.

It should be emphasized that the sample collected did not involve movement of the jaw or hyoid bones. This is a significant restriction, since the tongue rests on the jaw and hyoid, and the structures are able to move relative to one another. This limitation is further discussed in Section 7.2.1.

5.2 Dimension Reduction

This section discusses principal components analyses of the data. Principal components analysis has been used fairly extensively in analyzing vocal tract shapes. The technique is appropriate because, while many measurements are required to quantify a vocal tract shape, the shape of the vocal tract is determined by a relatively small number of parameters. Previous principal components analyses have focused on the English vowel space: Harshman et al. (1977) performed components analyses on X-ray tracings of the English vowels. Story and Titze (1998) performed a similar analysis, but using three-dimensional vocal tract shapes from MR images. Sanguineti et al. (1997, 1998) perform analyses on X-ray data, and on control parameters from a particular motor control model. The present work is most closely related to that of Perrier et al. (2000). These authors performed a principal component analysis on the set of tongue postures sampled from a biomechanical tongue

model, a two-dimensional model based on Payan and Perrier (1997). They found that they could account for a great deal of the variance in the data with just two components. The results of these studies are compared to the results of the present study in Section 5.6.1.

A brief introduction to principal component analysis is given in Section 5.2.1. Section 5.2.2 describes an initial dimension reduction that was necessary before performing the principal component analysis. Sections 5.2.3 and 5.2.4 present analyses of D_1 and D_2, respectively. Section 5.2.5 considers the projection of D_2 onto the basis derived from D_1.

5.2.1 Introduction to principal component analysis

A principle component analysis consists in rotating the axes of a data set to reduce correlation between dimensions, and in discarding dimensions that seem unimportant. A graphical example of this is considered in Figure 5.3. Suppose that an experimenter conducted an experiment and obtained the data points that are plotted in the left graph. Each point is characterized by X and Y coordinates, and so can be plotted on two axes. With a principal component analysis, a new pair of axes is determined to represent the data, X' and Y'. The first axis is the line that best fits the correlation between X and Y. The second axis is the line perpendicular to the first. The data are replotted in the right graph. Note that the data are the same: the difference is only that the axes have been rotated.

In this particular case, 99.7% of the variance in the data lies along X', and Y' has just 0.3% of the variance. An analyst might conclude that the Y' axis has unimportant variation, discard it, and instead represent each data point with a single number, X'. The advantage is each data point is represented by one number instead of two, which is easier computationally and conceptually. In a two-dimensional example the advantage is negligible, but in a case with thousands of dimensions, the advantage is much greater.

To perform a principal component analysis, the mean value of each dimension is first subtracted from the data set. Then the sample covariance matrix is calculated.

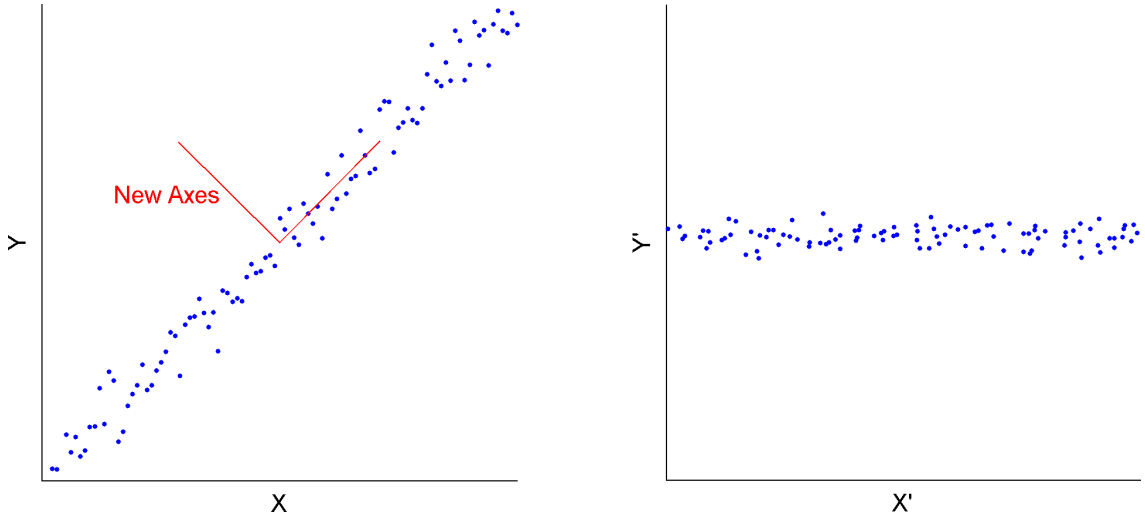

Figure 5.3: Geometric illustration of a principal component analysis. Collected data, before (left) and after (right) a transformation to a new coordinate system.

The eigenvectors of this matrix are the basis for the new vector space. The values of the new data are found by projecting the data into the new basis. The new basis has as many dimensions as the old, but again, dimensions can be discarded if they account for little of the variance in the data.

A particular PCA is based on a sample of data, so the calculated covariance matrix is an estimate of the true covariance matrix of the complete data set (hence *sample* covariance matrix). A basis calculated from a principal component analysis is associated with a particular data set. Here, $\mathbb{P}_1$ is the basis calculated from the analysis of D_1, and $\mathbb{P}_2$ is the basis calculated from the analysis of D_2. (The number of vectors in each basis is described below.) When a data set is projected onto another data set, the new basis is indicated with a superscript. $D_1^{\mathbb{P}1}$ is D_1 projected onto $\mathbb{P}_1$. $D_2^{\mathbb{P}1}$ is D_2 projected onto $\mathbb{P}_1$. $D_2^{\mathbb{P}2}$ is D_2 projected onto $\mathbb{P}_2$.

It is useful to be able to quantify the variance represented in a particular basis vector. Here a technique is used that is based on Jolliffe (2002, 112–113). If all eigenvectors of the covariance matrix are retained for the new basis, then the analysis is simply a rotation of the basis vectors. In this case the variance of the original and

projected data are the same. The difference is that after the analysis, variance is concentrated in the first few dimensions. Since the variances of the sets of coordinates are equal, the percentage of variance explained by the i^{th} component of the new basis is:

$$PE_i = \frac{r_i}{\sum_{k=1}^{p} s_k} \tag{5.3}$$

Where r_i is the variance of the data in the i^{th} vector of the new basis, and s_i is the variance of the data in the i^{th} vector of the old basis.

In the new bases, $\mathbb{P}_1$ and $\mathbb{P}_2$, basis vectors (or dimensions) will be referred to as "posture parameters" or "parameters," since they represent parameters along which tongue posture varies.

5.2.2 Initial downsampling of the mesh

The initial dimensions of the data set are the coordinates of each surface node. In three-dimensional space, this makes $3N$ dimensions for N nodes, and a $3N \times 3N$ covariance matrix. For the model used, this would mean a $33,993 \times 33,993$ covariance matrix. Presently, determining the eigenvectors of a such a large matrix is computationally impractical. Somewhat ironically then, a straight principal component analysis is made impossible by the number of dimensions in the model.

The number of nodes in the mesh was determined by the computational properties of the finite element model (Section 4.8), however, *not* by an appraisal of the necessary geometrical precision of the tongue model. For practical purposes (e.g., calculating vocal tract area functions, or forming a general impression about tongue movement), a coarser mesh is sufficient. Principal component analysis becomes tractable if we reduce the precision of the mesh before the analysis.

The task is to identify a proper subset of surface nodes that adequately represent the shape of the tongue. As it happens, there are well established algorithms for reducing the level of detail of a triangular mesh.[6] One commonly used algorithm is

[6]The technique was actually developed for computer games. It is computationally advantageous to represent computer characters with as few triangles as possible, particularly when they occupy only a small space on the screen.

that of Garland and Heckbert (1997).[7] This algorithm accepts a triangular mesh as an input, along with a target number of faces. The algorithm primary simplifies the model by eliminating nodes and edges. The judgment of the author was that the exterior finite element mesh was represented sufficiently with 2000 triangles, or 960 nodes. Figure 5.4 has a comparison of the full and reduced models.

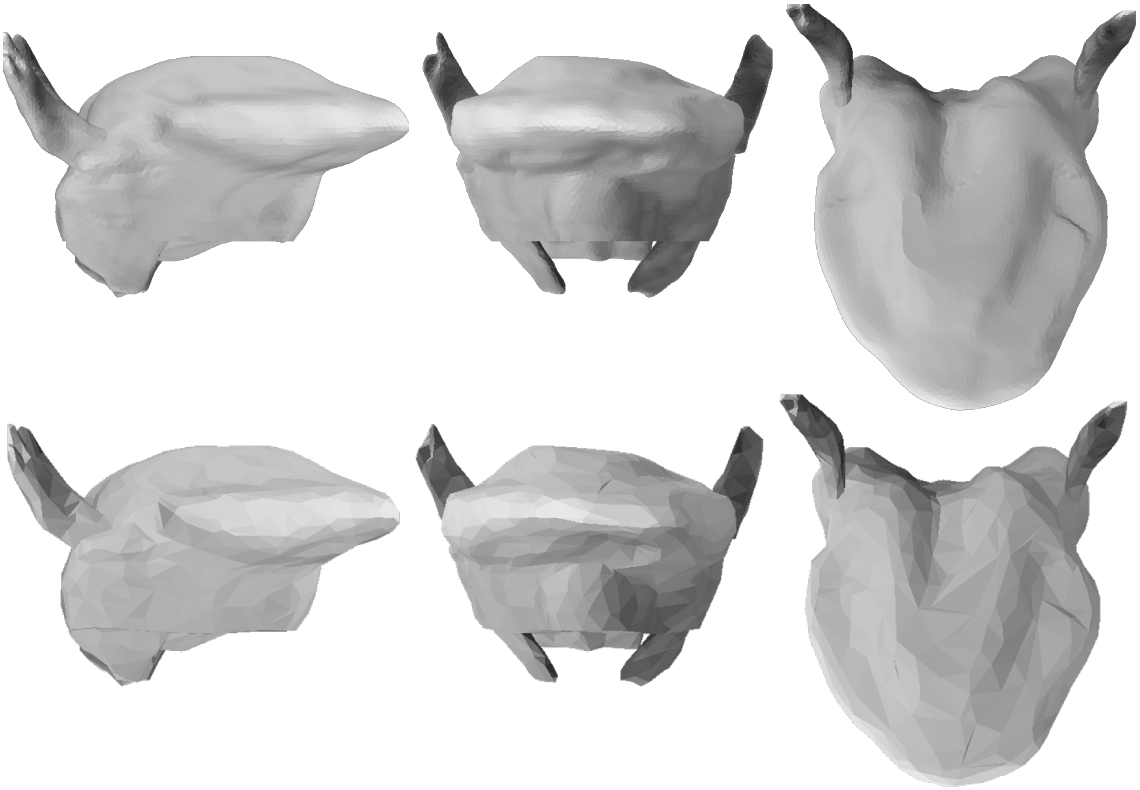

Figure 5.4: Full and reduced surface meshes.

5.2.3 $D_1^{\mathbb{P}1}$: D_1 projected onto $\mathbb{P}_1$

A principal component analysis on D_1—the set of postures resulting from all possible muscle activation patterns at *two* levels of activation—was performed. It was determined that nearly all of the variance in the node positions could be accounted

[7]The first author of this paper has made Qslim, an implementation of the algorithm, available at: `http://graphics.cs.uiuc.edu/~garland/software/qslim.html`.

for with seven parameters (hereafter, $p_1 \ldots p_2$). A Pareto diagram is shown in Figure 5.5. The bar plot shows the percentage of the variance explained by each parameter; the line plot shows the cumulative percentage of variance explained. The greater part of the variance (63.65%) is accounted for by p_1. The first three parameters explain 91.89% of the variance. There is a notable drop-off in the variance explained by each parameter after p_3. With seven parameters, 99.21% of the variance is explained. 100% of the variance is explained with 64 parameters. There is no absolute guideline that indicates how many parameters to retain in analysis. It was decided to include the first seven parameters in $\mathbb{P}_1$, so that greater than 99% of the variance could be accounted for.

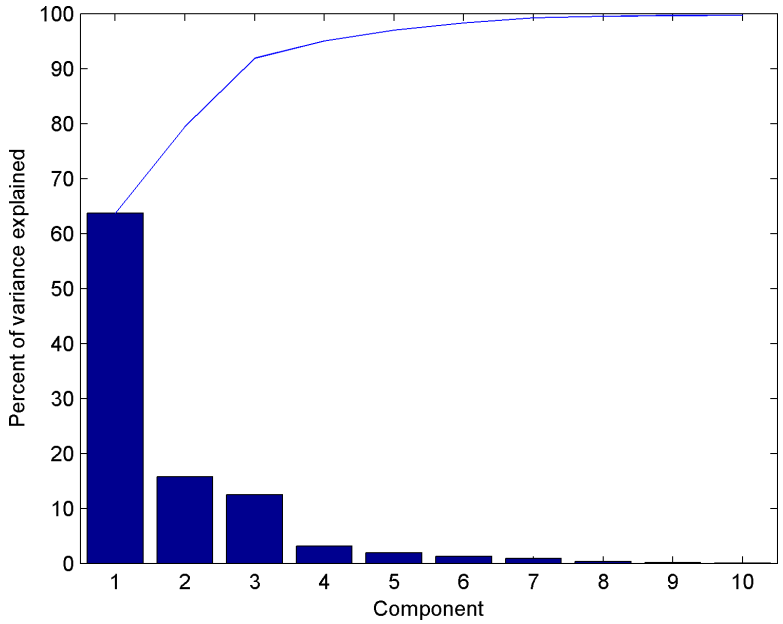

Figure 5.5: Pareto diagram of variance in D_1 accounted for by each parameter of $\mathbb{P}_1$.

5.2.3.1 Movements associated with components

In a principal component analysis, the question naturally arises whether any of the components have meaningful interpretations. To answer that question, a program was written to visualize the tongue deformations associated with different

components, Tongue Viewer, which is attached to the dissertation as supplemental material, and described in Appendix C. It is difficult to produce two-dimensional graphics that convey the deformations of a three-dimensional object, given variable parameters. Impressionistic descriptions of the tongue deformations associated with each component follow.[8] The descriptions are necessarily imprecise. Additionally, they represent a very small portion of the data. The result of varying each parameter is described, with the other parameters having been set to the values of the undeformed shape. The parameters combine, of course, to make other shapes. The reader is referred to Tongue Viewer for more complete understanding.

p_1 is associated with movements of the tongue tip. At negative values, the tongue is retroflexed slightly; at positive values the tongue is lowered and protruded (illustrated in Figure 5.6 and Figure 5.7, respectively).[9] For p_2, at negative values the tip is slightly drawn back and retroflexed, while the tongue body lowers, and the tongue root is advanced. At positive values the tip is lowered and protruded somewhat, the body of the tongue is narrowed, and the tongue root is retracted slightly (Figure 5.8; Figure 5.9). Variation along p_3 corresponds to bunching of the dorsum (with consequent dorsal raising) or flattening of the tongue dorsum (with some lowering) (Figure 5.10; Figure 5.11). p_4 shortens and lengthens the tongue along the anteroposterior axis (Figure 5.12; Figure 5.13). p_5 is difficult to describe: as it becomes more negative, the tongue dorsum raises slightly, lifting the sides of the tongue, and pulling the tip back slightly (Figure 5.14; Figure 5.15). The primary movement associated with p_6 is widening and narrowing of the tongue root (Figure 5.16; Figure 5.17). Movement along p_7 corresponds to a slight raising of the tongue dorsum and retraction of the tongue tip (Figure 5.18; Figure 5.19).

[8]It is somewhat ironic that a dissertation concerned with a *quantitative* tongue model should be confined to informal description of tongue deformation. These are provided merely as a guide. The quantitative data are of course available.

[9]Since the direction of the eigenvectors is arbitrary, these signs are without meaningful interpretation. They are used here purely for expository purposes.

Figure 5.6: $D_1^{\mathbb{P}1}$: p_1 at minimum value.

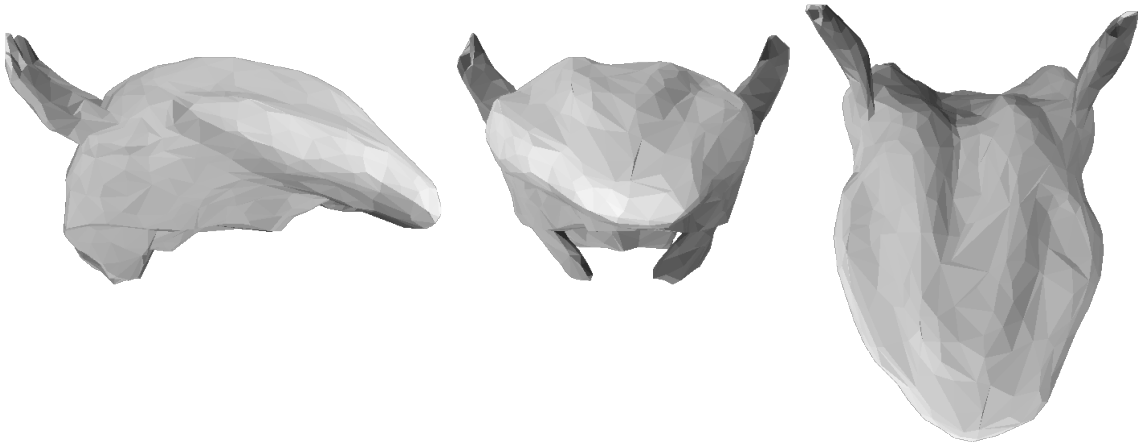

Figure 5.7: $D_1^{\mathbb{P}1}$: p_1 at maximum value.

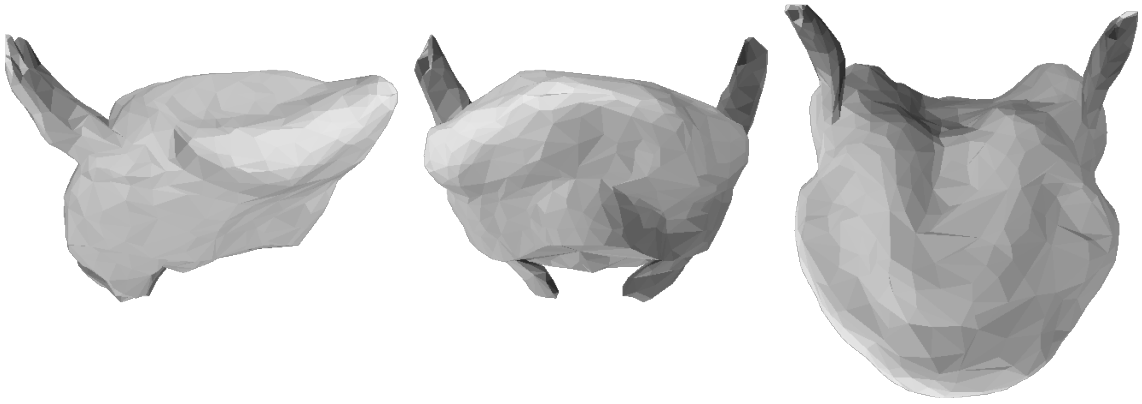

Figure 5.8: $D_1^{\mathbb{P}1}$: p_2 at minimum value.

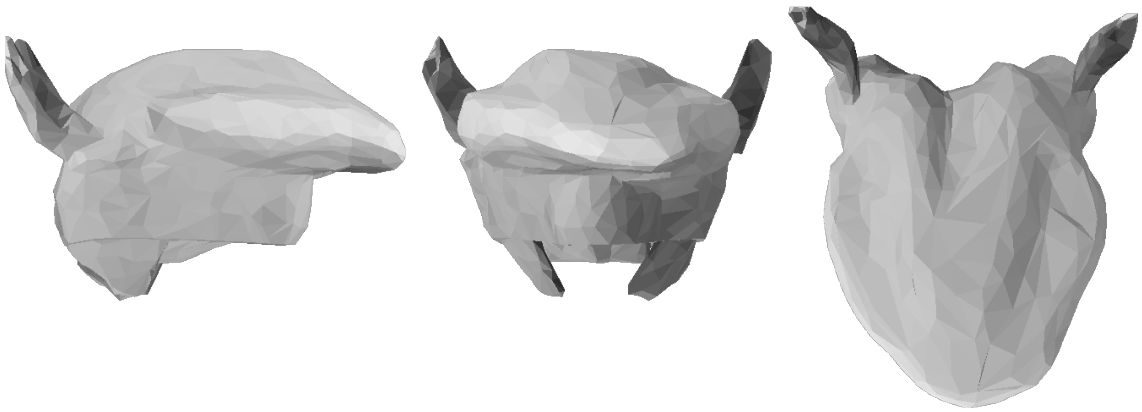

Figure 5.9: $D_1^{\mathbb{P}1}$: p_2 at maximum value.

Figure 5.10: $D_1^{\mathbb{P}1}$: p_3 at minimum value.

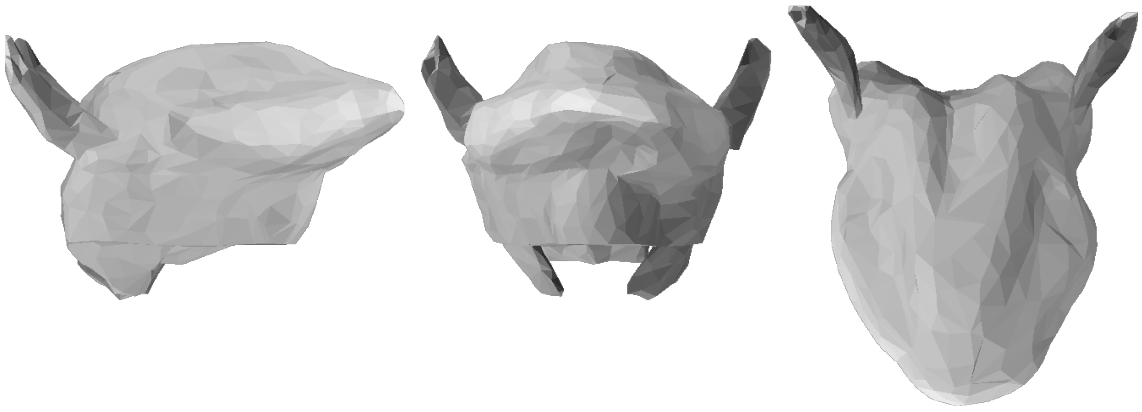

Figure 5.11: $D_1^{\mathbb{P}1}$: p_3 at maximum value.

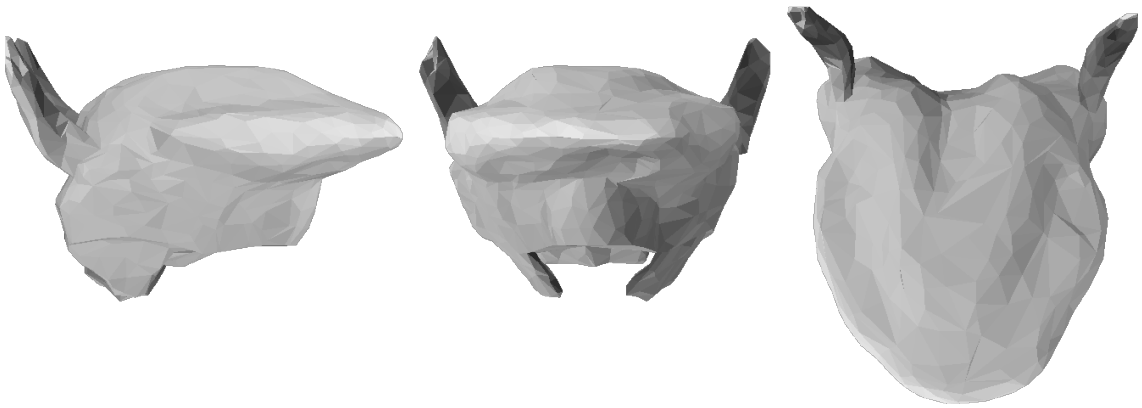

Figure 5.12: $D_1^{\mathbb{P}1}$: p_4 at minimum value.

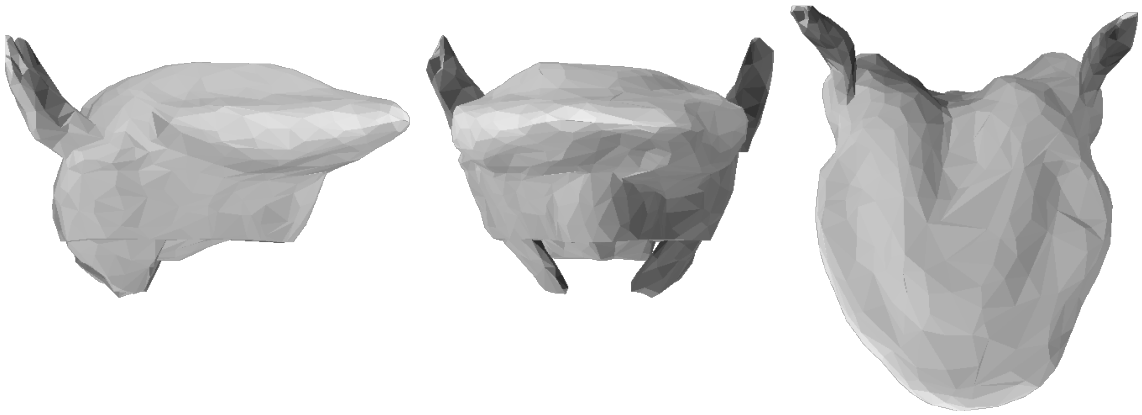

Figure 5.13: $D_1^{\mathbb{P}1}$: p_4 at maximum value.

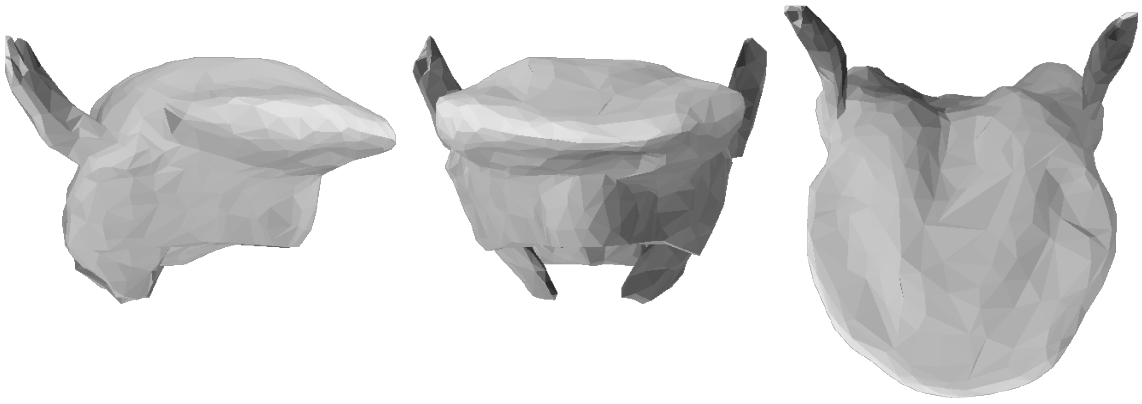

Figure 5.14: $D_1^{\mathbb{P}1}$: p_5 at minimum value.

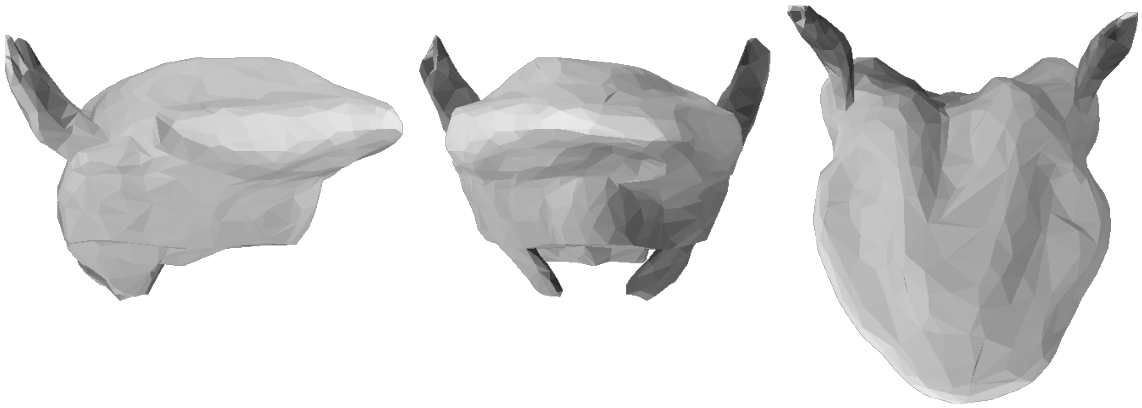

Figure 5.15: $D_1^{\mathbb{P}1}$: p_5 at maximum value.

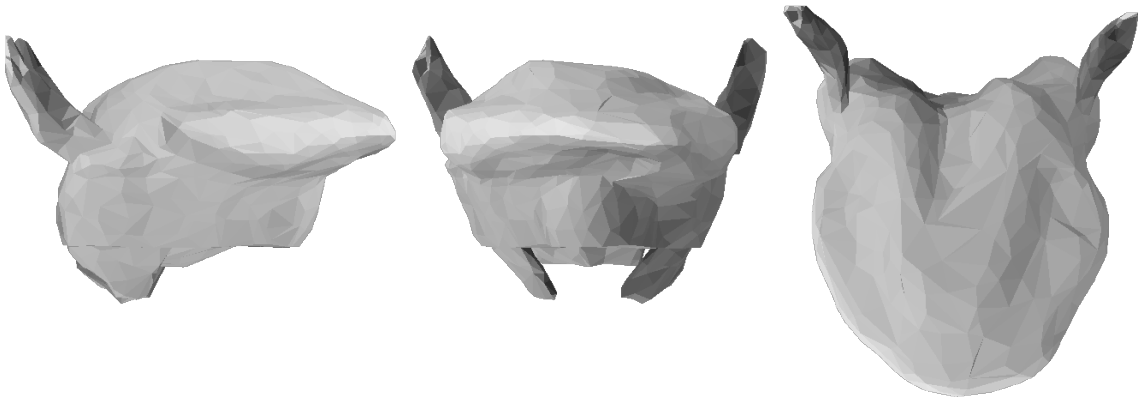

Figure 5.16: $D_1^{\mathbb{P}1}$: p_6 at minimum value.

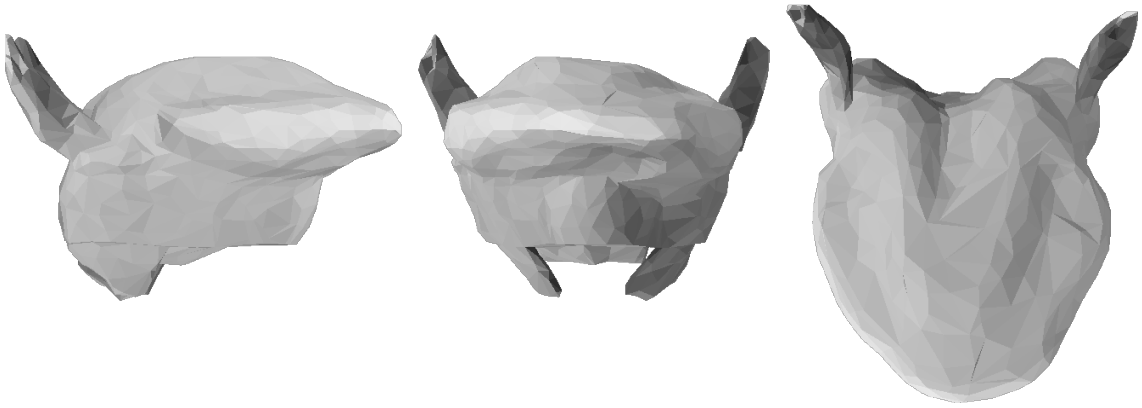

Figure 5.17: $D_1^{\mathbb{P}1}$: p_6 at maximum value.

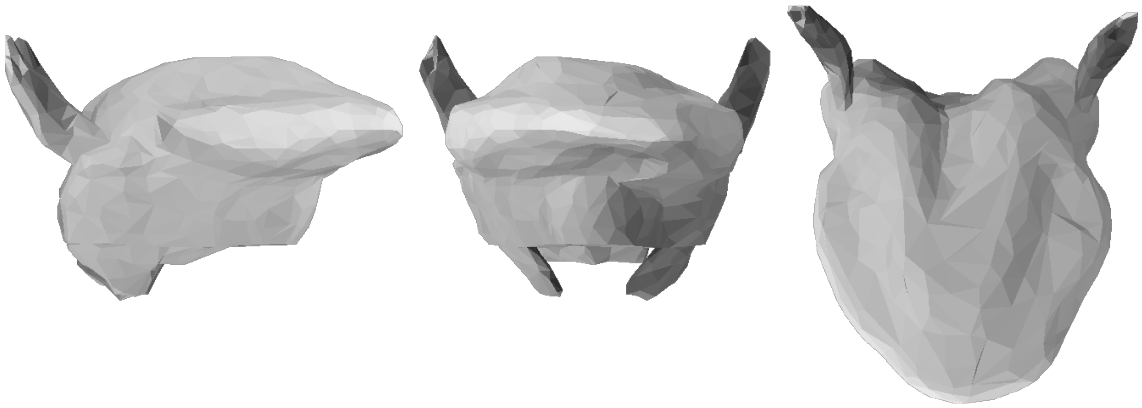

Figure 5.18: $D_1^{\mathbb{P}1}$: p_7 at minimum value.

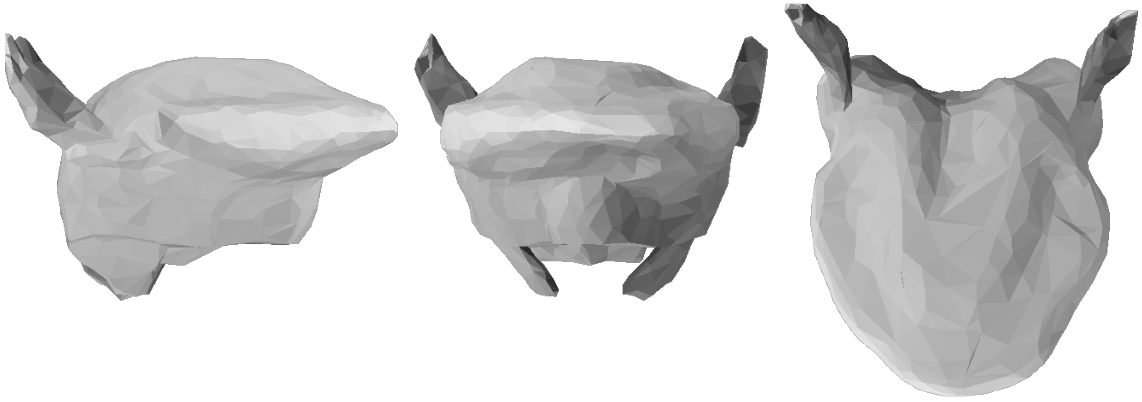

Figure 5.19: $D_1^{\mathbb{P}1}$: p_7 at maximum value.

5.2.4 $D_2^{\mathbb{P}2}$: D_2 projected onto $\mathbb{P}_2$

A principal component analysis was performed on D_2—the set of postures resulting from all possible muscle activation patterns at *three* levels of activation. A Pareto diagram is given in Figure 5.20. The results are nearly identical to those obtained for the D_1 data set. p_1 accounts for 63.60% of the variance. The first three components account for 92.13% of the variance, and the first seven account for 99.31%. It took 109 components to account for 100% of the variance. In this case as well, then, the decision was made to include only the first seven parameters in $\mathbb{P}_2$.

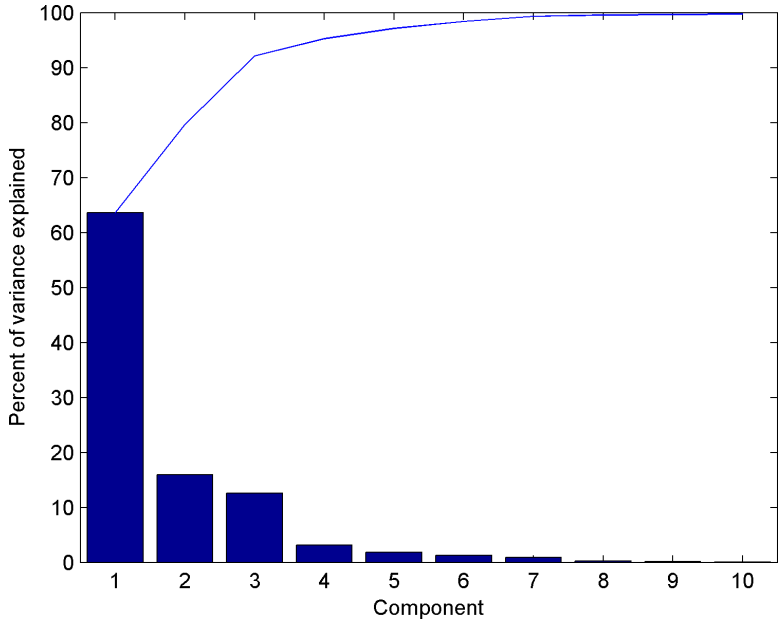

Figure 5.20: Pareto diagram of variance in D_2 accounted for by each parameter of $\mathbb{P}_2$.

5.2.4.1 Movements associated with components

Impressionistically, the changes in tongue posture associated with p_1 through p_7 for $D_2^{\mathbb{P}1}$ were identical to p_1 through p_7 for $D_1^{\mathbb{P}1}$. Figures 5.21 through 5.34 correspond to Figures 5.6 through 5.19.[10]

[10]In this and subsequent narration, the sign of the parameters has occasionally been flipped so that the positive and negative values correspond to those for $D_1^{\mathbb{P}1}$. This is only for ease of

Figure 5.21: $D_2^{\mathbb{P}2}$: p_1 at minimum value.

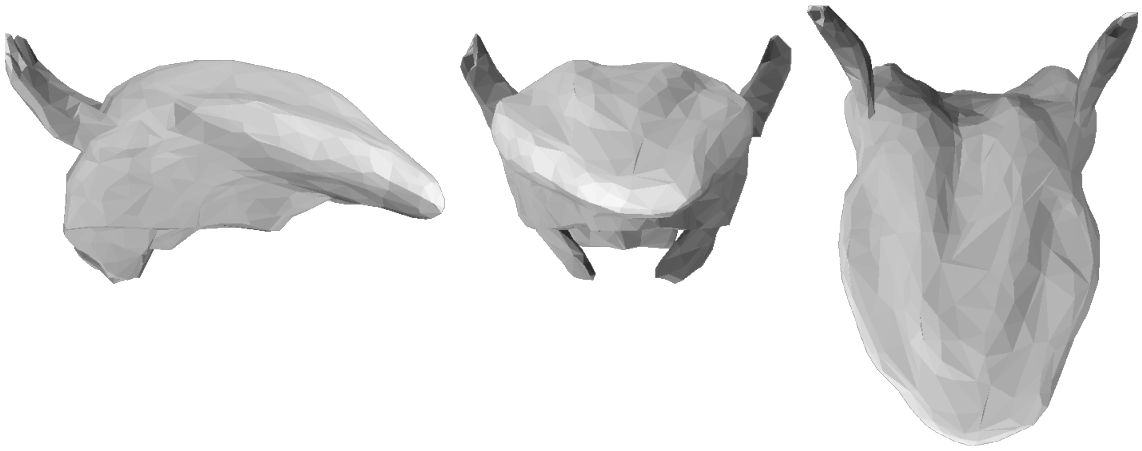

Figure 5.22: $D_2^{\mathbb{P}2}$: p_1 at maximum value.

Figure 5.23: $D_2^{\mathbb{P}2}$: p_2 at minimum value.

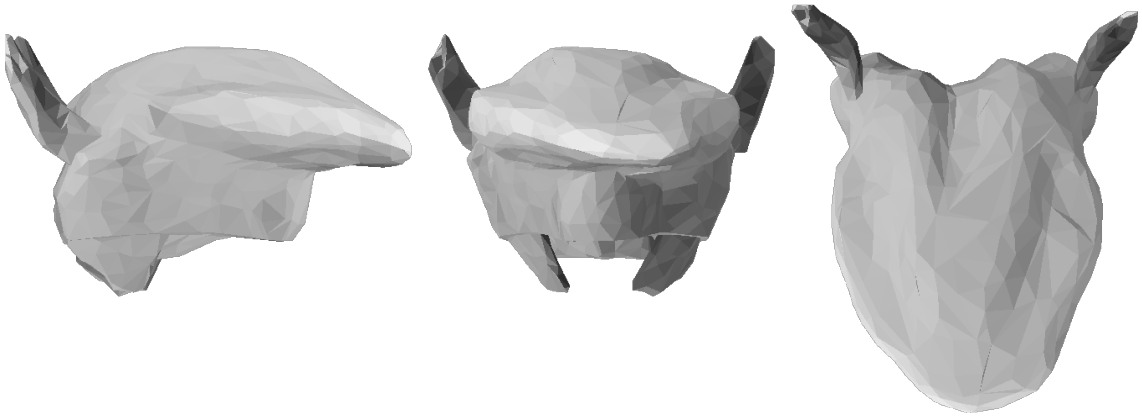

Figure 5.24: $D_2^{\mathbb{P}2}$: p_2 at maximum value.

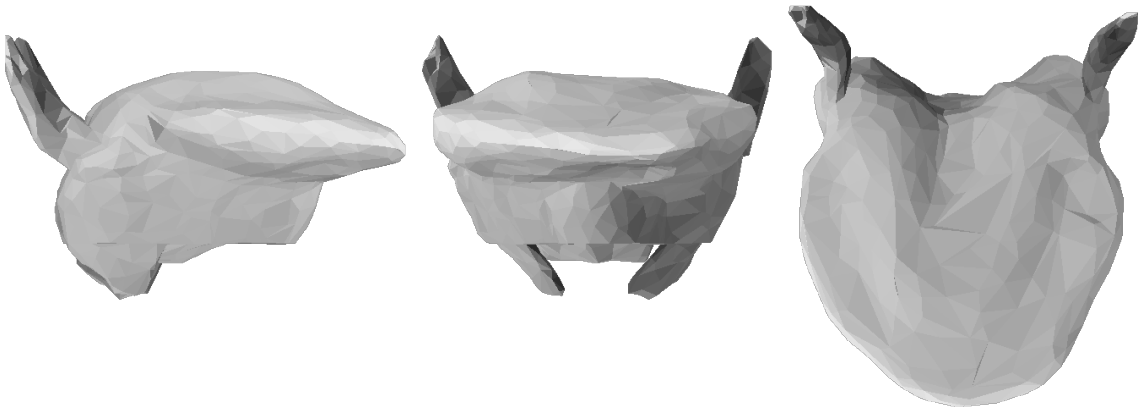

Figure 5.25: $D_2^{\mathbb{P}2}$: p_3 at minimum value.

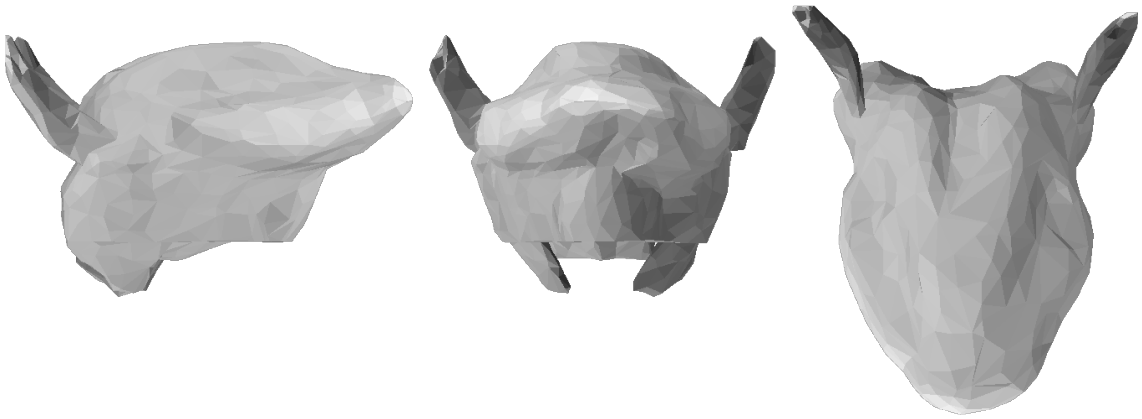

Figure 5.26: $D_2^{\mathbb{P}2}$: p_3 at maximum value.

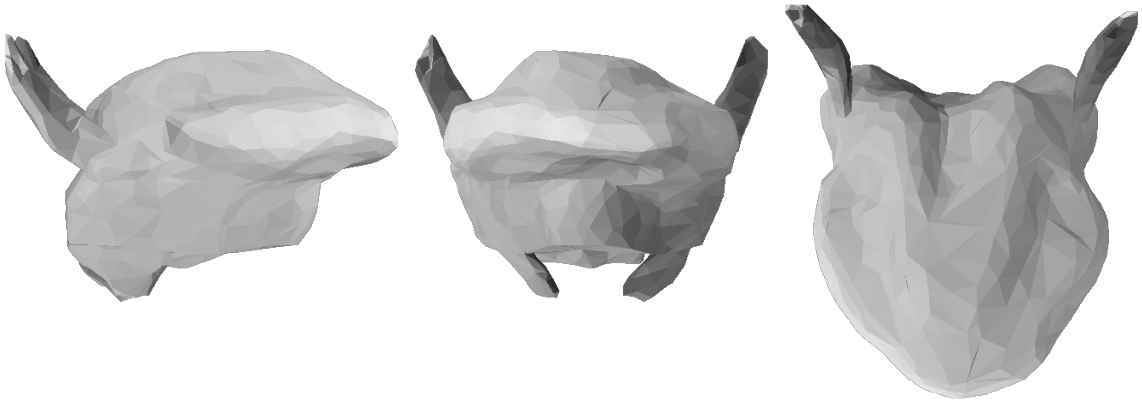

Figure 5.27: $D_2^{\mathbb{P}2}$: p_4 at minimum value.

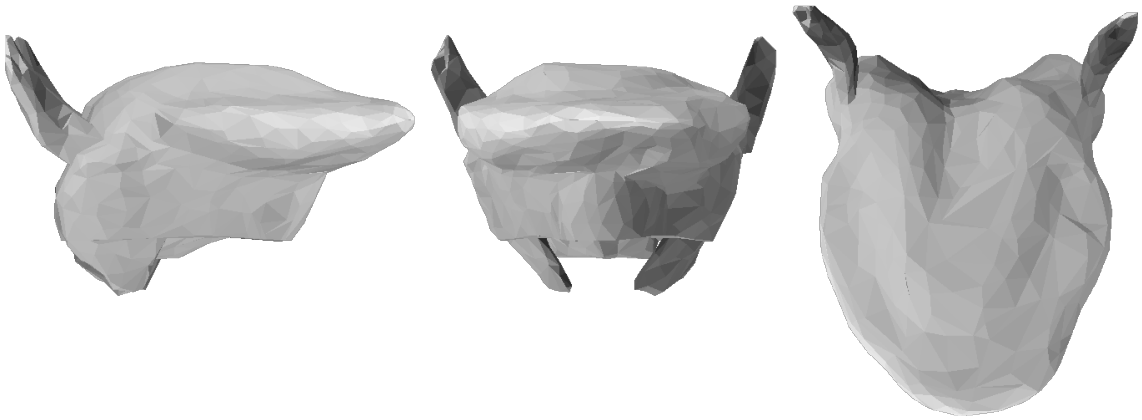

Figure 5.28: $D_2^{\mathbb{P}2}$: p_4 at maximum value.

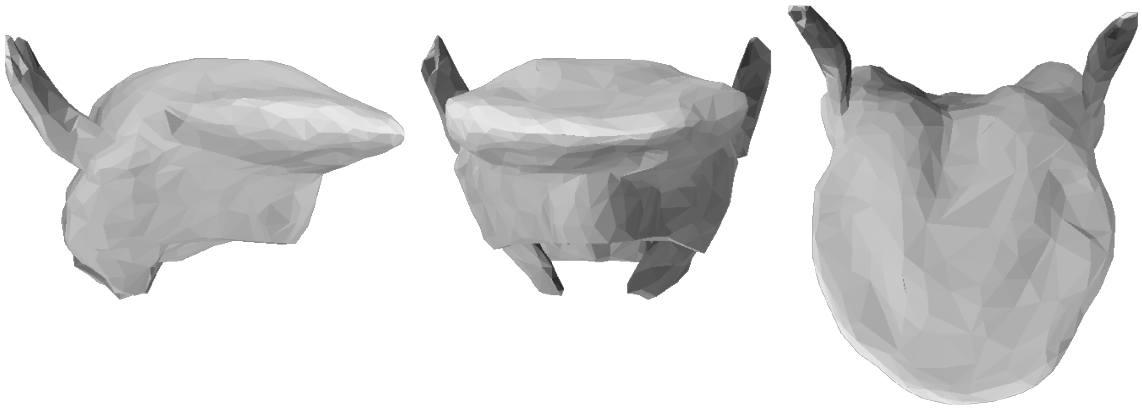

Figure 5.29: $D_2^{\mathbb{P}2}$: p_5 at minimum value.

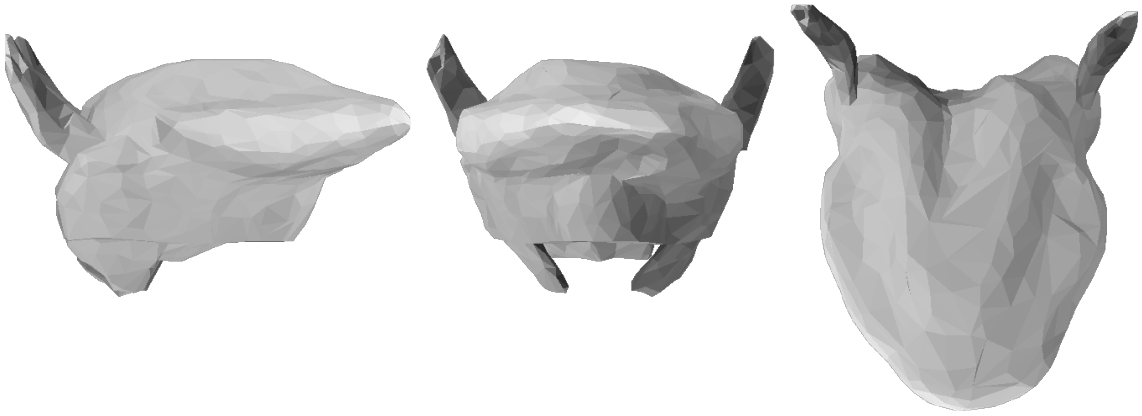

Figure 5.30: $D_2^{\mathbb{P}2}$: p_5 at maximum value.

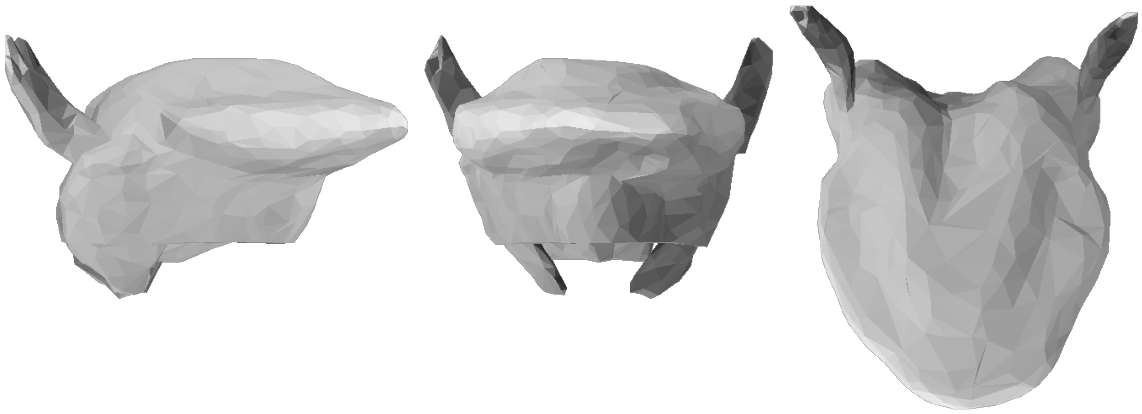

Figure 5.31: $D_2^{\mathbb{P}2}$: p_6 at minimum value.

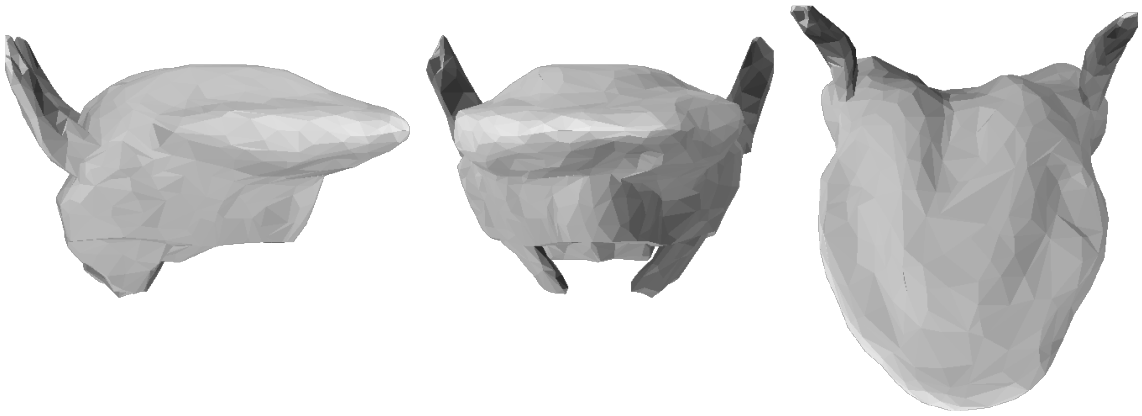

Figure 5.32: $D_2^{\mathbb{P}2}$: p_6 at maximum value.

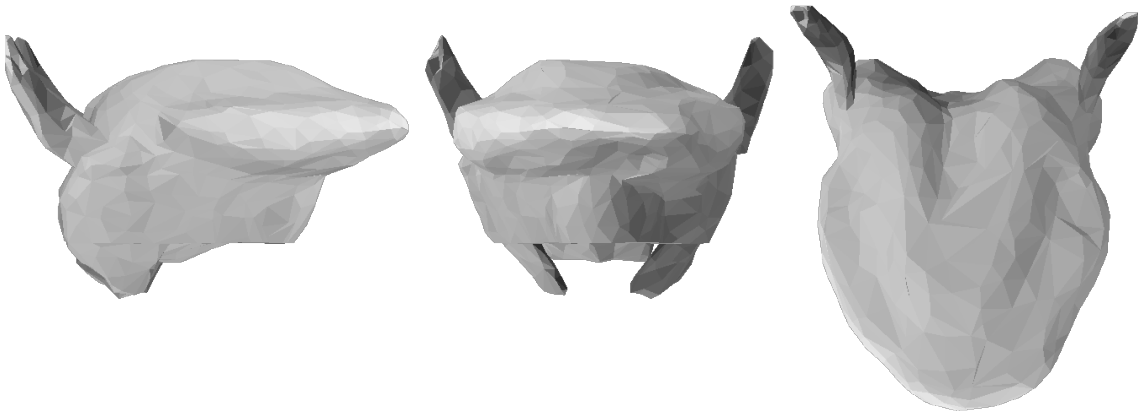

Figure 5.33: $D_2^{\mathbb{P}2}$: p_7 at minimum value.

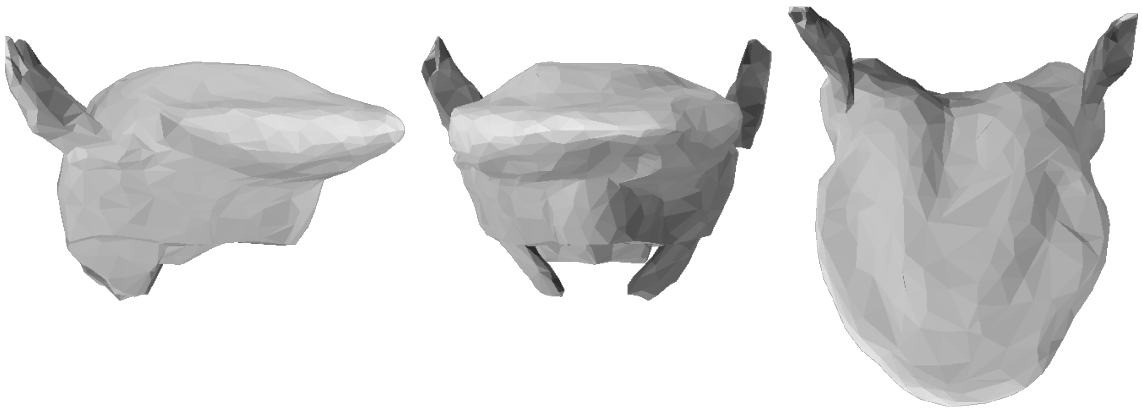

Figure 5.34: $D_2^{\mathbb{P}2}$: p_7 at maximum value.

5.2.5 $D_2^{\mathbb{P}_1}$: D_2 projected onto $\mathbb{P}_1$

One of the more important questions for evaluating the success of sampling is to determine whether $\mathbb{P}_1$ was sufficient, not just to represent the variance in D_1, but to represent the variance in D_2 as well. It is possible that the larger data set would require significantly more parameters to account for the data. To determine whether this was the case, the D_2 data were projected onto the *full* basis derived from D_1 (i.e., including all eigenvectors). The ratio of the variance of D_2 and this transformed data set was greater than 0.999999, which indicates that, as a whole, the vector space derived from D_1 is sufficient to represent the variance in D_2. Since these variances are nearly equal, it is legitimate to compare the variances of the original D_2 data set to those of the transformed data set.

The quantity of variance described by each dimension was assessed; this is shown in Figure 5.35. The first component accounted for 63.58% of the variance. Three components accounted for 92.10% of the variance. Seven components accounted for 99.28% of the variance. These results are quite similar to those for the $D_1^{\mathbb{P}_1}$ and $D_2^{\mathbb{P}_2}$ analyses, which indicate that $\mathbb{P}_1$ is well suited to describe D_2.

It is interesting to compare the range of the parameter values for $D_1^{\mathbb{P}_1}$ and $D_2^{\mathbb{P}_1}$; these are shown in Figure 5.36. In general, the range of the $D_1^{\mathbb{P}_1}$ parameters is the same as that of the $D_2^{\mathbb{P}_1}$ parameters. This is sensible since $D_1^{\mathbb{P}_1} \subset D_2^{\mathbb{P}_1}$. It is only in the fourth component that $D_2^{\mathbb{P}_1}$ has a slightly higher maximum than $D_1^{\mathbb{P}_1}$. This indicates that even the range of possible parameter values is also well predicted by $D_1^{\mathbb{P}_1}$. (This same point is made in Section 5.4, using a different technique for specifying the range of the data.)

5.3 Statistical Models

This section describes the attempt to use statistical models to predict parameter values from activation levels, and activation levels from parameter values. The motivation for the former is that a statistical model that maps muscle activations to

---

exposition; the sign is arbitrary in a principal component analysis.

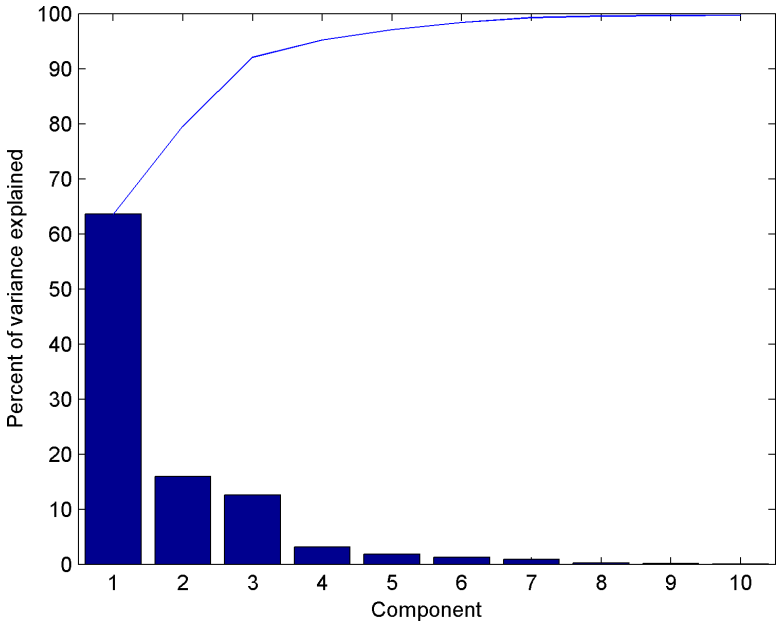

Figure 5.35: Pareto diagram of variance in D_2 accounted for by each parameter of $\mathbb{P}_1$.

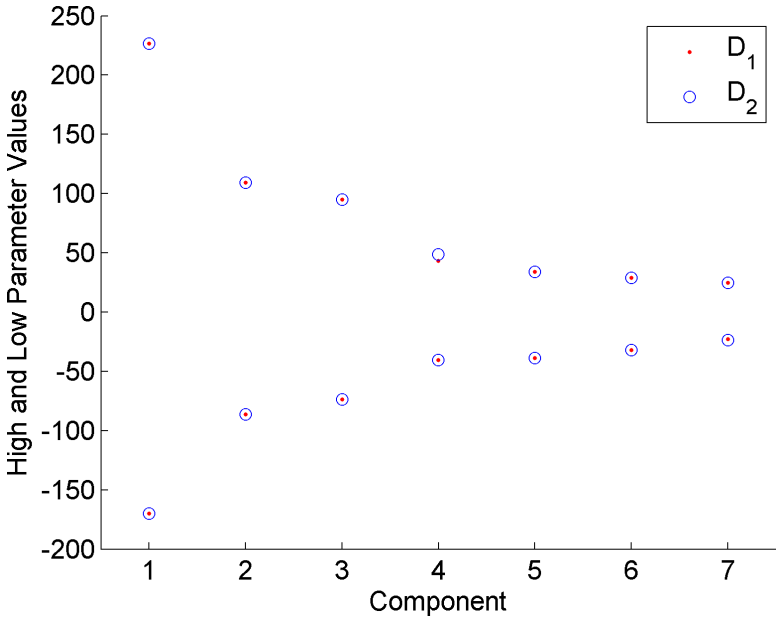

Figure 5.36: Parameter ranges for $D_1^{\mathbb{P}1}$ and $D_2^{\mathbb{P}1}$.

tongue postures is much less computationally expensive than a biomechanical model. The latter is motivated by the desire to go the other direction: to predict muscle activations, given a tongue posture. Two kinds of models are used for this purpose: a simple linear regression model and an artificial neural network (ANN). The results in this section demonstrate that statistical models can perform these mappings very well, which should broaden the appeal of the model considerably. Interestingly, different kinds of models are better suited to different mappings. In predicting parameter values from activation levels, the ANN produces better results, while for predicting activation levels from parameter values, linear models make better predictions.

The general strategy was to fit the statistical model using a particular data set, and test it with a set of data to which it had not been fit, in order to judge the extent to which the model generalizes successfully. This was done in two ways. First, a model would be fit to the D_1 data, and tested on data from the larger D_2 set.[11] In this case, the test is whether the model can generalize from a data set with two levels of activation to a data set with three levels of activation.

The second way of testing the statistical model was to determine whether it could make accurate predictions at levels of activation other than $\{0, 0.175, 0.35\}$. For this purpose, a special "testing" data set was generated. Ninety-nine activation patterns were created by setting each muscle activation to a value drawn from a random sample with a uniform distribution over the interval $[0, 0.35]$.[12] The tongue posture resulting from these activation was calculated using the biomechanical model, and the nodal coordinates were projected into $\mathbb{P}_1$. The statistical models were tested on their ability to map between the posture parameters and muscle activation levels, or from muscle activation levels to posture parameters. It is again emphasized that

[11] Since in Section 5.2.5 the $\mathbb{P}_1$ parameter space was found to be sufficient to represent D_2, in this section all data sets have been projected into $\mathbb{P}_1$. D_1 and D_2 are referred to rather than $D_1^{\mathbb{P}_1}$ and $D_2^{\mathbb{P}_1}$ for notational simplicity.

[12] One hundred such activations were generated initially, but only 99 are included in the testing set. For one activation pattern, the finite element solution did not converge, even after several runs through the solver. This activation pattern was therefore discarded. Since it was random to begin with, this is not expected to have systematically biased the results presented here.

the testing data set differed from any of the training data sets in that the muscle activation levels were drawn randomly from the interval $[0, 0.35]$, instead of from the set $\{0, 0.175, 0.35\}$.

Following the question raised in Section 5.2.5—i.e., whether a smaller data set (in this case, D_1) can be used to make accurate generalizations about larger one (here, D_2), it was of interest to see whether a statistical model fitted to the smaller data set made accurate predictions of a larger one. These tests were performed on the linear models. As non-parametric statistical models, artificial neural networks rely to a greater extent than do linear models on the training data, and are more susceptible to producing poor results with smaller training data sets; therefore in the neural networks no test was made whether D_1 could be used to generalize to D_2.

For the linear regression models, an intercept term and the simple effects are included, but not interactions of the independent variables. Since the mapping in either direction is multivariate, each report is actually a series of linear models. Adjusted r^2 values are reported for each dependent variable to indicate the success of the model in matching the data.

The neural networks are multilayer feedforward networks, trained with a conjugate-gradient backpropagation algorithm. A hyperbolic tangent transfer function was used for all neurons. Each neuron had a bias. The training set was divided randomly into a training and validation set. The validation consisted of either 10% or 20% of the data, depending on the run. During training the weights of the network were updated based on the training data, but not on the validation data; when root mean square error associated with the validation data increased, training was halted—the "early stopping" technique to prevent the network from overfitting the data. Specific network architectures are given below. In each case the optimal network architecture was determined informally by the researcher, with several runs each with different initial conditions. Root mean square error for the validation data was used as a guide.

Component	D_1	$(D_2 - D_1)$	Test
p_1	0.9651	0.9204	0.8514
p_2	0.9717	0.8977	0.5969
p_3	0.9715	0.9610	0.9565
p_4	0.9248	0.9237	0.9480
p_5	0.9489	0.9298	0.9067
p_6	0.9584	0.9337	0.9166
p_7	0.9748	0.8603	0.2088

Table 5.1: Adjusted r^2 for different data sets, predicted from the linear regression of D_1 on A_1.

5.3.1 Predicting parameter values from muscle activation

5.3.1.1 Linear model

A regression of D_1 on A_1 is quite successful in accounting for the D_1 data, indicating a fairly linear relationship between the posture parameter and muscle activation levels. The results are shown in Table 5.1. The adjusted r^2 values range from 0.9248 to 0.9748. These values are quite high; a perfect model would have an adjusted r^2 value of 1. Ideally, the goodness-of-fit could be assessed by comparing the values to other regressions models with similar independent and dependent variables, but such models do not seem to exist.[13]

To consider whether this model generalizes to further data, Table 5.1 also shows how the same linear model accounts for the variance in $(D_2 - D_1)$. This data set is D_2 with the overlapping data from D_1 removed, which ensures that the model does get "credit" for accounting for the data to which it was originally fit. The table shows diminished, but reasonably robust, performance. Adjusted r^2 values range from 0.8603 to 0.9610, which shows that the linear model generalized quite well to the larger data set. For the testing data, performance is diminished again, with adjusted r^2 values ranging from 0.2088 to 0.9480.

[13]Broadly similar studies do exist. Anderson and Fuglevand (2008) achieve r^2 values about $\sim$0.4 in predicting muscle activation from limb position in behaving human subjects. This is similar to

	Training (D_2)	Testing
p_1	0.9315	0.9126
p_2	0.9320	0.8362
p_3	0.9637	0.9689
p_4	0.9270	0.9614
p_5	0.9366	0.9397
p_6	0.9400	0.9382
p_7	0.9450	0.7789

Table 5.2: Adjusted r^2 values for training and testing data sets, predicted from the linear regression of D_2 on A_2.

When a linear model is fit to the full D_2 data set, generalization to the testing data is much better, with adjusted r^2 values ranging from 0.7789 to 0.9689 (Table 5.2).

5.3.1.2 Neural network

For mapping from muscle activations to posture parameters, the optimal network had seven units each in the first and second layers, and seven output units. Performance of the network was very good, both on the training and testing data sets. Table 5.3 provides adjusted r^2 values, the smallest of which is 0.9595.

5.3.2 Predicting muscle activation from parameter values

Here the converse situation is modeled: predicting muscle activation, given a tongue posture.

---

what is done below in Section 5.3.2. Anderson and Fuglevand's problem is more complex in many ways, however. For instance their muscle activation data have experimental noise, whereas in this dissertaton they do not.

	Training/Validation (D_2)	Testing
p_1	0.9939	0.9945
p_2	0.9919	0.9972
p_3	0.9981	0.9949
p_4	0.9892	0.9995
p_5	0.9969	0.9595
p_6	0.9994	0.9994
p_7	0.9991	0.9949

Table 5.3: Adjusted r^2 values for training and testing data sets, from a neural net trained to map muscle activations to tongue shape parameters.

	D_1	$(D_2 - D_1)$	Test
Genioglossus	0.9690	0.8556	0.5834
Hyoglossus	0.9577	0.9167	0.8530
Styloglossus	0.9789	0.8992	0.6914
Transversus	0.9746	0.9468	0.8574
Verticalis	0.9829	0.9536	0.9032
Superior Longitudinalis	0.9527	0.9110	0.8089
Inferior Longitudinalis	0.9000	0.8992	0.9077

Table 5.4: Adjusted r^2 values for different data sets, predicted from the linear regression of A_1 on D_1.

5.3.2.1 Linear model

Table 5.4 provides adjusted r^2 values from linear models that predict muscle activation from parameter values, where the linear models have been trained on the D_1 data. The model matches the D_1 quite well, with adjusted r^2 values ranging from 0.9000 to 0.9829. When the model generalizes to the data to which it was not fit, $D_2 - D_1$, values range from 0.8556 to 0.9536, which indicate good generalization. Generalization to the testing data is less successful, with adjusted r^2 values as low as 0.5834, though as high as 0.9077.

When a linear model is fit to the entire D_2 data set, generalization to the test data is more robust, as shown in Table 5.5. The adjusted r^2 values range from 0.8258 to 0.9598.

	Training (D_2)	Testing
Genioglossus	0.9131	0.8258
Hyoglossus	0.9372	0.9245
Styloglossus	0.9425	0.8922
Transversus	0.9645	0.9454
Verticalis	0.9689	0.9598
Superior Longitudinalis	0.9362	0.9162
Inferior Longitudinalis	0.9111	0.9393

Table 5.5: Adjusted r^2 values for training and testing data sets, predicted from the linear regression of A_2 on D_2.

	Training/Validation (D_2)	Testing
Genioglossus	0.9964	0.7298
Hyoglossus	0.9984	0.7377
Styloglossus	0.9967	0.6384
Transversus	0.9989	0.7203
Verticalis	0.9996	0.5926
Superior Longitudinalis	0.9997	0.7621
Inferior Longitudinalis	0.9997	-0.1711

Table 5.6: Adjusted r^2 values for training and testing data sets, from a neural net trained to map tongue shape parameters to muscle activations.

5.3.2.2 Neural network

The best neural network for predicting muscle activation levels from posture parameters had 20 units in the first layer, with seven output units. For the training data, performance is quite good; Table 5.6 shows the lowest adjusted r^2 value for is 0.9967 for the training data. Generalization to the test data was much poorer, however, with adjusted r^2 values ranging from -0.1711 to 0.7621. Reasons for this poor generalization are discussed in the next section.

5.3.3 Intermediate discussion

The important result from this section is that simple statistical models can be used to make very accurate predictions about tongue posture, given muscle activation,

or about muscle activation, given tongue posture. One application for these models would be to use the models to test an iterative motor learning algorithm. If each new tongue posture required hours to compute, it would not be feasible to test the algorithm with a biomechanical model. Since the statistical models produce their result more quickly (much less than 1 s), their use is to be preferred for such an application.

It is curious that the neural network outperforms the linear model in mapping activation levels to posture parameters, but underperforms it in mapping posture parameters to activation levels. The reason for this disparity seems to be the nature of the training sets. Figures 5.37 and 5.38 show the performance of the network, and also illustrate the differences in the range of the target values of the training and testing data. In mapping from muscle activation levels to posture parameters, both the training and testing data sets are continuously distributed over the range of the data.[14] This is because the posture parameters are taken from the biomechanical model, and are essentially a random sample of posture parameter space. The training data are good predictors of what the testing data look like. Conversely, in Figure 5.38 it is evident that the range of muscle activation levels in the training and testing data are quite different. In the training data the values come from the set $\{0, 0.175, 0.35\}$, whereas in the testing set they are randomly drawn from the interval $[0, 0.35]$. Training the network minimizes error for the targets $\{0, 0.175, 0.35\}$, but does not directly optimize the mapping for other values. It has no "experience" on the intervals $(0, 0.175)$ and $(0.175, 0.35)$. With this in mind, it seems likely that if a neural network were given as input a data set with a more uniform distribution of muscle activation levels, it would fit the data quite well.

5.4 Note on the range of the sample

Previously, in describing the range of tongue parameter values, the maximum and minimum values of D_1 and D_2 (projected onto $\mathbb{P}_1$ or $\mathbb{P}_2$) have been specified. Al-

[14] In fact, the distribution of each parameter value is approximately normal.

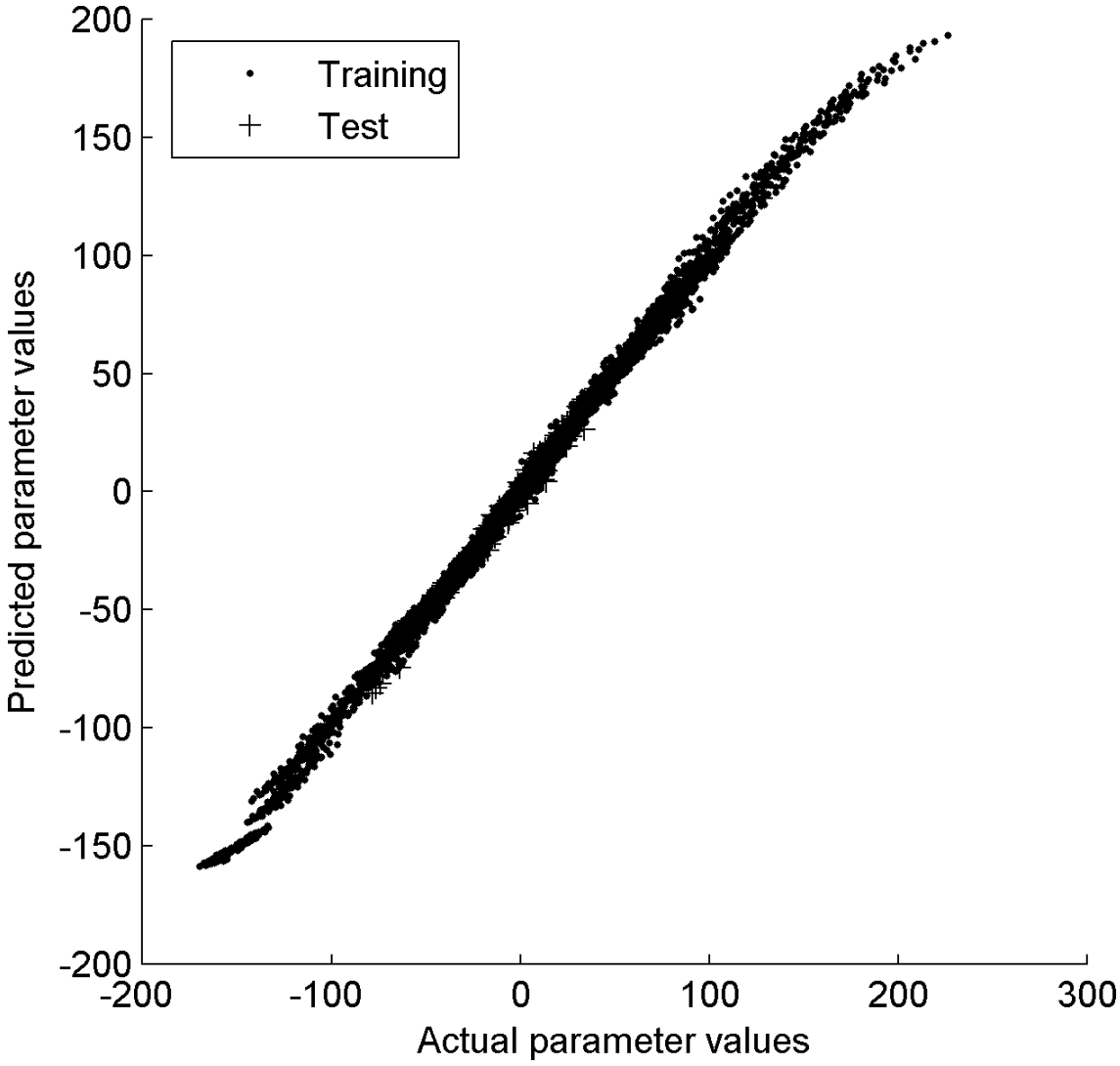

Figure 5.37: Network performance on testing and training data, mapping activation levels to posture parameters. Target values are plotted on the abscissa, and predicted values on the ordinate. If network performance were perfect, the points would fall on a straight line.

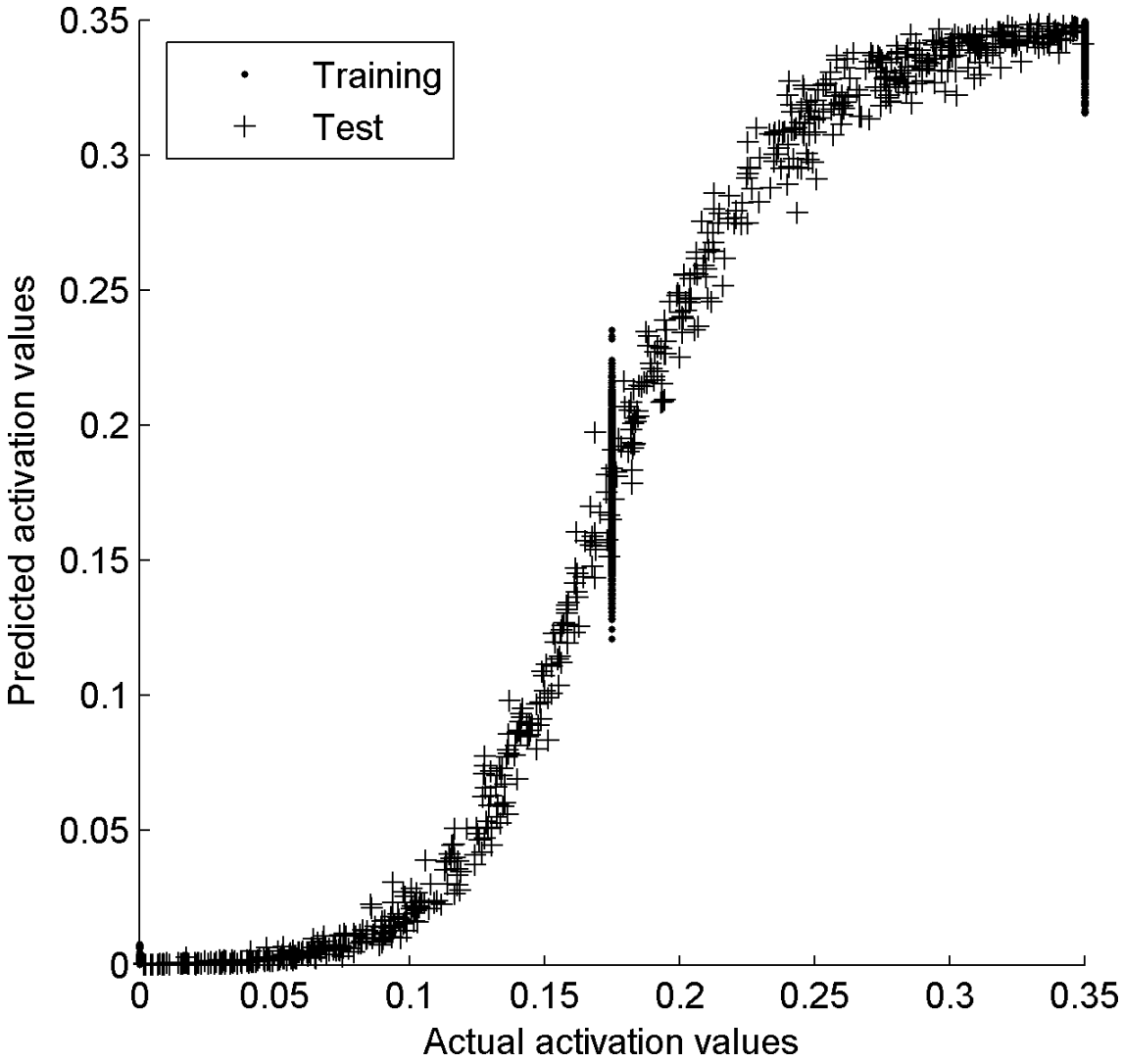

Figure 5.38: Network performance on testing and training data, mapping posture parameters to activation levels. Same interpretation as Figure 5.37.

though this is conceptually simple, it is important not to make inferences about the range of the sample based solely on minimum and maximum parameter values. The range of data values does not fill the parallelogram whose boundaries are determined by extreme values. This is illustrated in Figure 5.39 for two dimensions. Here the first two parameters of $D_2^{\mathbb{P}_1}$ are shown, with p_1 on the abscissa and p_2 on the ordinate (black dots). All of the points fall in a roughly diamond-shaped region. The dotted blue line shows the implied range obtained by specifying just the range of parameter values. It is implied for instance that $p_1 = 226.567$, $p_2 = 109.274$ is a possible shape of the model. This shape however, which is shown in Figure 5.40, is quite far from any of the data points in the actual sample. Knowing whether a point is within the range of the sample is important for knowing whether the posture associated with that point is posture that can be plausibly generated by the model. Particularly in machine learning applications, it is important not to test a model fit on data that are outside of the range of the training data.

A more conservative estimate of the range of the sample is obtained by computing the convex hull of the data set. In two dimensions, the convex hull is the smallest convex polygon that includes all of the points of the sample. A commonly used illustration is that of a rubber band being stretched around the data points. In Figure 5.39 the convex hull is shown with a solid red line. In three dimensions the convex hull is the polyhedron that just encompasses the points. The convex hull for the first three dimensions of $D_2^{\mathbb{P}_1}$ is shown in Figure 5.41. In higher dimensions the convex hull is a polytope, which is the equivalent of a polyhedron in three dimensions. The convex hull cannot be visualized in greater than three dimensions, but it is possible to tell whether a particular point falls within the hull.

The convex hull is a more conservative way to specify the range of a data set. It can be stipulated that if a point is within the range of the original data set only if it is within the convex hull defined by the data set.

This approach assumes that the set of tongue postures in $\mathbb{P}_1$ is convex. It may be true that this set is non-convex, being concave in places, or having gaps. For the present data, however, the assumption of convexity seems reasonable.

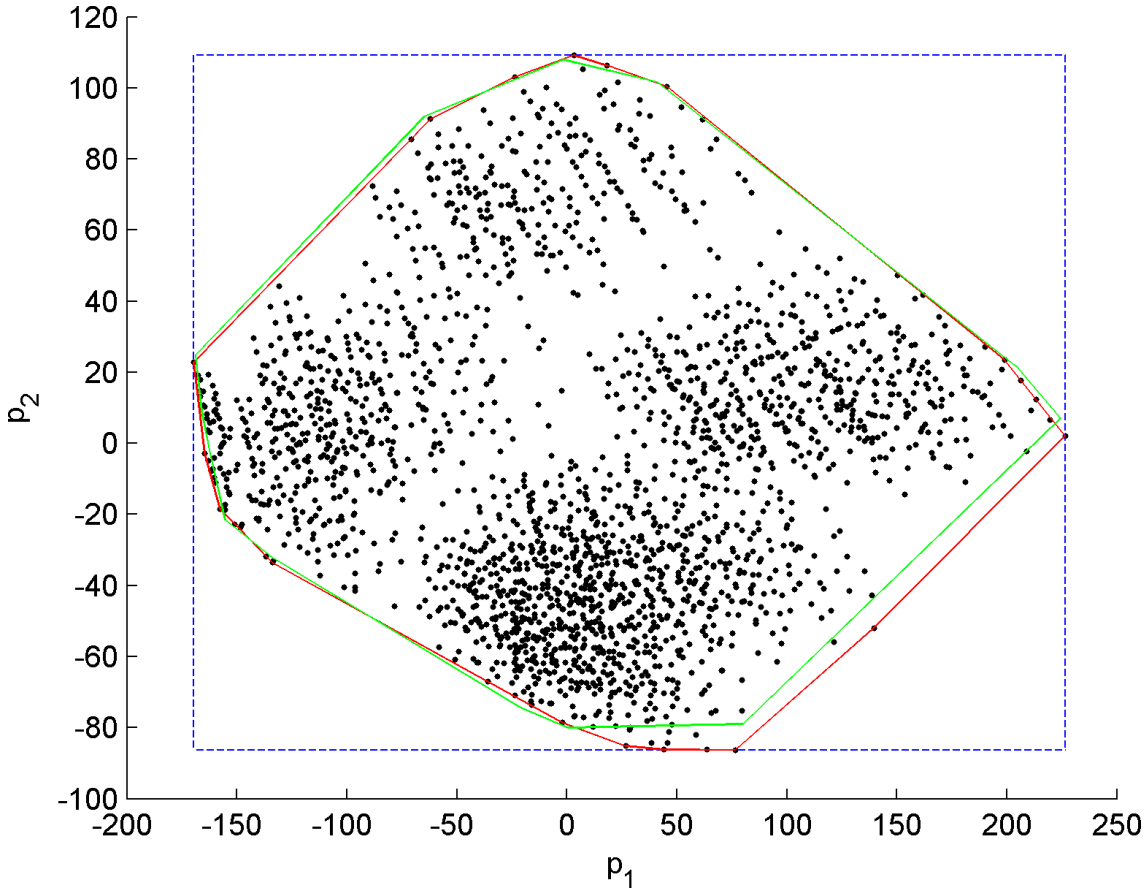

Figure 5.39: D_2 projected onto the first two parameters of $\mathbb{P}_1$. The range of the sample implied by specifying minimum and maximum parameter values is indicated with the dashed blue line. The more conservative range specified by the convex hull of $D_2^{\mathbb{P}_1}$ is shown with a solid red line. The convex hull of $D_1^{\mathbb{P}_1}$ is shown in green.

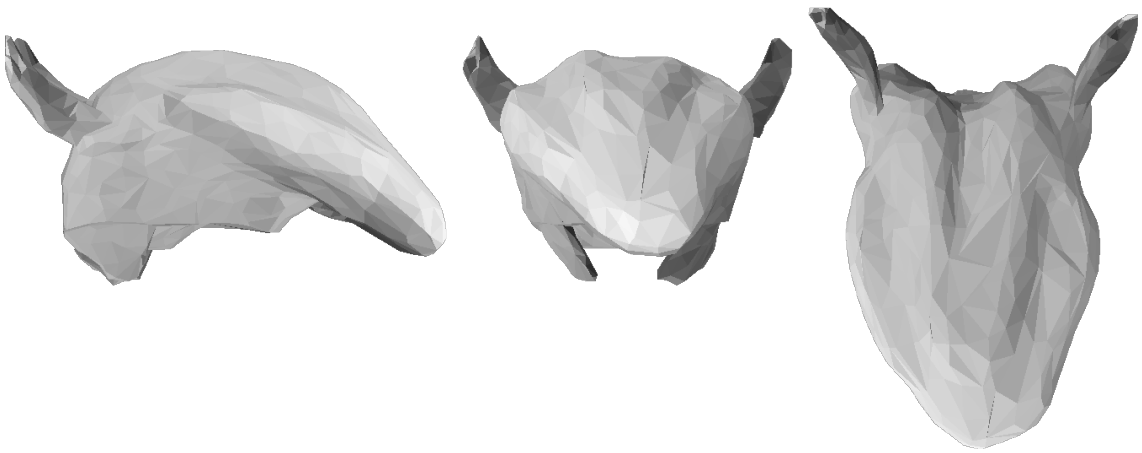

Figure 5.40: Not a valid tongue shape: $p_1 = 226.567$, $p_2 = 109.274$ (All other parameter values are those of the rest posture.)

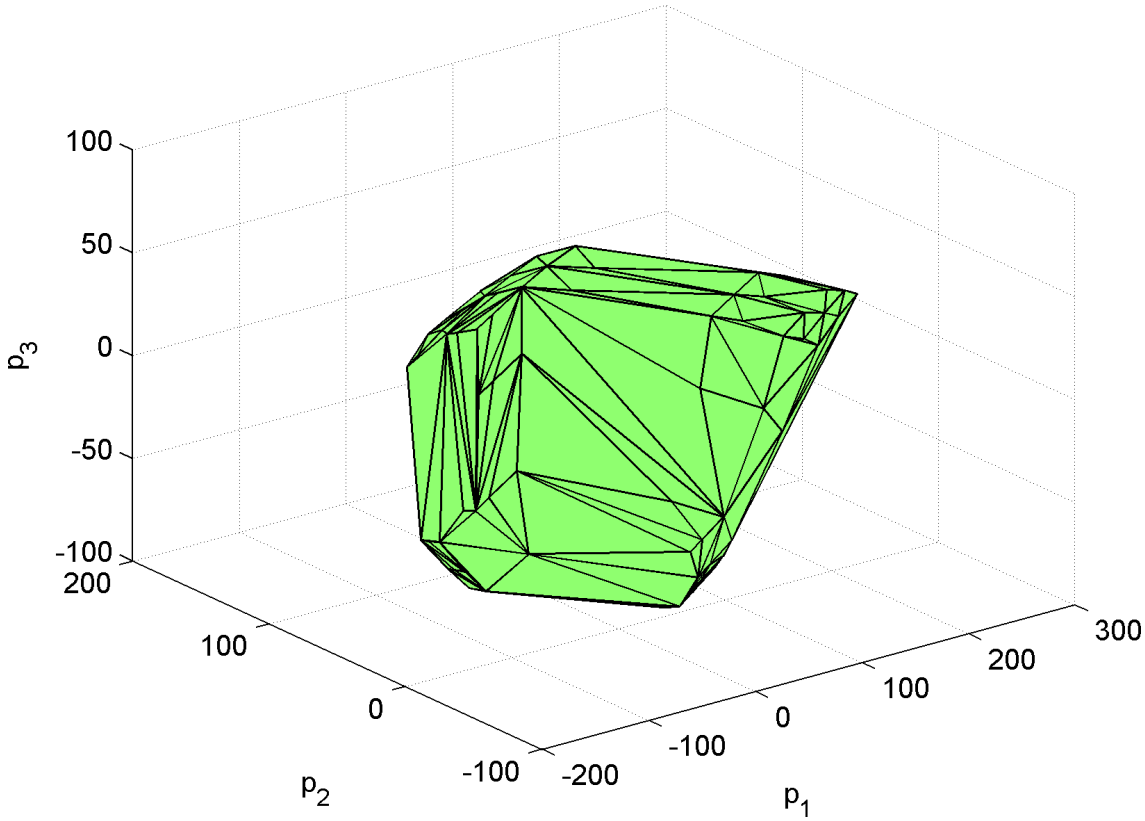

Figure 5.41: Convex hull of D_2 projected onto the first three parameters of $\mathbb{P}_1$.

In Figure 5.39 the convex hull for D_1 is shown in green. It is evident that the D_1 is a good predictor of the range of D_2 using convex hulls as well as parameter minima and maxima.

5.5 Antisymmetrical Tongue Movement

This section considers antisymmetrical tongue postures. It was observed in Section 5.1 that an exhaustive sample of antisymmetrical tongue shapes is prohibitively computationally expensive. Therefore such a sample is not collected here. Rather, an attempt is made to determine the parameters of antisymmetrical tongue movement from a more limited data sample. Since the sample size is limited, the number of tasks that can be performed are limited.

In determining this smaller sample, the fundamental observation is that the analyses in Section 5.2.3 and Section 5.2.5 yielded seven parameters of tongue movement and the range of values for each parameter within which tongue postures lay. In Section 5.2.5, for instance, 2,186 tongue postures were collected, but only the fourteen postures that occur in the periphery of the sample in $\mathbb{P}_1$ actually contribute to the result (the discussion of convex hulls notwithstanding). These fourteen postures indicate the extremes of tongue deformation, projected onto $\mathbb{P}_1$. The muscle activations associated with these extreme values are given in Table 5.7.

If the activation patterns in Table 5.7 are associated with the extrema of symmetrical tongue deformation, then it is reasonable to suppose that the same patterns of activation in the left or right side of the tongue would produce the extreme antisymmetrical postures. Under this assumption, a sufficient sample of the set of antisymmetrical tongue postures might involve all of the postures in which the left and right sides of the tongue are activated by some combination of the postures in Table 5.7. There are only $14^2 = 196$ such combinations, plus 28 postures more for the case in which one half of the tongue is active and the other is not active. A sample size of 224 is computationally tractable, and a fair reduction in the number of postures that would otherwise be required. This data set is called D_3.

		GG	HG	SG	TV	VR	SL	IL
Minima	p_1	0	0.35	0.35	0.35	0	0.35	0.35
	p_2	0.35	0	0	0.175	0	0.35	0
	p_3	0.35	0.35	0.35	0	0.35	0	0.35
	p_4	0	0	0.35	0.35	0.35	0.175	0
	p_5	0	0	0	0.35	0.35	0.35	0.35
	p_6	0.35	0.35	0	0.35	0.35	0	0.35
	p_7	0.35	0.35	0.35	0.35	0.35	0.35	0.35
Maxima	p_1	0.35	0	0	0.35	0.35	0	0
	p_2	0	0.35	0.35	0.35	0.35	0	0.35
	p_3	0	0	0	0.35	0	0.35	0
	p_4	0.35	0.35	0	0.35	0	0	0.35
	p_5	0.175	0.35	0.35	0	0	0	0
	p_6	0.35	0	0.35	0	0	0	0
	p_7	0	0	0	0	0.175	0	0

Table 5.7: Muscle activations associated with the peripheral elements of D_2 in $\mathbb{P}_1$.

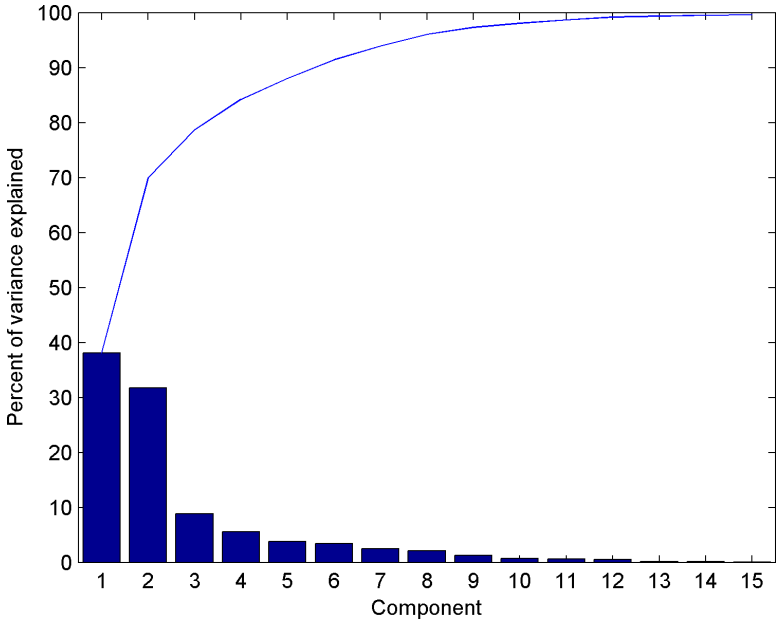

Figure 5.42: Pareto diagram of variance in D_3 accounted for by each parameter of $\mathbb{P}_3$.

A principal components analysis was performed on this antisymmetrical data set. The variance explained for each of the components is shown in Figure 5.42. As would be expected, more parameters are required to explain an antisymmetrical data set than a symmetrical one. In this case, 12 parameters are required to account for greater than 99% of the variance. These parameters are denoted $\mathbb{P}_3$. Some of the parameters have similar physical interpretations as those in $\mathbb{P}_1$ or $\mathbb{P}_2$, in that they correspond to symmetrical variation, while others correspond to asymmetrical variation.[15]

p_1 in $\mathbb{P}_3$ is similar to p_1 in $\mathbb{P}_1$; it corresponds to raising and lowering of the tongue tip. Variation along p_2 corresponds to a rocking of the tongue about the anteroposterior axis, and deflecting the tongue tip to the left or right. p_3 corresponds to p_2 in $\mathbb{P}_1$ (primarily, raising and lowering of the tongue tip). p_4 involves complex asymmetrical rotation, with the primary effect being a twisting of the tongue tip about the anteroposterior axis. p_5 corresponds to p_3 in $\mathbb{P}_1$ (raising and lowering of the tongue body). p_6 corresponds to depressing the left posterior part of the tongue. Variation along p_7 involves a sort of twisting of the tongue tip, and deflection of the tip to the right. p_8 corresponds to p_6, but for the right side (though the deformations appear smaller). p_9 involves flattening of the tongue, primarily on the left side. p_{10} is the contralateral equivalent to p_7. p_{11} is the contralateral equivalent of p_9. The effect of variation along p_{12} is difficult to describe.

5.6 Discussion

In analyzing the smaller data set D_1 (the set of postures resulting from all activation patterns at just two levels of activation), it was shown that the variance of the data did not need to be represented in a 2,880-dimensional space, but could instead be represented quite faithfully in a 7-parameter space ($\mathbb{P}_1$). The first several parameters had reasonably straightforward interpretations: tongue protrusion, retroflexion,

[15]The author's ability to describe tongue deformations, which was strained earlier in this chapter, is about to fall apart completely. The reader is referred again to the Tongue Viewer program.

dorsum bunching, and so forth. Further, it was shown that, once the data had been projected into the new space ($D_1^{\mathbb{P}_1}$), they could be regressed on the muscle activation levels using a simple linear model, with quite high adjusted r^2 values. Muscle activation could also been predicted fairly well from a given posture.

In analyzing D_2 (the set of postures resulting from all activation schema at three levels of activation), the results were essentially identical.

5.6.1 Interpretation of tongue components

In comparing and interpreting the results of different studies it is important to bear in mind the dataset upon which the analyses were based. Studies have been performed on the set of tongue shapes collected from a systematic sample of a biomechanical model's behavior (e.g., Perrier et al., 2000), the approach taken in the present work. Collecting a systematic and unbiased sample of tongue postures is of interest because the resultant factor analysis reflects the biomechanically natural movements of the tongue. "Natural" is intended to refer to behaviors that are congruent with an organism's anatomical and physiological disposition.[16] These can be discovered in part with factor analysis.

The sample of data analyzed in the present work is most similar to that in Perrier et al. (2000). The authors collected a systematic sample of tongue postures from the two-dimensional biomechanical tongue model developed by Payan and Perrier (1997, reviewed in Section 2.2.1). Figure 5.43 shows the effects of varying the first three parameters of movement from their data set, which together account for 94.2% of the variance in the data. Their first parameter corresponded to movement of the tongue toward the palato-alveolar region. The second parameter corresponded to movement toward the velum. The third parameter was a more subtle, seeming to correspond to an advancement and raising of the tongue.

In fact the results of Perrier et al. bear very little resemblance to those in the

[16]To take an example from another motor behavior, it is *possible* to walk backwards, but it seems reasonably clear that our lower limbs are optimized for forward locomotion. Likewise there are bound to be certain tongue postures for which the tongue is optimized.

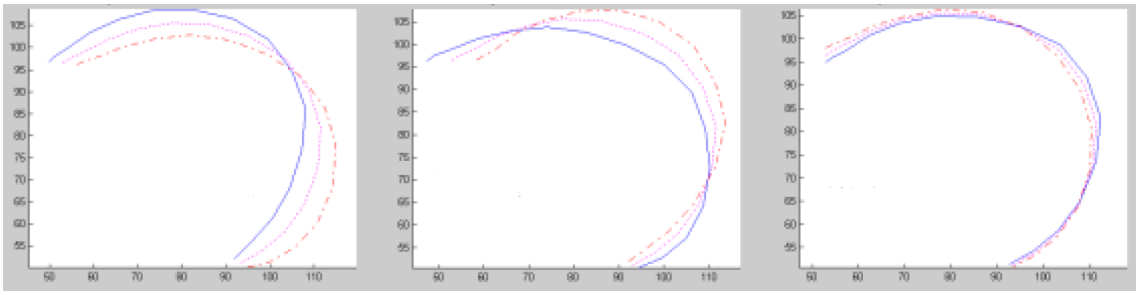

Figure 5.43: Effect of parameter variation for the first three parameters from Perrier et al. (2000). (Figure adapted from Perrier et al. 2000.)

present work. Compared to the results of the present model, Perrier et al. report a very limited range of tongue motion. For instance, in their sample the tongue tip moves relatively little (it is always angled downward), whereas in the present model movement of the tongue tip dominated the factor analysis. The difference in the results is probable attributable to the fact that the Payan and Perrier (1997) model is two-dimensional. Limitations of two-dimensional models in general are discussed in Section 2.2.3. Although firm conclusions about this specific model cannot be drawn, it seems likely that the range of vertical motion was limited by the absence of transversus (cf. the importance of transversus for the present model, described in Section 6.1). It may also be that, in a two-dimensional model, superior longitudinalis is smaller relative to the size of the tongue, which could account for why the tip specifically seems relatively less mobile in the Payan and Perrier (1997) model.

Studies have also been been performed on vocal tract postures from phonetic segments, typically a vowel space (e.g., Harshman et al., 1977; Story and Titze, 1998; Engwall, 2003, ; also Sanguineti et al. 1998 for such a sample drawn from a computational model). Here the goal is to determine the parameters of tongue movement pertinent to a specific task, which for the papers cited was speech production.

The two types of studies—analyzing systematic samples of model behavior, and samples from a specific behavior—are complementary. Understanding the biomechanically natural parameters of movement is relevant to, but distinct from, the in-

vestigation of the parameters of a specific behavior. It is of interest to know whether the parameters of movement for a particular behavior coincide with the biomechanically natural paremeters of tongue movement. One cannot know the parameters of movement pertinent to a particular behavior, however, without analyzing data drawn from that behavior.

Since the language faculty in particular is believed by some to be the result of a language-specific innate endowment, there is a temptation to look for language-specific parameters of tongue behavior. No speculation about innateness will be made here, for the following reasons. First, as the exposition below will note, many of the parameters that can be observed have applications both in language and in survival-related activities, so that claiming the existence of a language-specific adaptation would be very difficult to justify. Second, there is no corresponding animal study, to study whether the characteristic parameters human tongue posture are different from those of any other animal. Third, even if such a comparison were available, any differences would be uninterpretable. Humans differ from primates not just in their ability to speak, but also in the shape of the upper airway, i.e., the descended larynx. It is not unreasonable to expect special tongue adaptations to cope with this riskier configuration; there does not seem to be any way of differentiating adaptations for this purpose and hypothesized adaptations for speech, barring the development of vastly superior anatomical and physiological knowledge than we now possess. Consequently, the remarks will be limited to observing correlations between tongue parameters and specific behaviors.

Variation along p_1 corresponds to protrusion and retrusion of the tongue. Tongue protrusion is only rarely a feature of speech production (an exception is the Kagayanen interdental approximant, in which the tongue can protrude from the lips; Olson et al. forthcoming). The movement is reminiscent of lapping behavior. Humans do not lap liquids for survival purposes, though other mammals do, so it is not very surprising to find that lapping is a "natural" behavior for a mammalian tongue. Variation along p_1 advances and retracts the tongue root. This is a crucial movement for maintaining a patent airway (e.g., Bailey and Fregosi, 2004).

Movement on p_2 produces tongue tip elevation and depression. Vertical tongue tip movement is associated with a number of behaviors. Probably the most biologically crucial of these is suckling, in which the mother's nipple is squeezed between the anterior portion of the tongue and the hard palate (Smith et al., 1985; Hayashi et al., 1997). Coronal speech sounds make use of the tongue tip. Further, p_1 and p_2 can be varied to retroflex the tongue tip. The aforementioned lapping behavior also requires tongue tip movement. Additionally, as with p_1, p_2 involves movement of the tongue root, which is important for certain consonant and vowel distinctions.

p_3 and p_4 have mostly to do with bunching the tongue, which is an important part of both deglutition and vowel production.

5.6.2 Generalization from smaller to larger datasets

The results in Sections 5.2.5 and 5.3 show that the component analysis and regression model from the smaller data set generalizes quite well the larger data set. These results are significant because they indicate small sets of tongue postures can be used to generalize successfully to much larger ones.

From a developmental perspective, it becomes much easier to imagine how an infant or child can "solve" the tongue motor control "problem." As complex as the tongue is, its mechanical properties are such that its posture can be represented with a fairly small number of parameters. Moreover, these parameters can be calculated from a reasonably small data set. Again, the claim is not that the central nervous system is performing the analyses given in this chapter, but rather that *some* sort of simple representation is available to it. The regression results indicate that, based on a small a sample, the tongue posture can be predicted fairly well from a muscle activation pattern (and vice versa). This generalization was not as robust as that from the principal components analysis, but they do provide a good estimate.

From a research standpoint, the results indicate that sampling the behavior of an organ as complex as the tongue is not a fool's errand. In fact, since much of the variance in the three-level (2,186-sample) data set is predictable from the two-level (127-sample) data set, researchers are justified in sampling tongue behavior with

more control parameters, at just two activation levels. If one were to take a sample of antisymmetrical whole muscle contractions, for instance, one could collect 16,384 samples instead of the 4,782,969 that would be required for a three-level sample—a considerable computational savings.

The regression models have broad application. They can be used for exploring the muscle activations associated with particular tongue postures, for instance. They could also form a component of a more complex vocal tract model. A particularly apt application would be in testing theories of motor learning. One could imagine, for instance, an iterative motor learning model, for which a 4–6 hour delay between model increments would be prohibitive. Calculating a posture from a linear model or neural network is much faster.

CHAPTER 6

Particular Results

> Moses said to the LORD, "O Lord, I have never been eloquent, neither in the past nor since you have spoken to your servant. I am slow of speech and tongue."
>
> —Exodus 4:10 (NIV)

This chapter consists of an assortment of results, including tests of simple muscle contraction, tests of predictions made in the literature, and a model of flaccid dysarthria. In Section 6.1 images are presented that illustrate the result of activating each muscle of the tongue, which are the stock in trade results of any biomechanical tongue model. This section also contains a discussion of styloglossus (Section 6.1.1), which has a much smaller effect on tongue posture in the present model than was expected to be the case. A comparison of the results of the present model with other models is made in Section 6.1.2. Section 6.1.3 investigates the role of coactivation of muscles in producing the full range of tongue postures, by performing a principal component analysis on the set of tongue postures produced by contracting just one muscle. Some specific predictions from the literature are tested, from Abd-El-Malek (1955) (Section 6.2.1), Kier and Smith (1985) (Section 6.2.2), and Halle (1983) (Section 6.2.3). Section 6.2.4 contains an investigation of a possible functional division between the intrinsic and extrinsic tongue muscles, which did not discover any principled mechanical division between these muscles. Finally, Section 6.3 has a model of flaccid dysarthria, in which half of the tongue receives no motor input. It is determined that different muscle activation patterns are required for a hemiplegic tongue than for a normal tongue, to achieve comparable tongue postures.

6.1 Consequences of whole muscle contraction

In this section the small subset of tongue postures in which only one muscle is active is considered. These results are more or less in line with what one would expect from looking at the anatomy of the tongue, with some exceptions, which are discussed below. The figures in this section can be compared to Figure 5.4, which shows the tongue in its undeformed state. It turns out that one of the surprises of a quantitative tongue model is how uninformative it is to look at postures in which only one muscle is active, and other muscles are set to produce no contractile force. This observation is considered quantitatively in Section 6.1.3.

The action of genioglossus is to pull the tongue forward and downward (Figure 6.1). Hyoglossus, shown in Figure 6.2, depresses the posterior of the tongue. This has the effect of raising the tongue tip slightly, and pulling the tongue backward slightly as well. Styloglossus pulls the body of the tongue up slightly, but only in the immediate area of its insertion (Figure 6.3). There is practically no effect on the tongue tip. This muscle had a much smaller effect on the shape of the tongue than was expected (cf. Chapter 2); this is further discussed in Section 6.1.1.

Contraction of transversus, shown in Figure 6.4, narrows the tongue body in the horizontal direction, expands it vertically, and produces a slight protrusion of the tongue tip. The action of verticalis is mainly to flatten the body of the tongue (Figure 6.5). It also depresses the tongue tip quite a bit, a motion previously not associated with verticalis in the literature. Superior longitudinalis, shown in Figure 6.6, raises the tongue tip, as one would expect. Finally, inferior longitudinalis shortens the body of the tongue, somewhat expanding the tongue vertically, and pulling the tongue tip backward and slightly downward (Figure 6.7).

These results are compared to those of other modeling studies below in Section 6.1.2.

Asymmetrical contractions of the tongue were also modeled, in which only the left or right side of the tongue was active. The results are the expected unilateral analogs of the postures that involve bilateral contractions. Figures 6.1 through

6.7 are paralleled in Figures 6.8 through 6.14 for right-hand contraction, and Figures 6.15 through 6.8 for left-hand contraction.

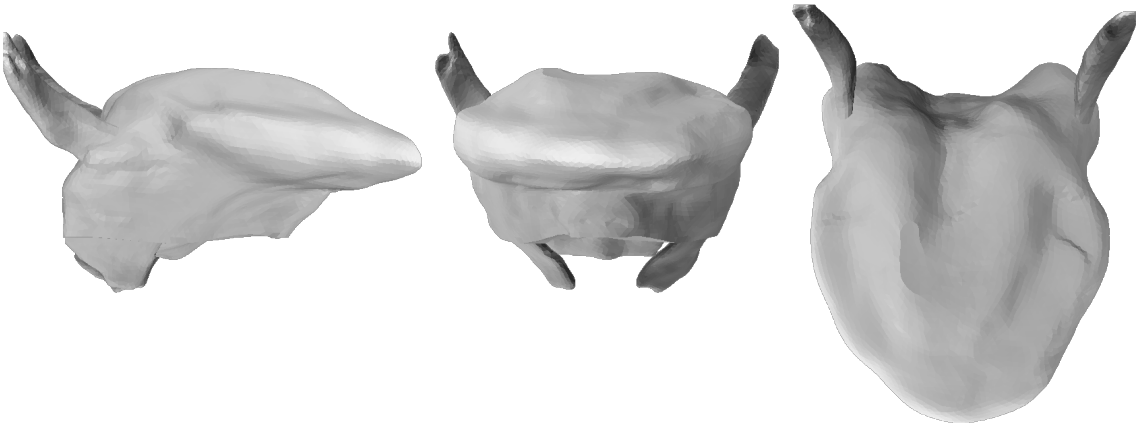

Figure 6.1: Symmetric full strength contraction of genioglossus

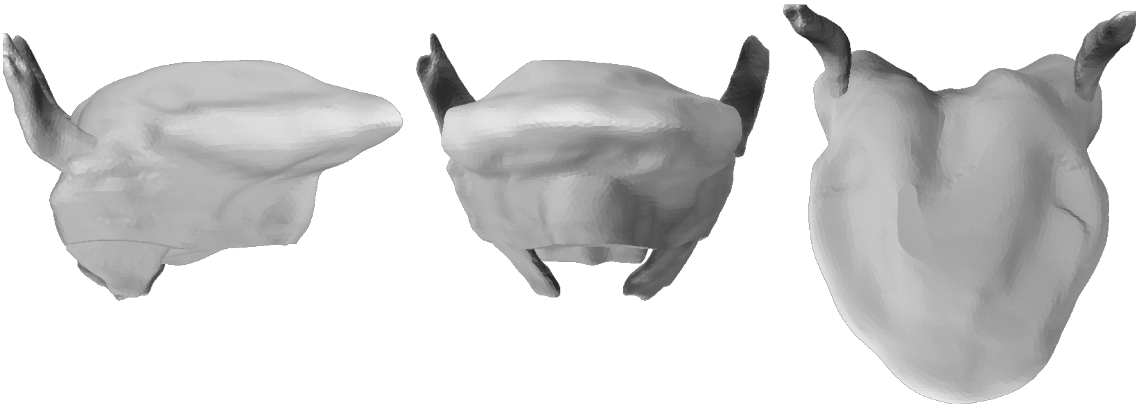

Figure 6.2: Symmetric full strength contraction of hyoglossus

6.1.1 Styloglossus

Styloglossus is curious in that, in the present model, it seems to have very little effect on the posture of the tongue. This muscle is frequently assumed to have a significant role in elevating and retracting the tongue, however (e.g., Abd-El-Malek, 1955; Hardcastle, 1976; Halle, 1983; Zemlin, 1988), a purpose that it seems to serve in lower animals (e.g., the rat, McClung and Goldberg, 2000). Additionally,

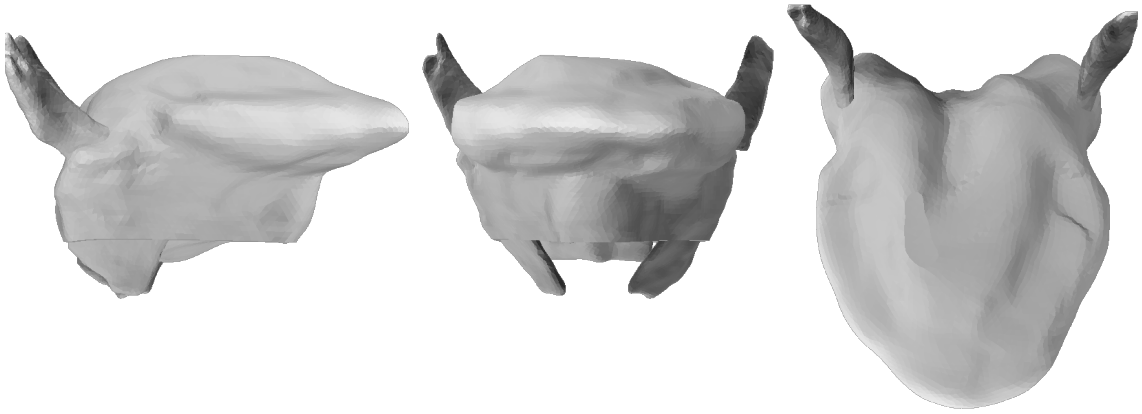

Figure 6.3: Symmetric full strength contraction of styloglossus

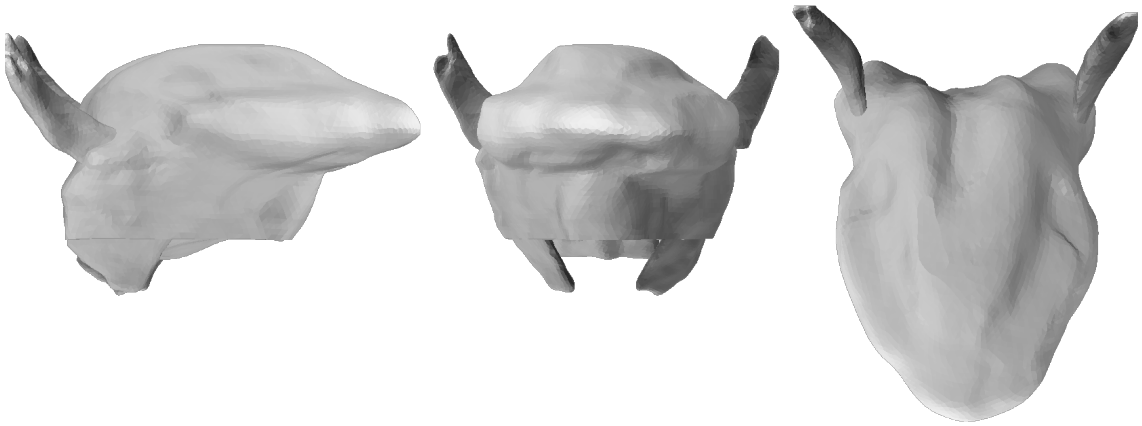

Figure 6.4: Symmetric full strength contraction of transversus

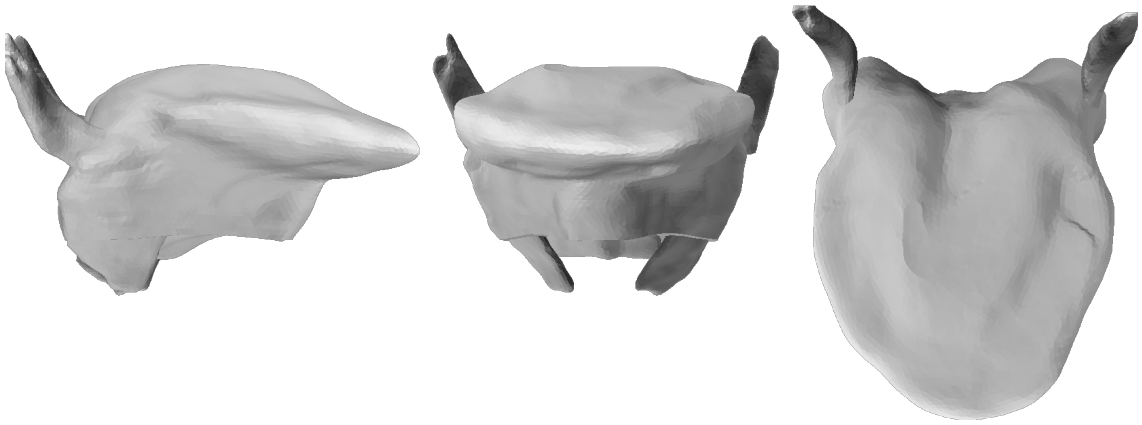

Figure 6.5: Symmetric full strength contraction of verticalis

Figure 6.6: Symmetric full strength contraction of superior longitudinalis

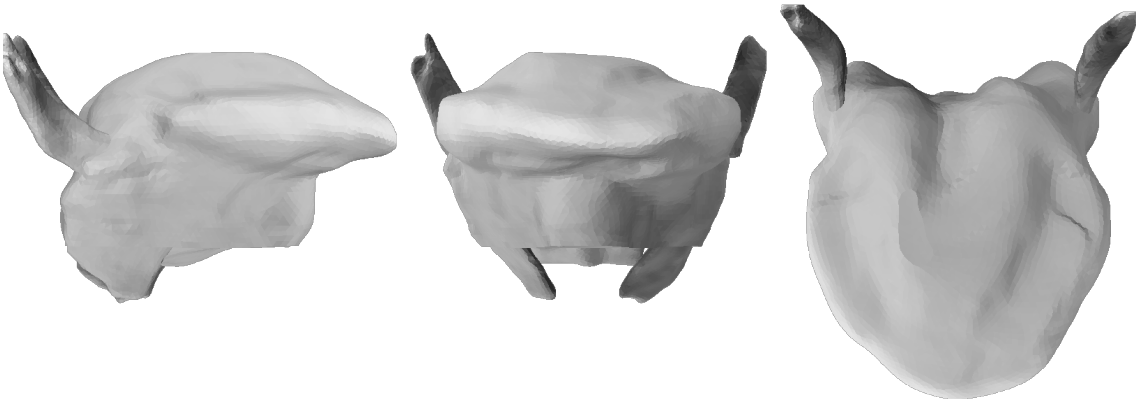

Figure 6.7: Symmetric full strength contraction of inferior longitudinalis

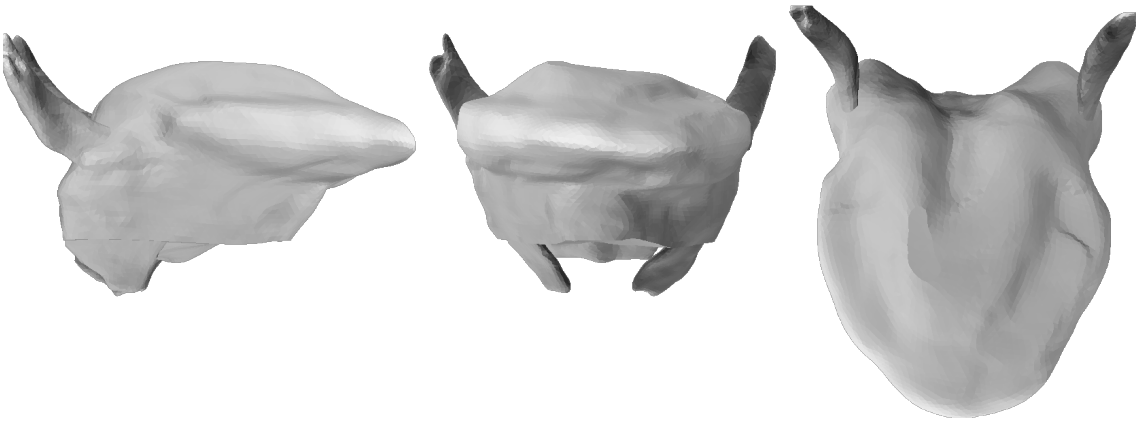

Figure 6.8: Right-hand full strength contraction of genioglossus

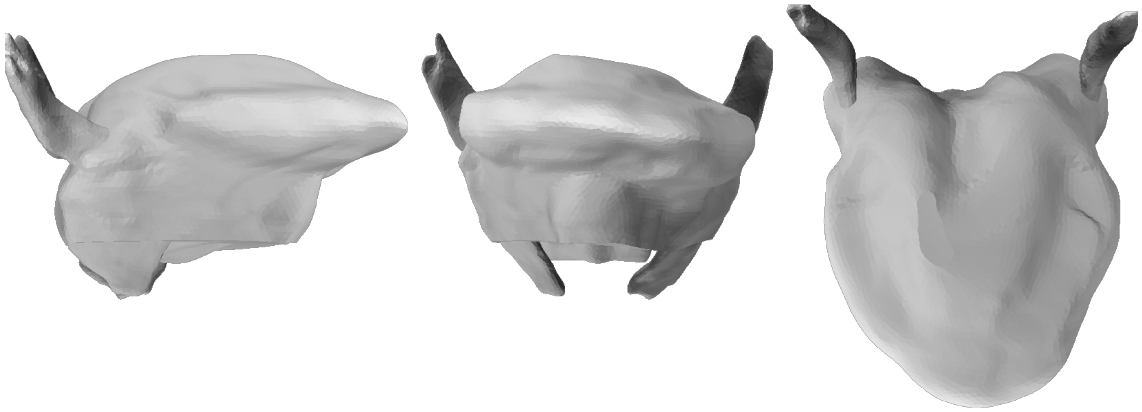

Figure 6.9: Right-hand full strength contraction of hyoglossus

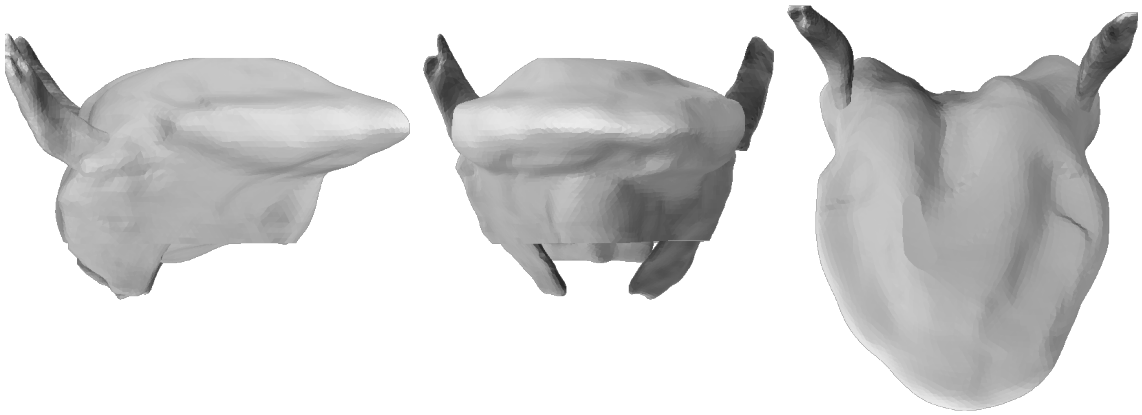

Figure 6.10: Right-hand full strength contraction of styloglossus

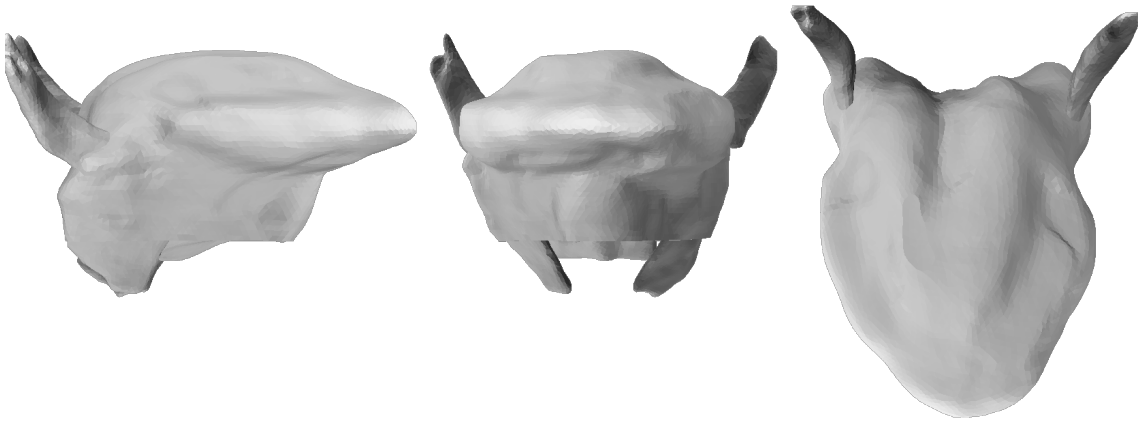

Figure 6.11: Right-hand full strength contraction of transversus

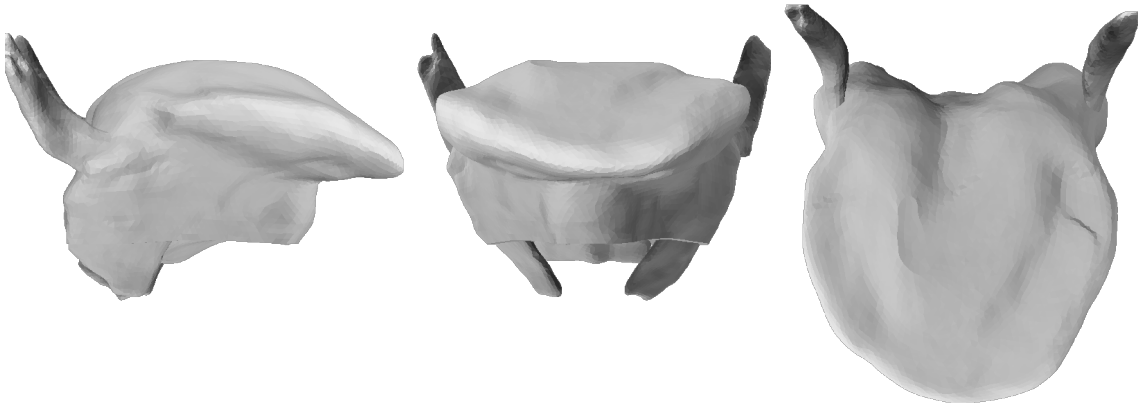

Figure 6.12: Right-hand full strength contraction of verticalis

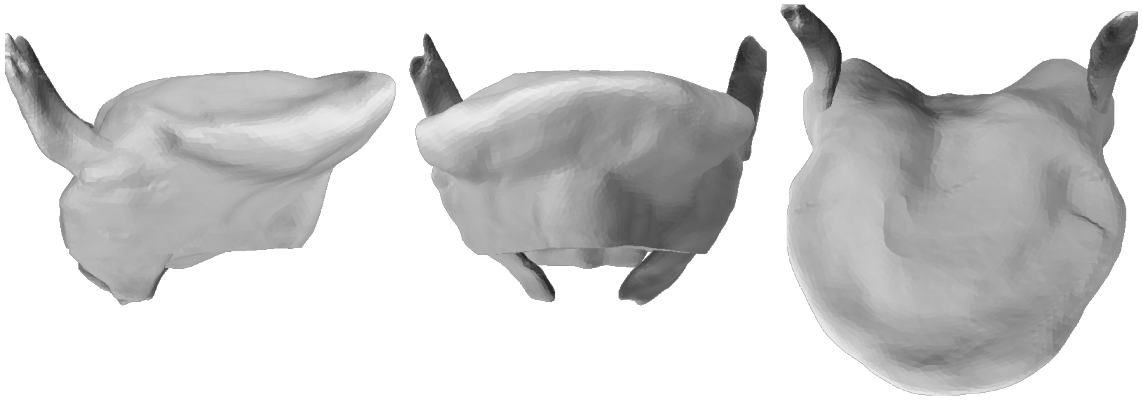

Figure 6.13: Right-hand full strength contraction of superior longitudinalis

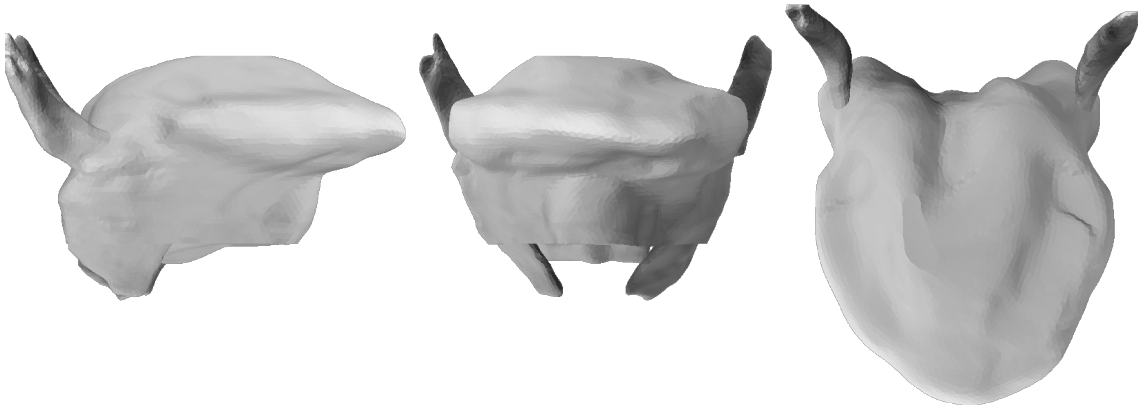

Figure 6.14: Right-hand full strength contraction of inferior longitudinalis

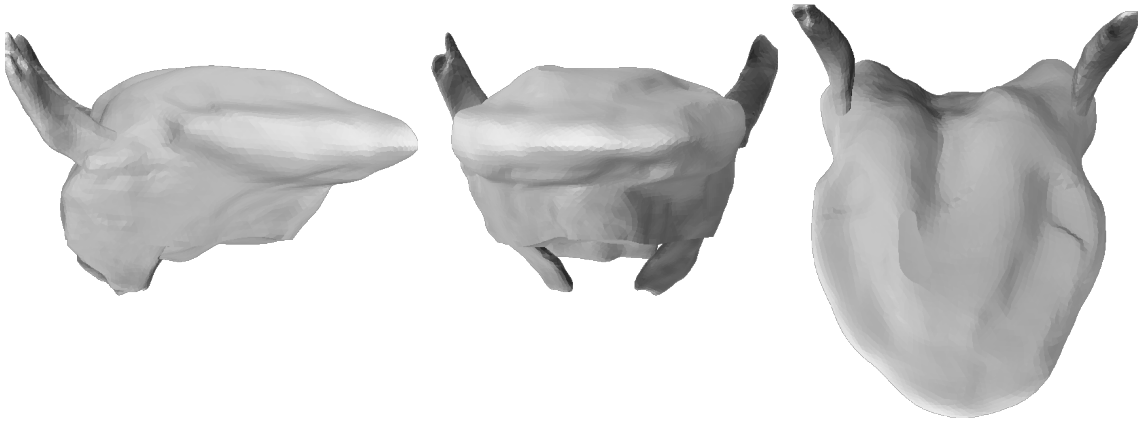

Figure 6.15: Left-hand full strength contraction of genioglossus

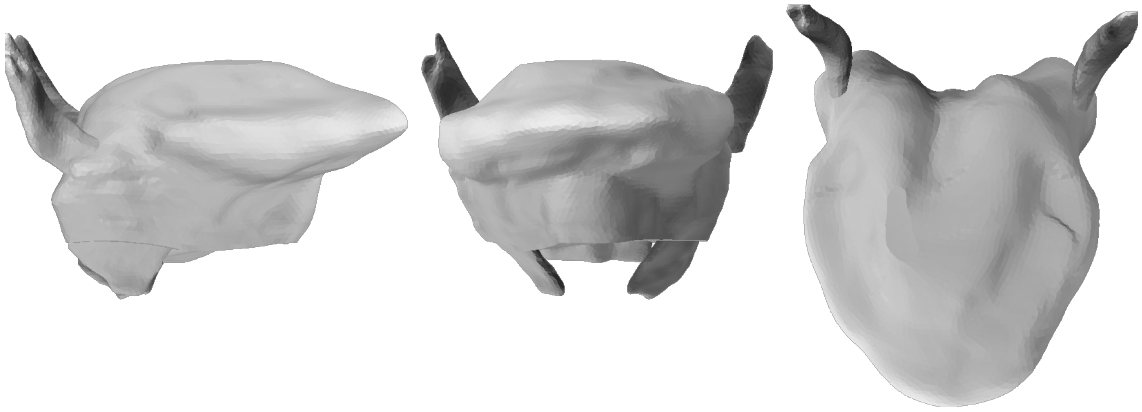

Figure 6.16: Left-hand full strength contraction of hyoglossus

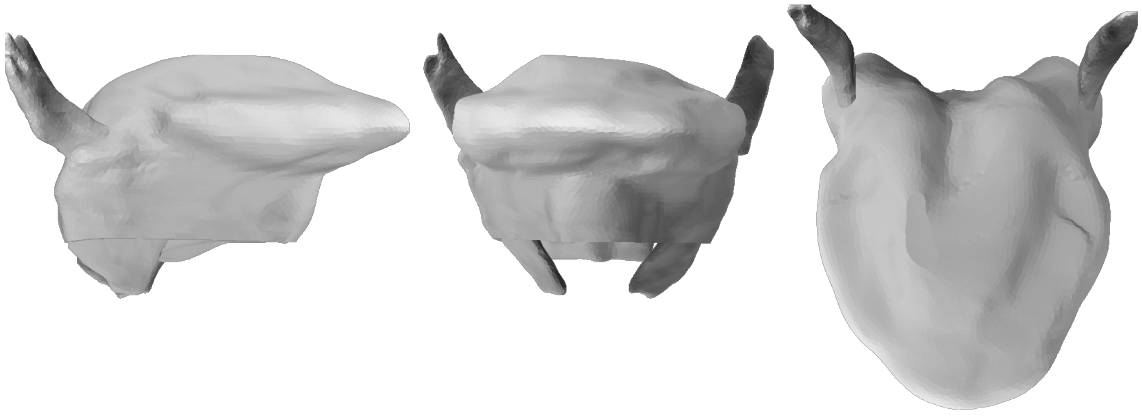

Figure 6.17: Left-hand full strength contraction of styloglossus

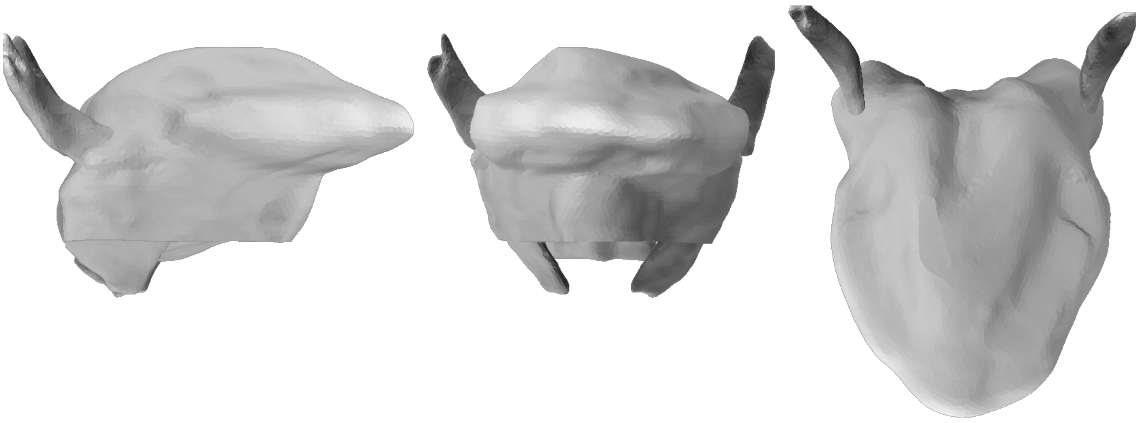

Figure 6.18: Left-hand full strength contraction of transversus

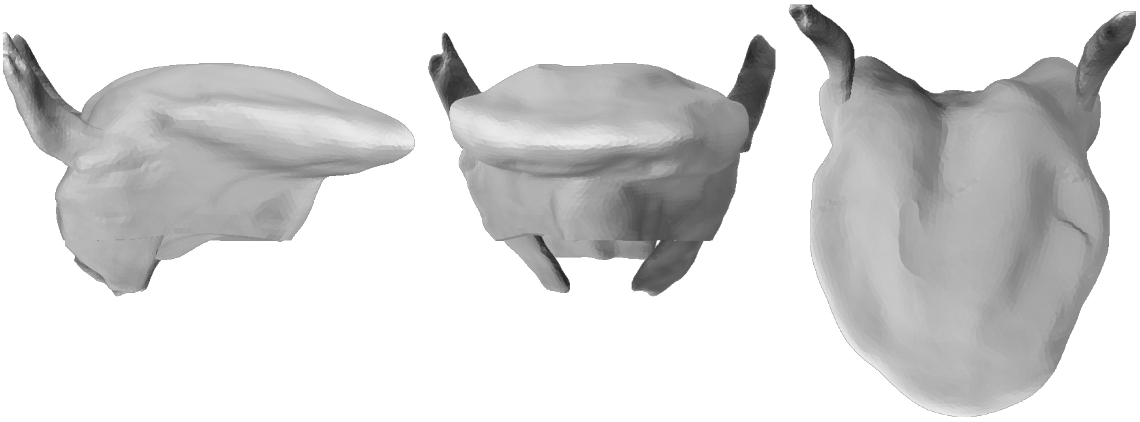

Figure 6.19: Left-hand full strength contraction of verticalis

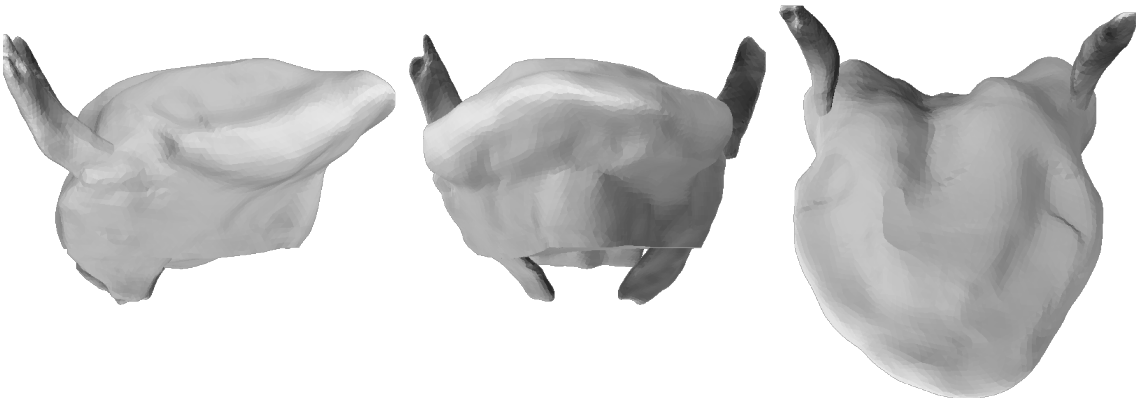

Figure 6.20: Left-hand full strength contraction of superior longitudinalis

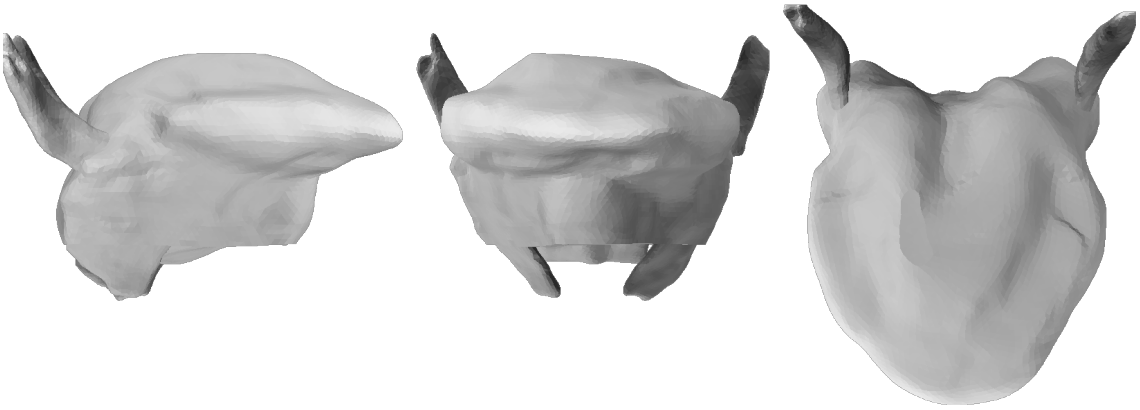

Figure 6.21: Left-hand full strength contraction of inferior longitudinalis

styloglossus is active in humans during the production of back, and especially high back, vowels (Baer et al., 1988). The absence of much movement in the model with full contraction of styloglossus is therefore surprising. Three possible explanations for this discrepancy between previous research and the present model are considered below.

It may be that inadequacies of the model in specifying the geometry of styloglossus significantly limit the effect of styloglossus activation. The geometry of styloglossus in the model differs from the anatomical reality in two ways. First, near its insertion on the styloid process, human styloglossus is surrounded by soft tissue. In the model, this tissue is not included. It is possible that this soft tissue directs the force generated by styloglossus in functionally a significant way. Second, as was noted in Section 3.1.0.3, several other anatomical studies trace the fibers of styloglossus to the tongue tip anteriorly, and across the dorsum posteriorly (shown schematically in Figure 3.8). The anatomical study on which the model was based did not detect these fibers, and represented styloglossus as terminating approximately in the place that it meets hyoglossus (see Figure 3.16). That had the consequence of assigning to other muscles (most likely, inferior longitudinalis) voxels that rightly belonged to styloglossus. This would have the effect of diminishing the apparent effect of styloglossus.

It is also possible that styloglossus generates more stress than 0.35 MPa. Although Zajac (1989) gives 0.35 MPa as a typical contraction strength for human skeletal muscle, Martins et al. (1998) note that values ranging from 0.16 MPa to 1 MPa have been reported in the literature. If human styloglossus generates for than 0.35 MPa, then the current model would underestimate its influence.

A final reason that styloglossus does appear to be very significant in this model is that the rest position of the tongue model—the cadaveric posture—was not one in which activation of styloglossus would be expected to contribute much to tongue shape anyway. In the specimen used (see Figure 3.10), the tongue was already high and back in the roof of the mouth. Since the anticipated function of styloglossus is to pull the tongue superiorly and posteriorly, it might be expected that activation of styloglossus would have little further effect.[1] Styloglossus would then be expected to have a greater effect if the tongue were lower or further forward in the mouth.

To test the last two ideas broadly, a series of simulations was run in which the position of the jaw and hyoid and the contraction magnitude of styloglossus were varied. The simulations had two phases. First the tongue was displaced, and then styloglossus was activated. For the displacements, the nodes of the finite element mesh that attach to the jaw and hyoid were given a specified displacement. This displacement was either 5 mm anterior and 5 mm inferior (total displacement: 7.07 mm), or 10 mm anterior and 10 mm inferior (total displacement: 14.14 mm). Contraction magnitude was either 0.35 MPa or 0.7 MPa. These values are not principled selections, but rather were chosen for convenience.

Figure 6.22 shows the results of the simulations. Displacement magnitude varies by row; contraction magnitude varies by column. The first row shows the effect of styloglossus contraction *without* displacement of the jaw and hyoid. The left posture is the rest position; the middle posture corresponds to Figure 6.3. Even in the right posture, in which contraction magnitude is doubled, there is little effect

[1] By way of illustration, if one constructed a model of the upper limb in which the rest position of the model involved a fully flexed elbow, biceps brachii would not appear to be a very significant muscle.

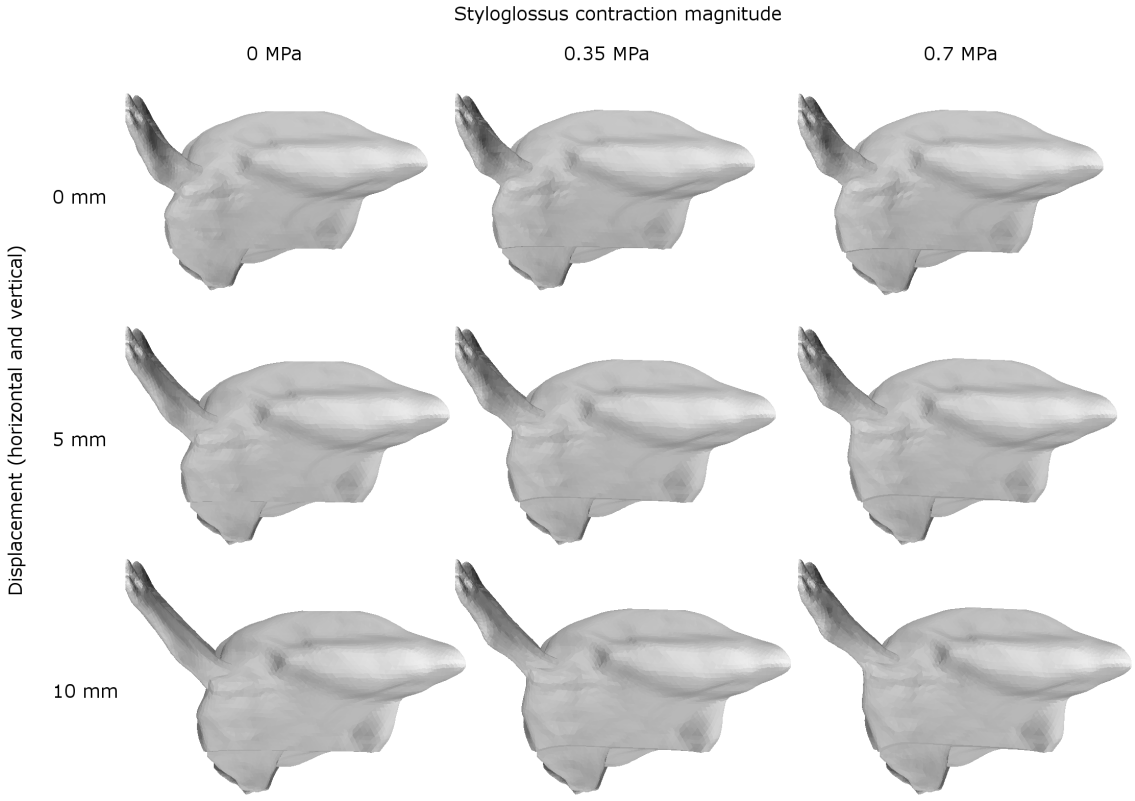

Figure 6.22: Effect of styloglossus under varying contraction magitude (columns), and displacement of the jaw and hyoid attachments (rows).

on the posture of the tongue. The second and third rows show postures for the corresponding activations after the 5 mm and 10 mm displacements, respectively. It is evident that, even for the 0.35 MPa activation, styloglossus has a greater effect if the tongue is displaced anteriorly and posteriorly. Activation of styloglossus had a greater effect after the 10 mm displacement than after the 5 mm displacement. There is also an increased displacement with the greater contraction magnitude (i.e., comparing the second and third columns).

To summarize, in the present model contraction of styloglossus did have much effect on the posture of the tongue. Three reasons were presented to suggest why this result should obtain. First, the model was based on an anatomical representation of styloglossus that was incomplete. Second, it is possible that styloglossus generates more than 0.35 MPa in active stress. Third, it was proposed that the rest position of the tongue restricted the effect of activating styloglossus. The second and third proposals were tested. Increased contraction magnitude in styloglossus did produce increased displacement, as would be expected. (It is not known, however, whether styloglossus actually generates more than 0.35 MPa of active stress.) Displacing the tongue body anteriorly and posteriorly also increased the effect of activating styloglossus. The displacements used in this chapter were selected arbitrarily; it is likely that other displacements would have a greater effect. Since the jaw and hyoid are mobile during speech production, it is likely that the position of the tongue body is an important factor in determining the extent to which styloglossus deforms the tongue. Subsequent studies of model behavior with systematic variation of jaw and hyoid position may further elucidate this relationship.

6.1.2 Comparison to previous models

The results presented here are broadly similar to those reported in other modeling studies. The results are compared with three studies here, namely the studies cited in Section 2.2 that illustrated the consequences of single-muscle contractions. Comparisons can be made when the control parameters of other studies match those of the current study. For instance, each of the other studies matches the current one

in controlling superior longitudinalis with a single parameter, which enables comparisons to be made; conversely, each controls genioglossus with several parameters, which was not done here.

In general, the differences are in degree, not in kind. This is to be expected. The models were grossly similar in geometry, so that the general function of muscles should be preserved. The models differed, however, in mesh density, material properties, muscle definitions, etc. It is also notable that, in the three works described below, the authors use different stimulation levels for different muscles.[2] The relative effects of different muscles in the same model are therefore not directly comparable, as they are for the results reported for the present work. Comparisons between these models and the present model are also fairly limited for this reason.

In the two-dimensional model reported by Perkell (1974), superior longitudinalis had a similarly large effect as in the present study. The action of inferior longitudinalis was also similar. Hyoglossus had a somewhat larger effect than it did in the current study. The effect of styloglossus was much larger, producing a full-fledged [u]-like posture.

The results of the two-dimensional model from Payan and Perrier (1997) are generally similar. Styloglossus has a larger role than would be expected from the current results, while superior longitudinalis has a smaller effect. Hyoglossus and inferior longitudinalis have similar effects. Verticalis has almost no effect in their model, however, which is quite different from the present model.

Gerárd et al. (2003) report that contracting styloglossus in their model produces greater retraction than it does in the current one; their styloglossus extends further anteriorly than it does in the present model, though not to the tongue tip. The effect of contracting superior longitudinalis is less pronounced in their model as well. Their simulation of contraction of hyoglossus was similar to the present one.

[2]None of the authors provide reasons for this decision, except for Perkell (1974), who indicates that it was necessary to alter model parameters given his solution procedure. One possible motivation for individually varying activation levels would be to achieve tongue postures that "looked right."

6.1.3 Role of muscle synergies in achieving full range of motion

It is easily notice that the tongue postures associated with the contraction of a single muscle (Figures 6.1 through 6.7) show a very limited range of motion in comparison to that found in the D_1 and D_2 datasets, which include all combinations of muscle activation (respectively, Figures 5.6 through 5.19 and Figures 5.21 through 5.34). Figures 6.23 through 6.34 show the results of a principal component analysis of performed on the postures that involved contraction of just one muscle.[3] All of the variance could be accounted for with just six parameters. The physical correlates of movement along each parameter are essentially the same as in $\mathbb{P}_1$, except that p_6 more closely resembles p_7 in $\mathbb{P}_1$ than p_6. For each component—most easily noticeable for p_1, p_2, and p_3—the range of the tongue is much more limited.

Figure 6.23: PCA of single muscle data: p_1 at minimum value.

Figure 6.35 offers another way to appreciate this limitation of range. Here the data with single muscle contractions have been projected onto $\mathbb{P}_1$. The blue circles indicate the minimum and maximum values of the D_2 data. The red dots indicate the range of parameters for just those samples in which one muscle is contracted. The single muscle contraction postures have a range of 34%–62% of the data with all combinations of muscles. This indicates the importance of the coactivation of tongue muscles in producing the full range of motion.

[3]Note that this data set is a subset of D_1, which is a subset of D_2.

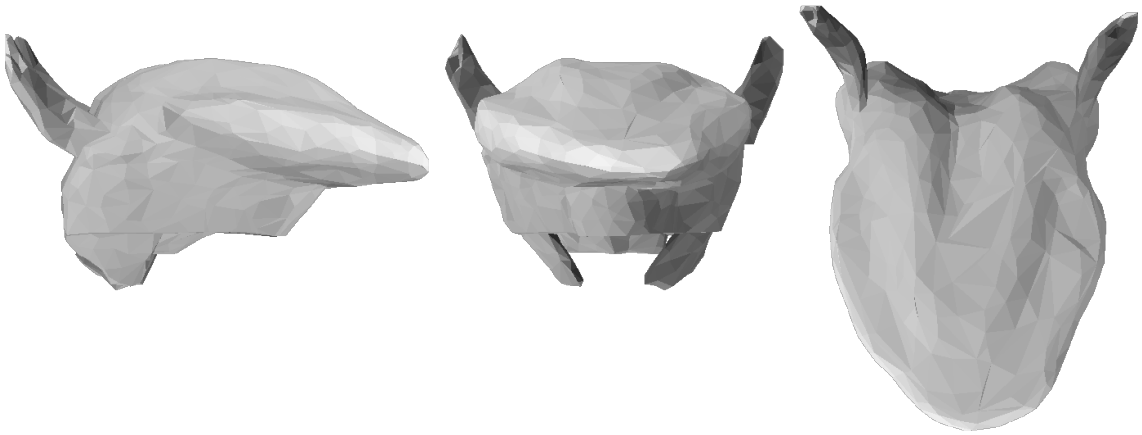

Figure 6.24: PCA of single muscle data: p_1 at maximum value.

Figure 6.25: PCA of single muscle data: p_2 at minimum value.

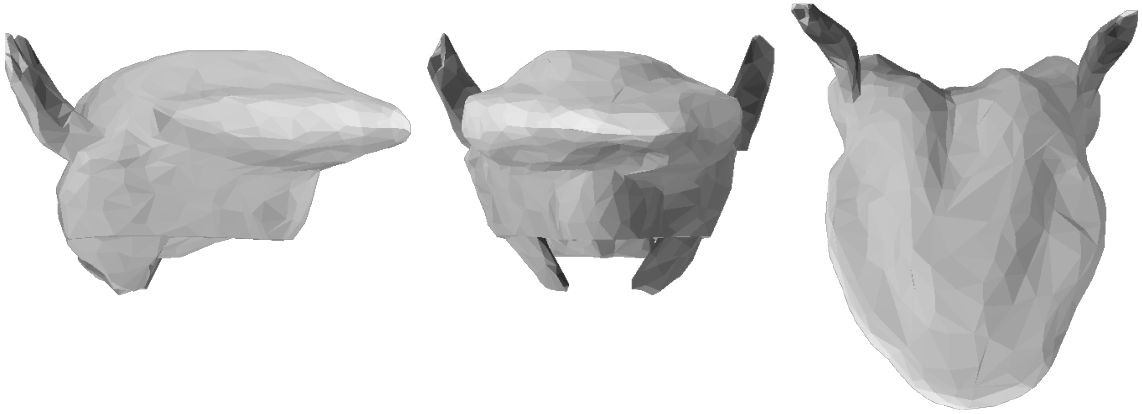

Figure 6.26: PCA of single muscle data: p_2 at maximum value.

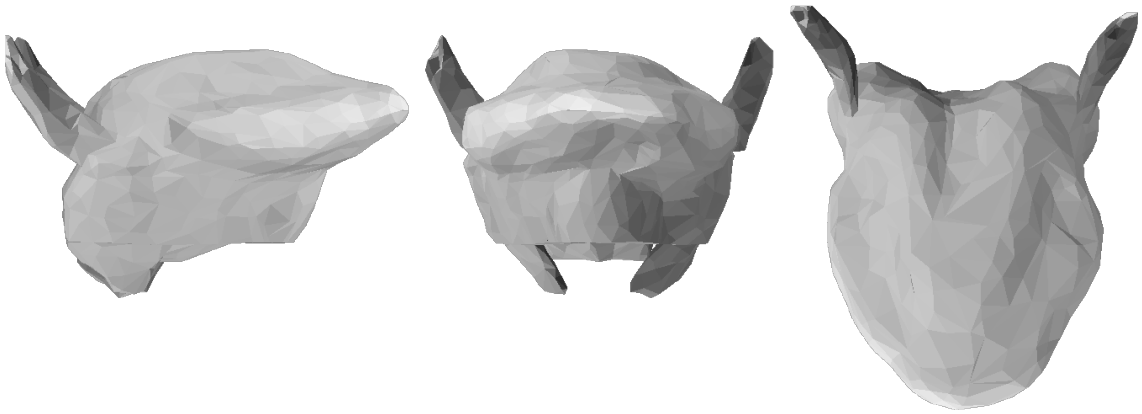

Figure 6.27: PCA of single muscle data: p_3 at minimum value.

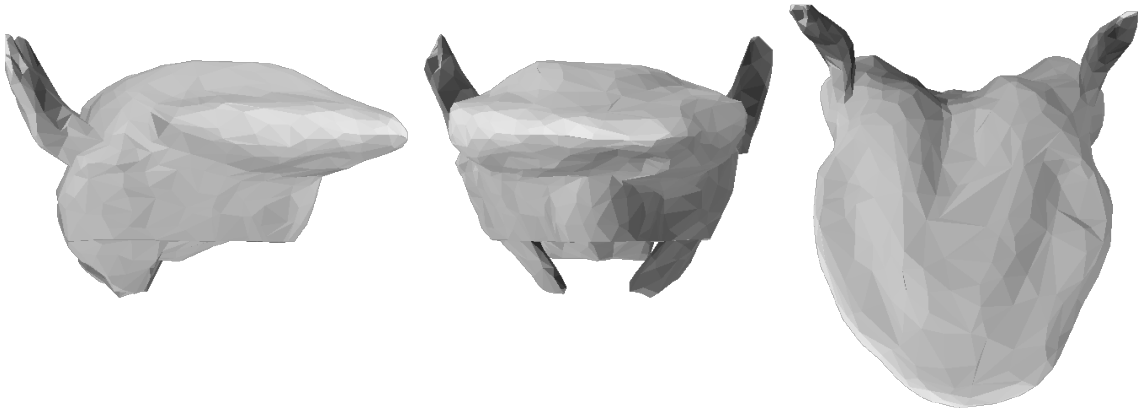

Figure 6.28: PCA of single muscle data: p_3 at maximum value.

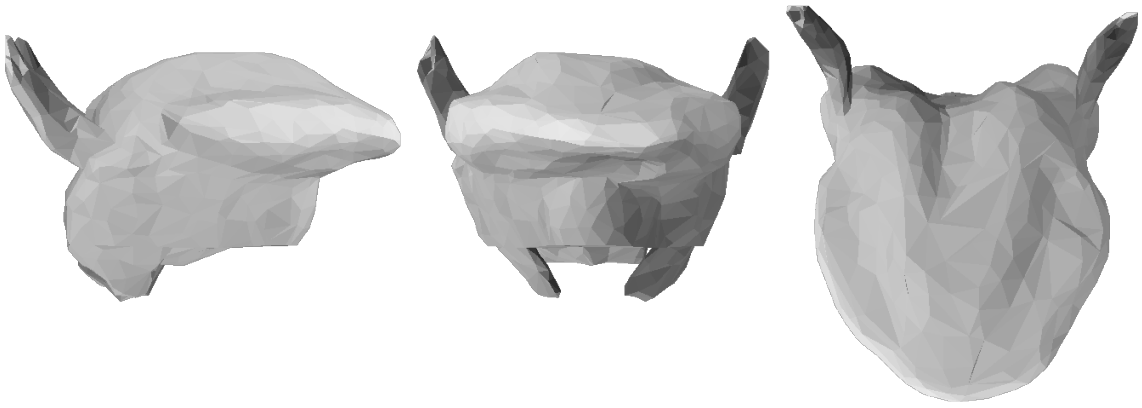

Figure 6.29: PCA of single muscle data: p_4 at minimum value.

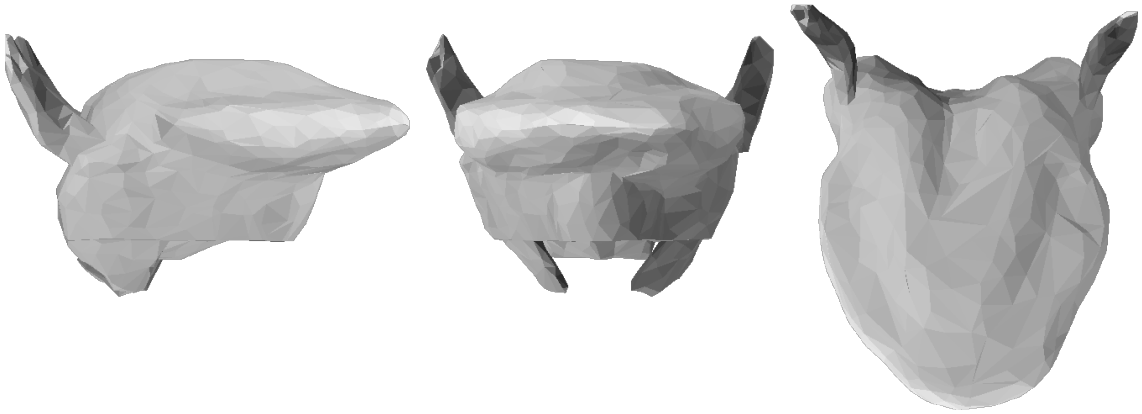

Figure 6.30: PCA of single muscle data: p_4 at maximum value.

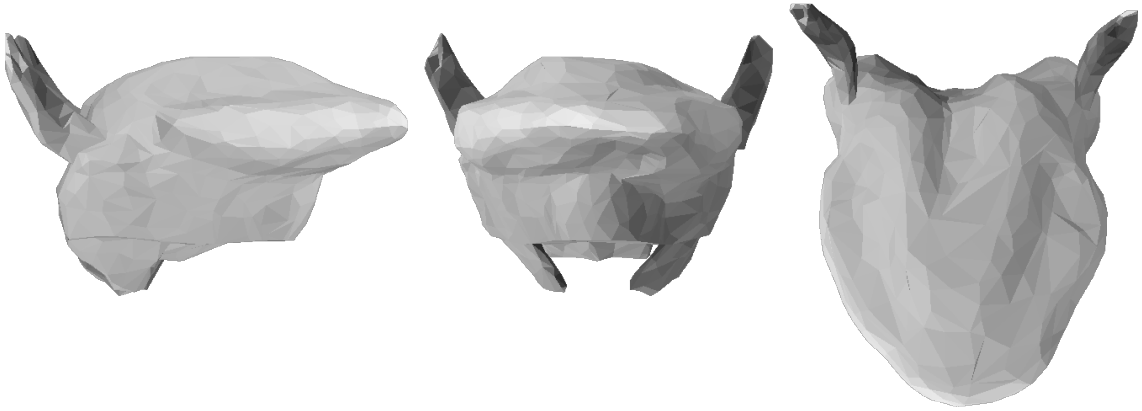

Figure 6.31: PCA of single muscle data: p_5 at minimum value.

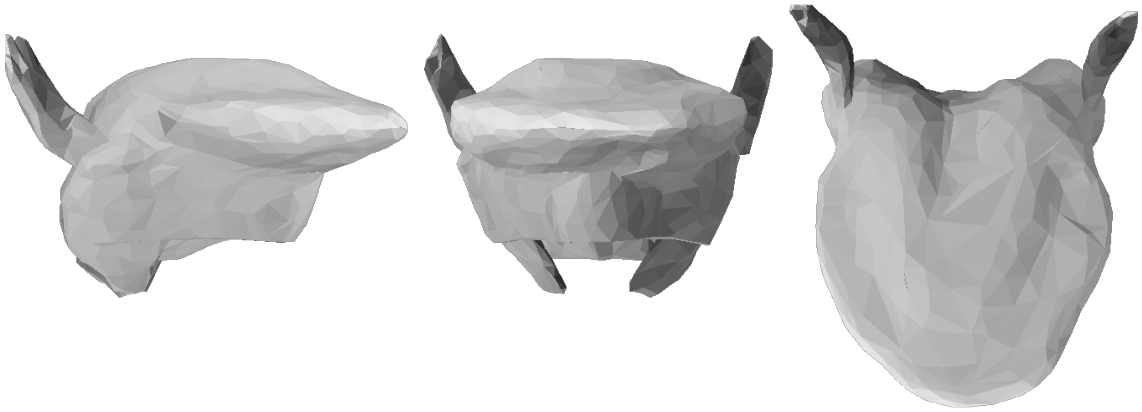

Figure 6.32: PCA of single muscle data: p_5 at maximum value.

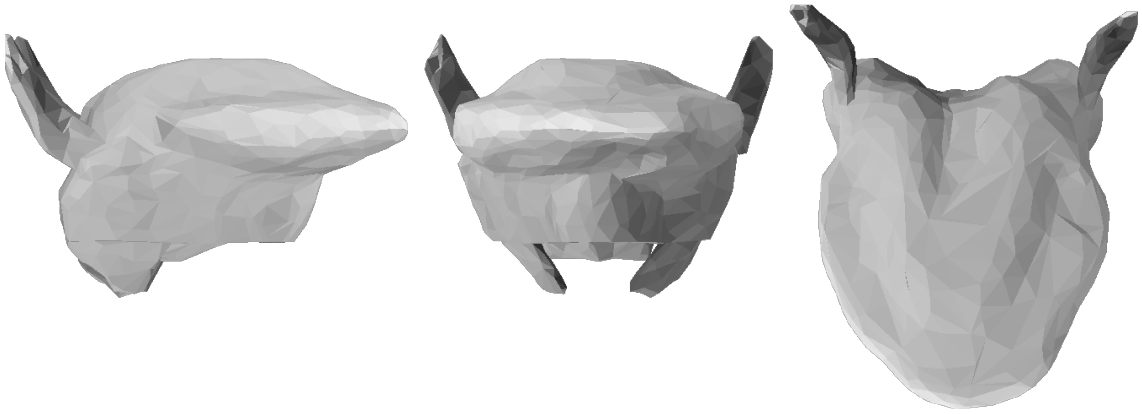

Figure 6.33: PCA of single muscle data: p_6 at minimum value.

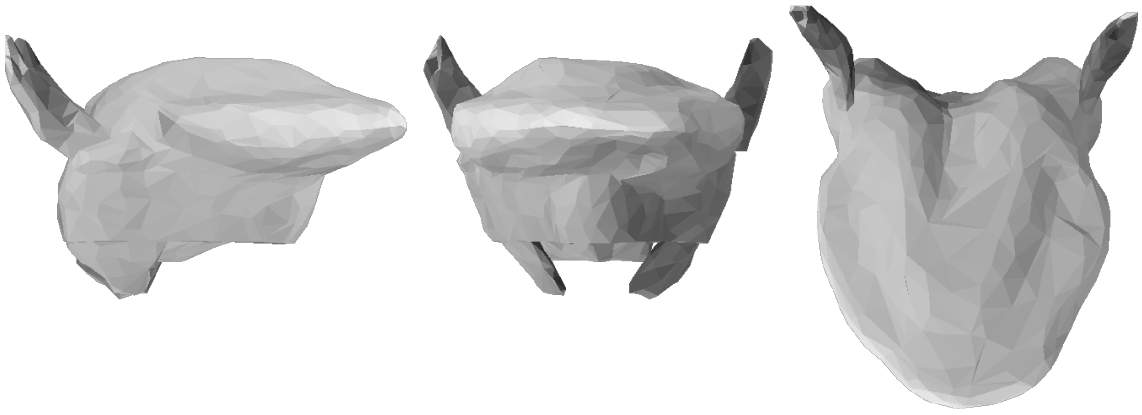

Figure 6.34: PCA of single muscle data: p_6 at maximum value.

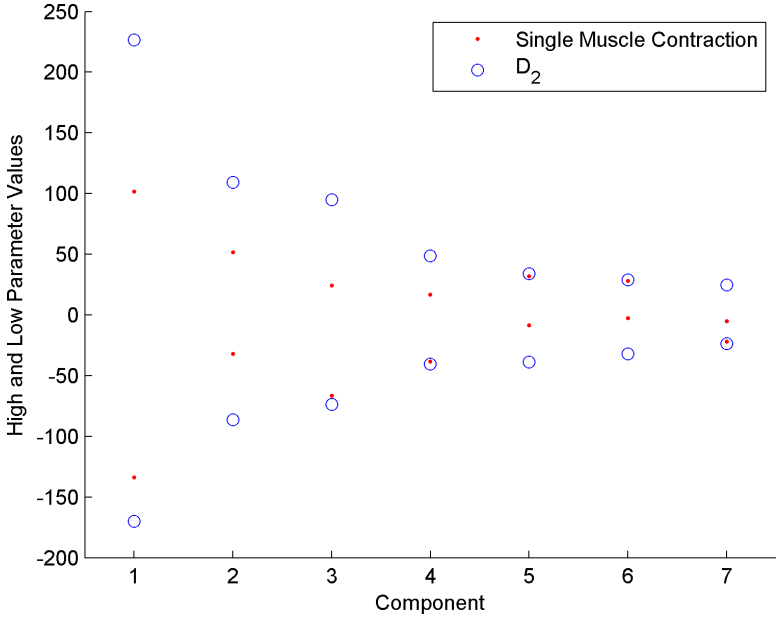

Figure 6.35: Range of $\mathbb{P}_1$ parameter values for full D_2 data set, compared the range with samples in which only one muscle is contracted.

6.2 Testing specific predictions

This section examines some of the predictions made by previous researchers. The purpose is to determine how the predictions made by other researchers match those made by the tongue model. In several of the cases reported here, the results do not match very well. This could be because the predictions are actually inaccurate, or because the model is lacking in the features necessary for the predictions to be shown true. With this caveat in mind, it seems nevertheless clear that certain predictions of previous researchers are inadequate for explaining tongue function. Previous researchers operated without very detailed anatomical information, and without biomechanical models to test their ideas, so this is quite understandable. In the tests below, both model error and incorrect predictions are considered as explanations for mismatches.

6.2.1 Predictions of Abd-El-Malek (1955)

Here the predictions that Abd-El-Malek (1955) made about his canonical feeding postures are tested, by simulating his proposed activation patterns with the biomechanical model. Abd-El-Malek predicts that a lowered medial surface with raised lateral surface is associated with contraction of genioglossus and hyoglossus, and bilateral contraction of styloglossus. The tongue posture associated with this contraction schema is shown in Figure 6.36, which should be compared with Figure 2.1.

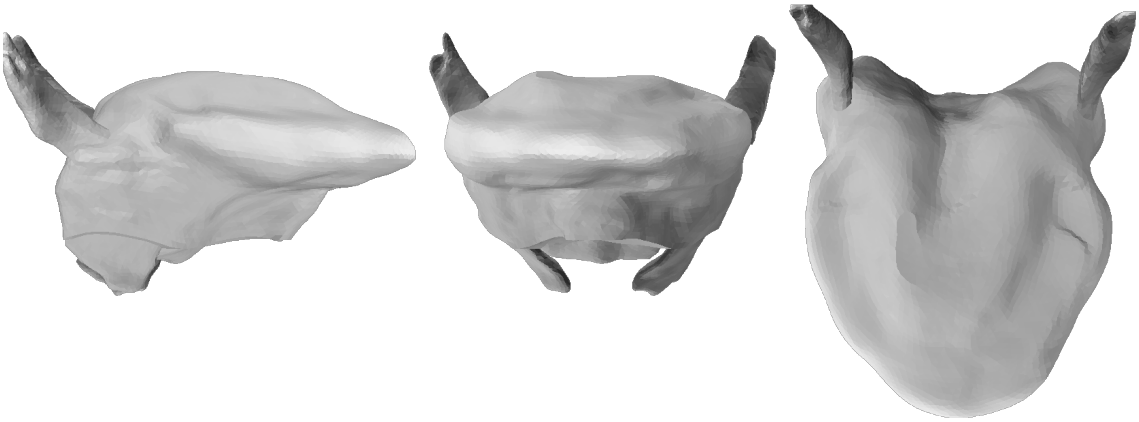

Figure 6.36: Symmetric full strength contraction of genioglossus, hyoglossus, and styloglossus.

Trough formation (Figure 2.2) is achieved with the same contraction schema, with the additional contraction of the superior longitudinal muscles. This is shown in Figure 6.37, which is to be compared with Figure 2.2.

Abd-El-Malek associates both the throwing and guarding stages with contralateral action of styloglossus. Depressing the tongue tip comes from contracting inferior longitudinalis (he specifies unilateral action; ipsilateral action is assumed here as well). Figure 6.38 shows contraction of right styloglossus and left inferior longitudinalis.

Comparisons of the simulated tongue postures with the tongue postures that they were intended to simulate reveals a substantial disparity. In fact, many of these postures were unable to be produced by the tongue model. The one shape that could be reproduced fairly well was Abd-El-Malek's preparatory stage. A tongue

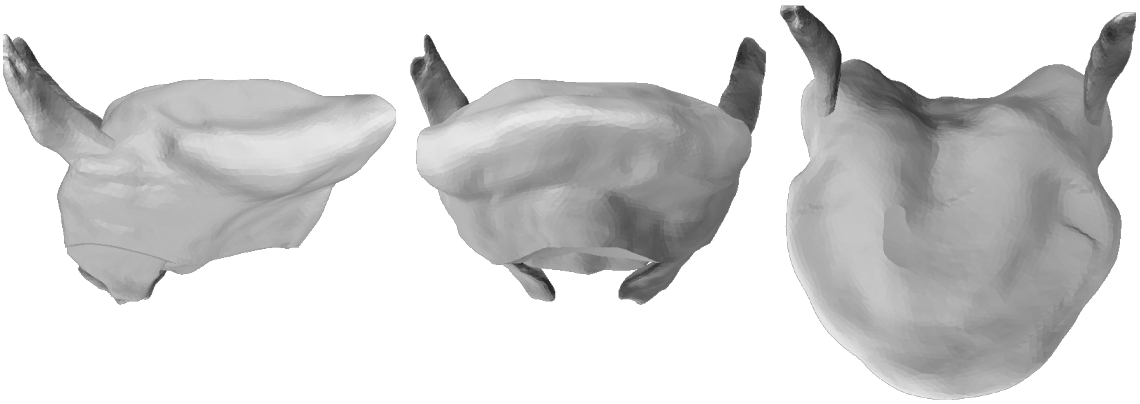

Figure 6.37: Symmetric full strength contraction of genioglossus, hyoglossus, styloglossus, and superior longitudinalis.

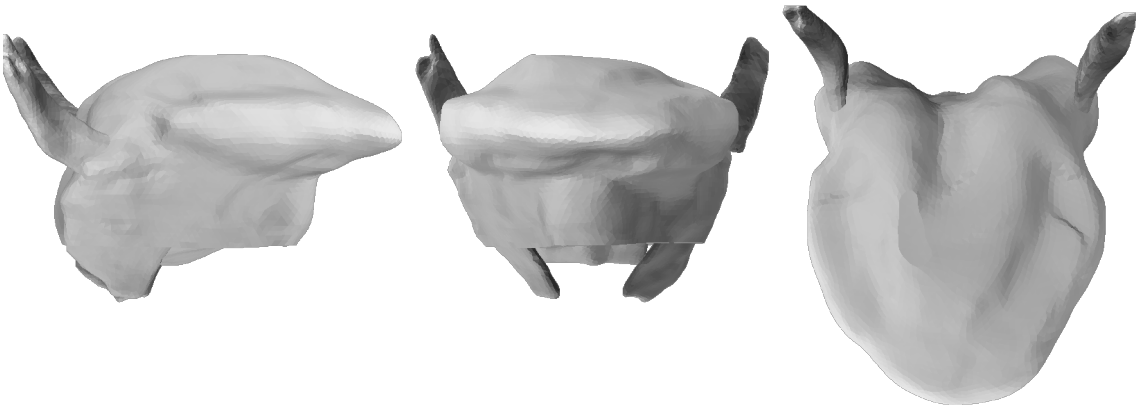

Figure 6.38: Full strength contraction of the *right* styloglossus and *left* inferior longitudinalis.

shape qualitatively similar to Figure 2.1 is shown in Figure 6.39 and Figure 6.40. Figure 6.39 was obtained by manipulating tongue parameter values to match the shape in Figure 2.1. Then muscle activation levels were calculated with a neural network, and a new shape was simulated, which is Figure 6.40. Muscle activation levels are provided in the caption of Figure 6.39, which indicate that the preparatory stage tongue posture is brought about mostly by verticalis, superior longitudinalis, and transversus.

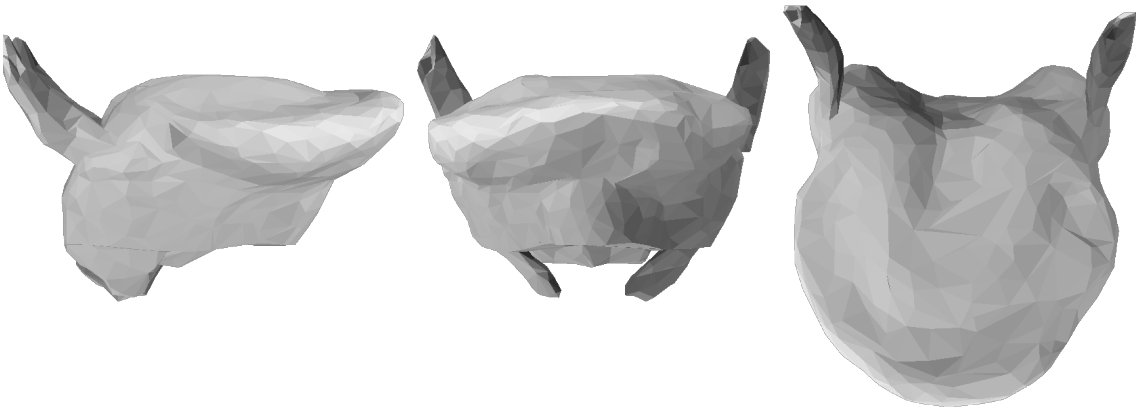

Figure 6.39: Author's best attempt to match the posture found in Abd-El-Malek (1955)'s Figure 1 (Figure 2.1). The muscle contractions—calculated with a neural network—are: GG: 0.07812; HG: 0.0002757; SG: 0.04657; TV: 0.01404; VR: 0.3474; SL: 0.3473; IL: 0.004803.

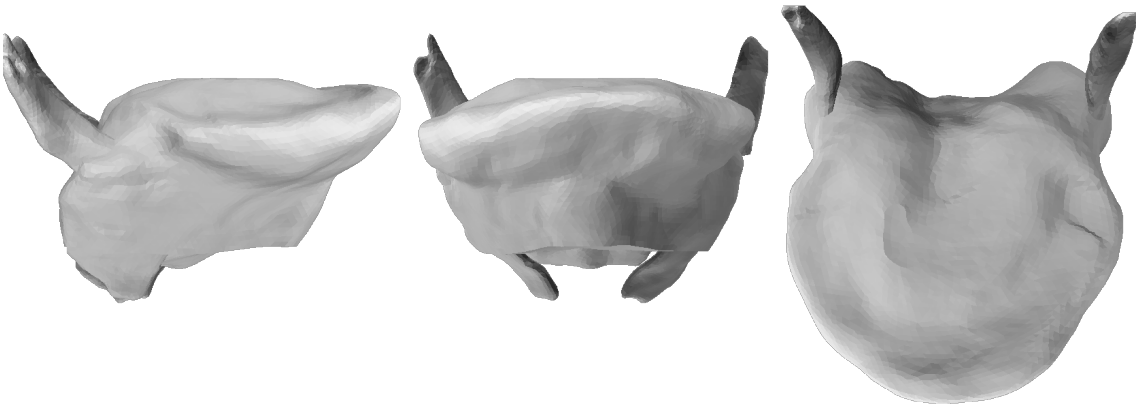

Figure 6.40: Simulation using the muscle activation levels given in the caption of Figure 6.39.

There are at least two explanations of why the tongue model could not generate

most of Abd-El-Malek's tongue postures (Figures 2.1 through 2.5). The first is that the model was lacking in the features necessary to produce the postures. Many of Abd-El-Malek's predictions involved styloglossus, but as discussed in Section 6.1.1, the model's representation of styloglossus was limited. It could also be that producing Abd-El-Malek's postures requires a finer control mechanism than whole muscle contraction. It may be necessary to specify activation at the level of the muscle compartment, for instance.

A second reason is that these "tongue" postures in fact crucially rely on other anatomical structures. This certainly seems true for Figure 2.5: the author of the present work is not able to raise the sides of his tongue in this way, except by squeezing the tongue with the lips. If this proves a general requirement, then it would seem that raising the sides of the tongue is actually accomplished with the combined action of the tongue and lips. More generally, the tongue model does not produce much tongue lowering, which is required for instance for both Figures 2.1 and 2.2.[4] It seems likely that in these and similar postures, the tongue is not being compressed vertically so much as lowering as the hyoid bone is lowered. Since the present model did not systematically study the effect of hyoid position on tongue posture, however, this remains conjecture.

6.2.2 Predictions of Kier and Smith 1985

Kier and Smith (1985) make clear predictions about what muscles are involved in elongation, shortening, and bending of the tongue.[5] These predictions are tested, and largely confirmed, in this section.

They predict that elongation of the tongue is achieved by contracting muscles that reduce the cross-sectional area of the tongue. To test this prediction, Figure 6.41 shows the posture that the tongue assumes with full contraction of transversus and verticalis. Tongue protrusion is evident in this posture.

Kier and Smith suggest that shortening of the tongue is achieved by contraction

[4]This can be appreciated by observing the "height" of the tongue in Figures 5.6 through 5.19.

[5]The authors also make predictions about stiffening torsion, which are not addressed here.

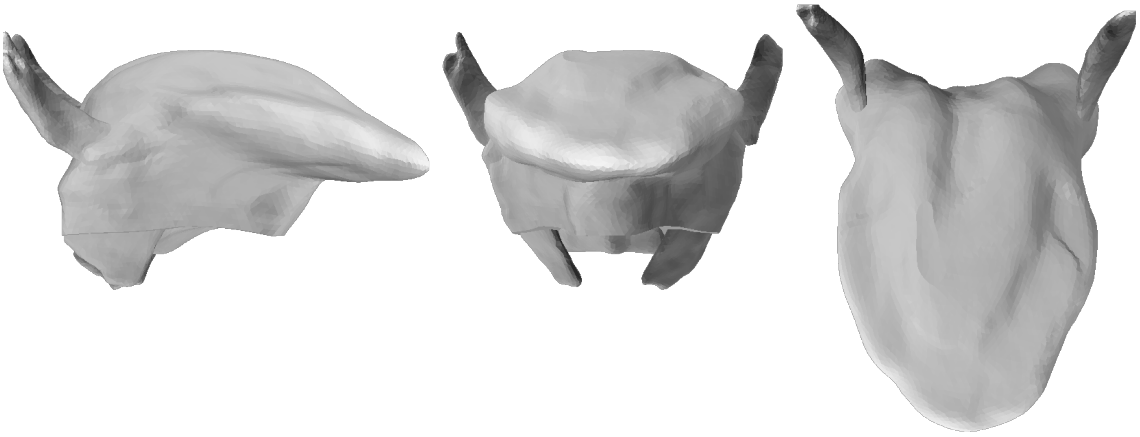

Figure 6.41: Tongue posture with full contraction (0.350 MPa active stress) of transversus and verticalis.

of the longitudinal muscles. Figure 6.42 shows the tongue posture when superior longitudinalis and inferior longitudinalis are fully contracted; shortening is evident.

Figure 6.42: Tongue posture with full contraction (0.350 MPa active stress) of superior longitudinalis and inferior longitudinalis.

Finally, bending is thought to occur only when either the superior or inferior longitudinal muscle are contracted, *and* when longitudinal compression is prevented by activation of transversely-oriented muscles. Testing this prediction requires comparison of several tongue postures. The effects of contracting *just* superior longitudinalis and *just* inferior longitudinalis are shown respectively in Figures 6.6 and 6.7. Both postures show bending in the expected direction, which is not consistent with Kier

and Smith's prediction that contraction of the transverse musculature is necessary. Figures 6.43 and 6.44 illustrate the effect of adding contraction of verticalis. For superior longitudinalis, the body of the tongue is raised by contraction of transversus, but the curvature does not appear to be greater. For inferior longitudinalis, the tongue body is also raised, *and* the curvature of the tongue is noticeably greater. Thus for the case of downward bending, contraction of transversus seems to aid bending, but is not required for bending.

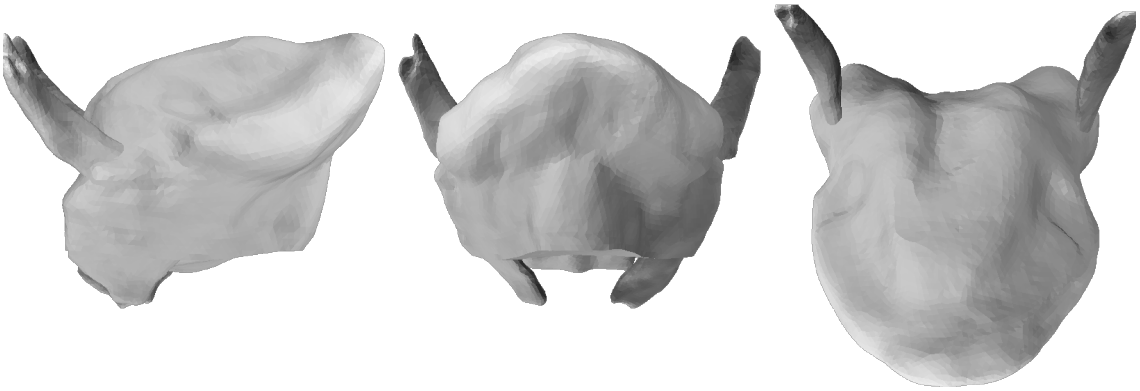

Figure 6.43: Symmetric full strength contraction of superior longitudinalis and transversus.

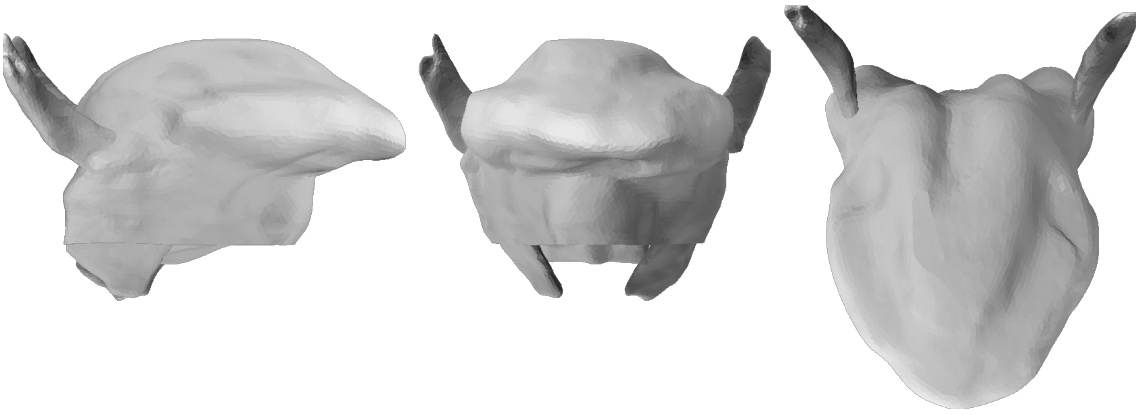

Figure 6.44: Symmetric full strength contraction of inferior longitudinalis and transversus.

A further test consists in determining what muscular activations are associated with the maximal elongation, shortening, and bending postures that could be pro-

	p_1	p_2	p_3	p_4	p_5	p_6	p_7
a.	145.72	26.59	16.24	23.21	-2.14	4.92	-1.59
b.	-31.73	-36.65	1.97	-31.37	3.63	0.37	1.26
c.	-110.58	-30.81	54.02	-0.54	5.07	3.41	-3.25
d.	47.12	-7.46	48.98	37.79	0.00	-0.01	6.48

Table 6.1: Parameter values (in $\mathbb{P}_1$) for particular tongue postures. a. Maximum elongation—also maximum downward bending. b. Maximum elongation without tongue tip lowering. c. Maximum shortening. d. Maximum upward bending.

	GG	**HG**	**SG**	**TV**	**VR**	**SL**	**IL**
a.	0.1491	0.0014	0.1346	0.2532	0.3259	0.0007	0.0012
b.	0.1707	0.0801	0.0173	0.1196	0.0005	0.3401	0.3471
c.	0.0017	0.0109	0.0119	0.2515	0.0004	0.3423	0.0033
d.	0.0138	0.0160	0.2547	0.3460	0.3451	0.0969	0.0007

Table 6.2: Predicted muscle activation levels (in MPa) for particular tongue postures. See the caption of Table 6.1 for descriptions of each posture.

duced by the model. For this purpose, tongue posture parameters were manipulated with the Tongue Viewer program, in order to produce the most extreme gestures that fell within the convex hull of D_2 (Section 5.4). Muscle activation levels were calculated from these postures, and new postures were calculated using these activation levels. The parameter values for these postures are given in Table 6.1, and the muscle activation levels in Table 6.2.

The maximal elongation of the tongue coincides with maximal downward bending (Figures 6.45 and 6.46). The main muscles involved in this posture were (in descending order of contraction magnitude):[6] verticalis, transversus, genioglossus, and styloglossus. Since "elongation" may be felicitously understood not to involve bending, another posture was produced in which the tongue was elongated but the tip was not depressed (Figures 6.47 and 6.48). For this posture the main muscles

[6]Note that "contraction magnitude" is not necessarily an indicator of relative importance, since a full-strength contraction in one muscle might have a much smaller mechanical effect than a half-strength contraction in another muscle.

were inferior longitudinalis, superior longitudinalis, genioglossus, and transversus.

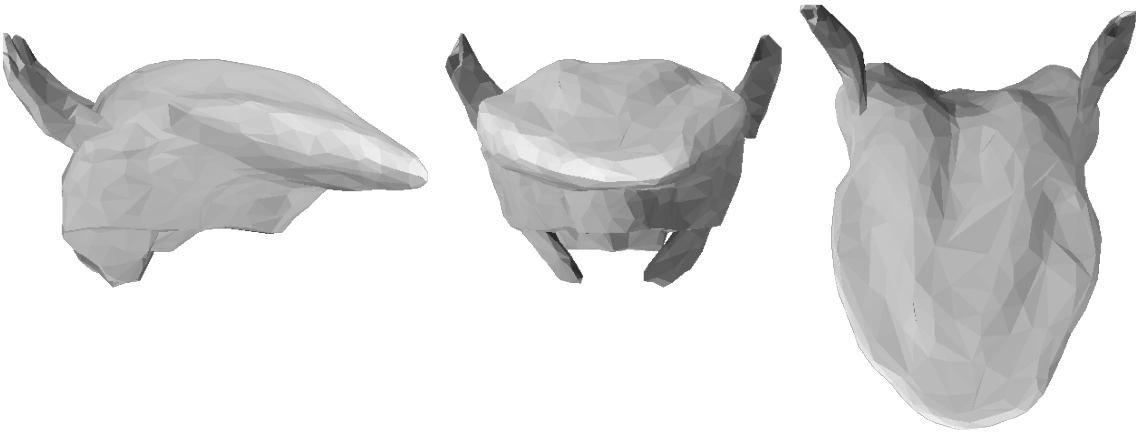

Figure 6.45: Author's best attempt to protrude the tongue.

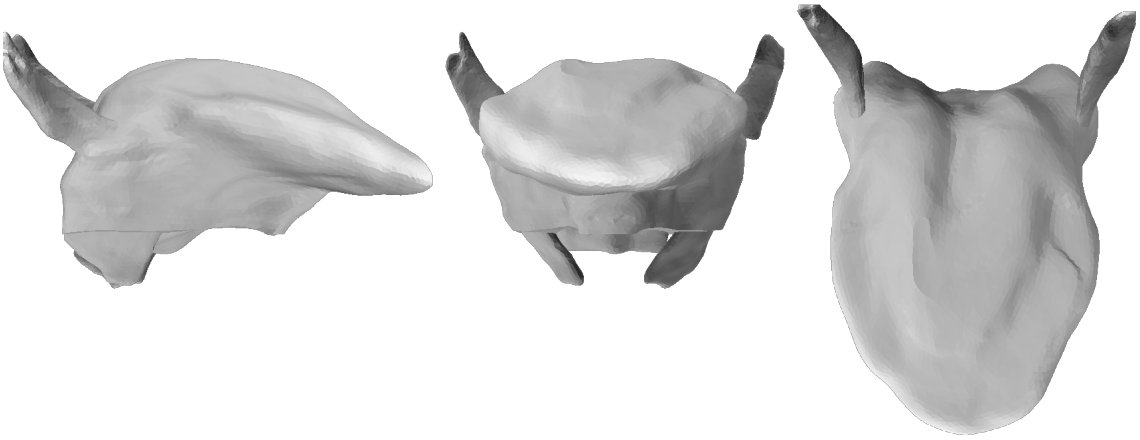

Figure 6.46: Maximum protrusion (corresponding to Figure 6.45. This posture is the result of an FEM simulation, based on the muscle activations predicted from the corresponding figure, which are given in Table 6.2a.

Maximum shortening of the tongue (Figures 6.49 and 6.50)[7] involved mainly superior longitudinalis and transversus. It is interesting to note that inferior longitudinalis played almost no role in this contraction.

[7]Compared to the posture resulting from the contraction of superior longitudinalis (Figure 6.42) and inferior longitudinalis, Figure 6.50 has tongue tip that is further forward, but the tip is raised much less as well.

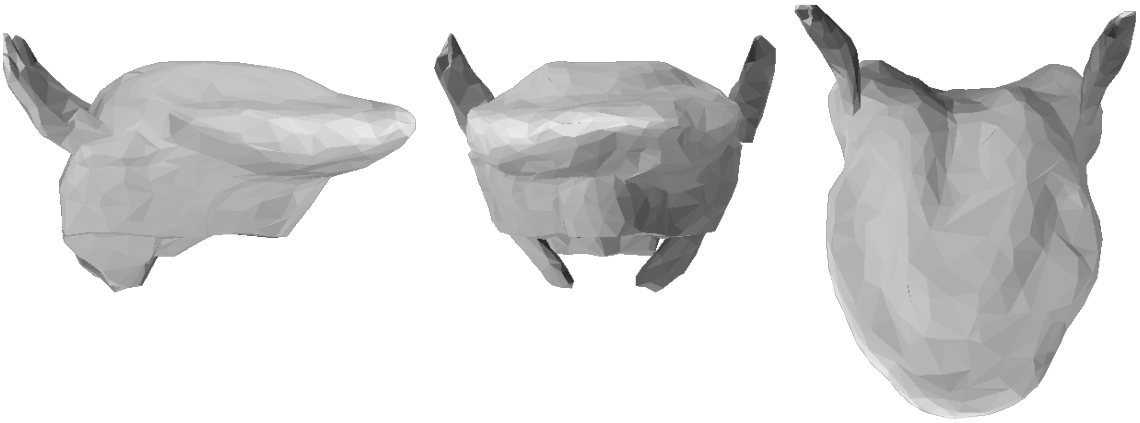

Figure 6.47: Author's best attempt to protrude the tongue without much bending.

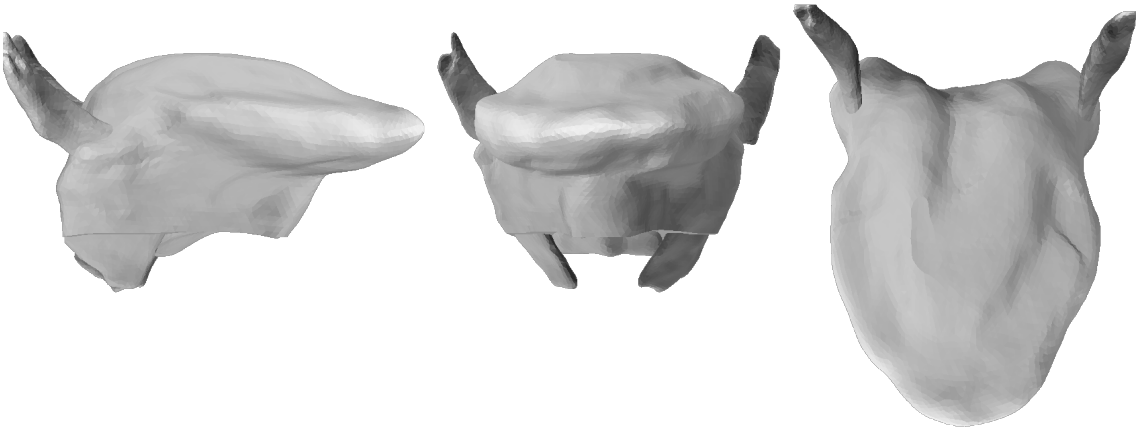

Figure 6.48: Maximum protrusion without much bending, (corresponding to Figure 6.47. This posture is the result of an FEM simulation, based on the muscle activations predicted from the corresponding figure, which are given in Table 6.2d.

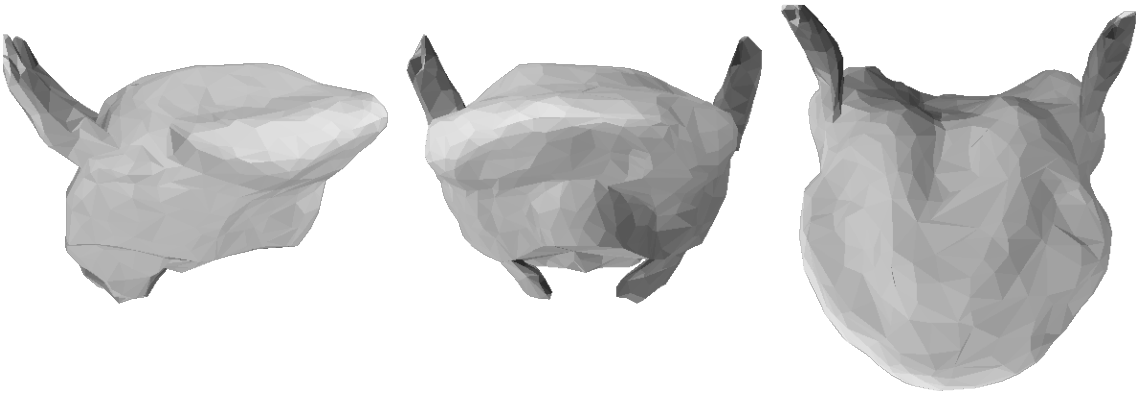

Figure 6.49: Author's best attempt to shorten the tongue.

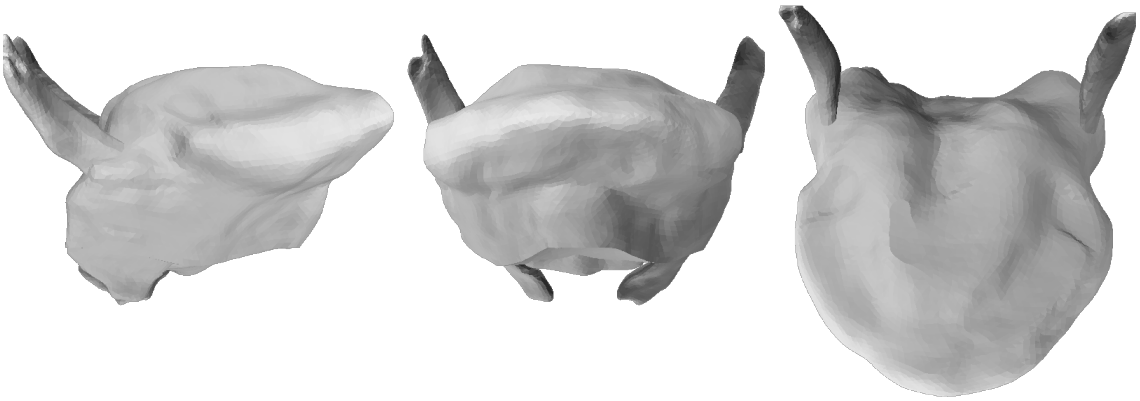

Figure 6.50: Maximum shortening (corresponding to Figure 6.49. This posture is the result of an FEM simulation, based on the muscle activations predicted from the corresponding figure, which are given in Table 6.2b.

Maximum upward bending (Figures 6.51 and 6.52) involved primarily contraction of transversus, verticalis, styloglossus, and superior longitudinalis.

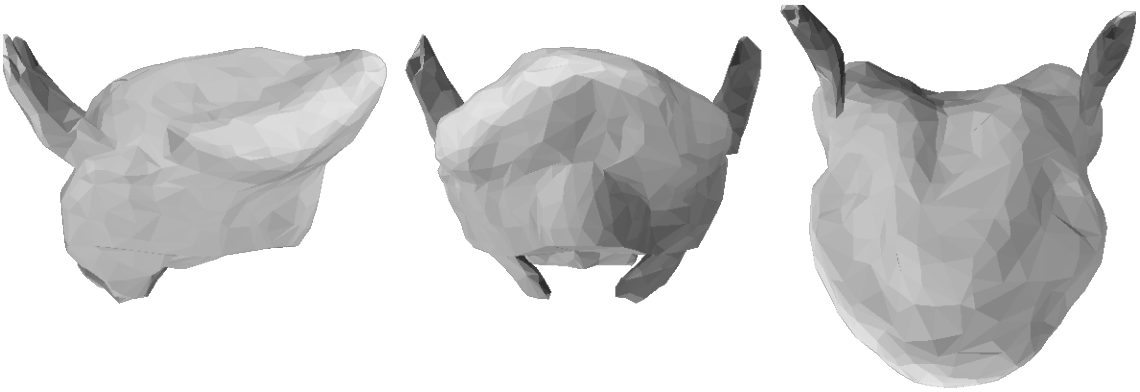

Figure 6.51: Author's best attempt to bend the tongue upward.

The results in this section largely support the predictions of Kier and Smith (1985). All of their predictions produced the kinds of postures that the authors predicted. When the extreme elongation, shortening, and bending postures were produced, all of the postures involved several muscles, typically both intrinsic and extrinsic, which also broadly supports the muscular hydrostat theory of tongue movement (Bailey and Fregosi, 2004).

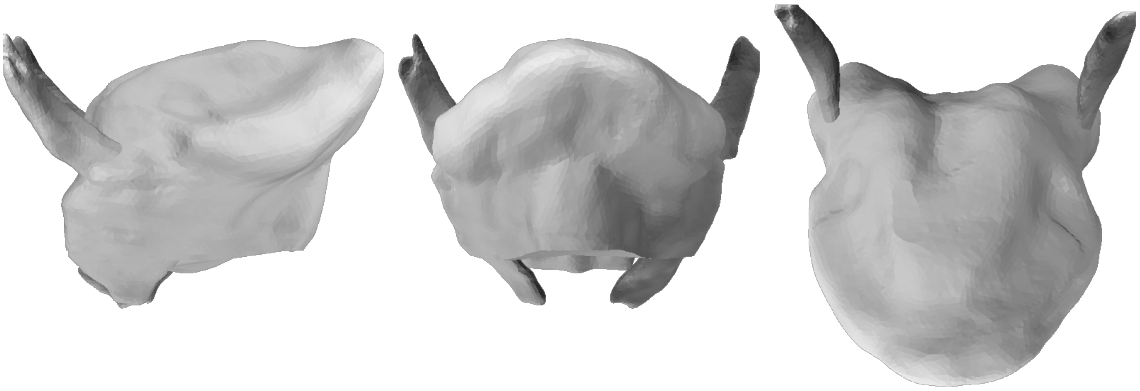

Figure 6.52: Maximum upward bending (corresponding to Figure 6.51. This posture is the result of an FEM simulation, based on the muscle activations predicted from the corresponding figure, which are given in Table 6.2c.

	high	low	back	AGG	PGG	GH	HG	SG
u	+	−	+	−	+	−	−	+
i	+	−	−	+	+	+	−	−
o	−	−	+	−	−	−	−	−
e	−	−	−	+	−	+	−	−
ɑ	−	+	+	−	−	−	+	−
æ	−	+	−	+	−	+	+	−

Table 6.3: Halle (1983)'s predictions of muscle activity for various vowels. AGG: anterior genioglossus. PGG: posterior genioglossus. GH: geniohyoid. GH: hyoglossus. SG: styloglossus.

6.2.3 Predictions of Halle 1983

In an effort to connect abstract phonological representations to concrete physiological parameters, Halle (1983) made some very specific predictions concerning the muscular activations associated with the vowels {u, i, o, e, ɑ, æ}. These predictions are reproduced in Table 6.3. The predictions are tested here, to the best of the model's ability.

A couple of adjustments had to be made to the model in order to perform tests of the predictions. Halle predicts that geniohyoid is active in vowel production. Since the hyoid bone is mobile in speech this is very reasonable, but geniohyoid

was not part of the present model. Therefore, a specified displacement of the hyoid bone was used as a proxy for geniohyoid activity. Since the geniohyoid originates on the superior part of the hyoid bone and inserts on the posterior of the mandible, it was assumed that the geniohyoid muscle would advance the hyoid anteriorly and superiorly along this path, which is shown in Figure 6.53. Hiiemae and Palmer (2003) recorded the position of the hyoid bone during speech, and their results (particularly their Figure 4) indicate that in the sagittal plane the hyoid moves in an area with a diameter of about 1 cm. Therefore, activity of the geniohyoid was associated with 0.5 cm of movement along the path of the geniohyoid. There was no side-to-side movement.

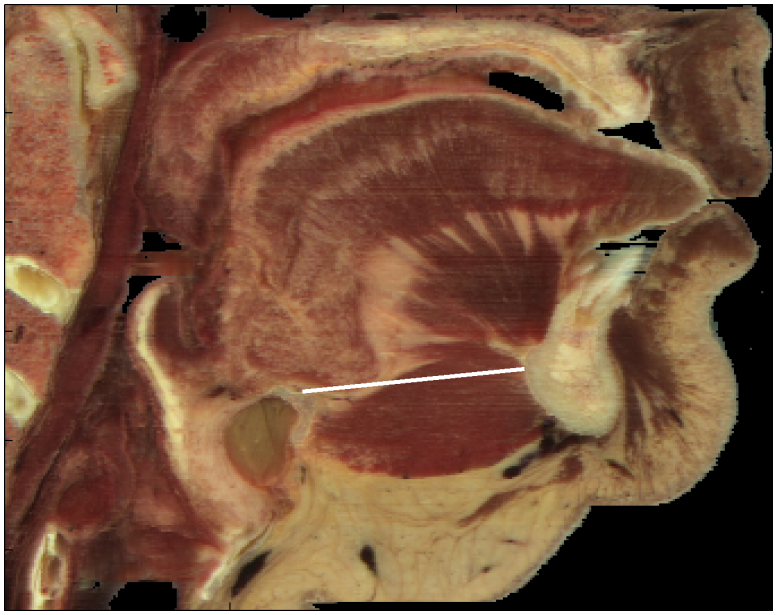

Figure 6.53: Path of hyoid movement.

Halle also differentiates between anterior and posterior genioglossus, which is not a distinction otherwise included for the present model. In fact Halle does not define an anatomical boundary. Therefore, the study of canine genioglossus compartmentalization by Mu and Sanders (2000) was taken as a guide, to create a "horizontal" and one "oblique" compartment. The division is shown in Figure 6.54.

Finally, since Halle predicted activation levels using binary (+/−) notation, mus-

cles were either inactive or fully active (0.35 MPa active stress).

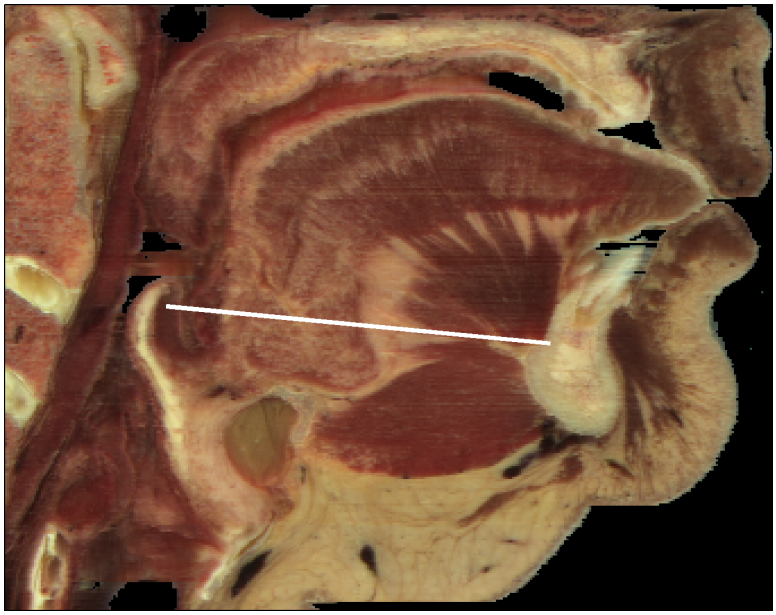

Figure 6.54: Partition of genioglossus into anterior and posterior parts.

The results of the tests of Halle's predictions for the vowels {u, i, o, e, ɑ, æ} are shown in Figures 6.55 through 6.60, respectively. Since Halle does not associated any muscle contraction with [o], Figure 6.57 is the rest position, and can be used for comparison. None of the postures bear a very close resemblance to the intended target. The predicted posture for [u] is distinct from [o] in its depressed tongue tip, which [i] features as well, but to a larger degree. [e] is similar in shape to [i], except that the tongue body is slightly higher. [ɑ] is just slightly pulled back from [o]. The predicted posture for [æ] is plausible, though, since the tongue tip is most depressed in that posture.

Generally, Halle's predictions seem not to generate appropriate postures because of reliance on styloglossus as an elevator and retrusor of the tongue, and a neglect of the intrinsic muscles of the tongue. As discussed in Section 6.1.1, in the present model the effect of styloglossus should be expected to be underestimated. It has also been demonstrated however (Section 6.1 and Section 6.2.4 below), that the intrinsic muscles have a substantial role in raising the tongue. The absence of a role for the

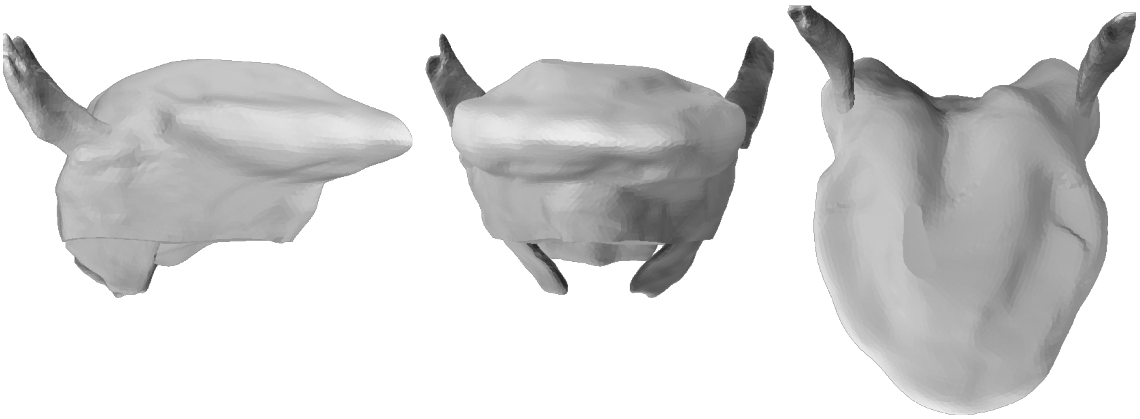

Figure 6.55: Position computed based on Halle's prediction for [u].

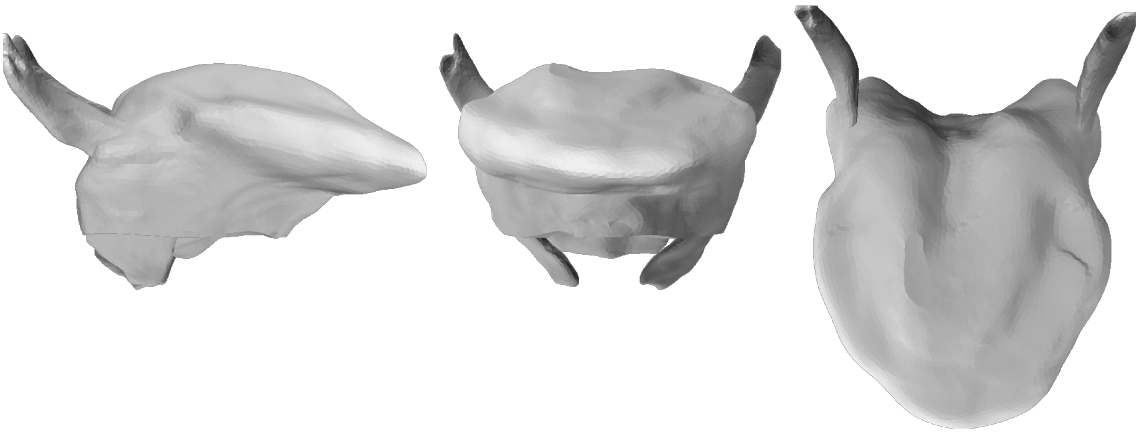

Figure 6.56: Position computed based on Halle's prediction for [i].

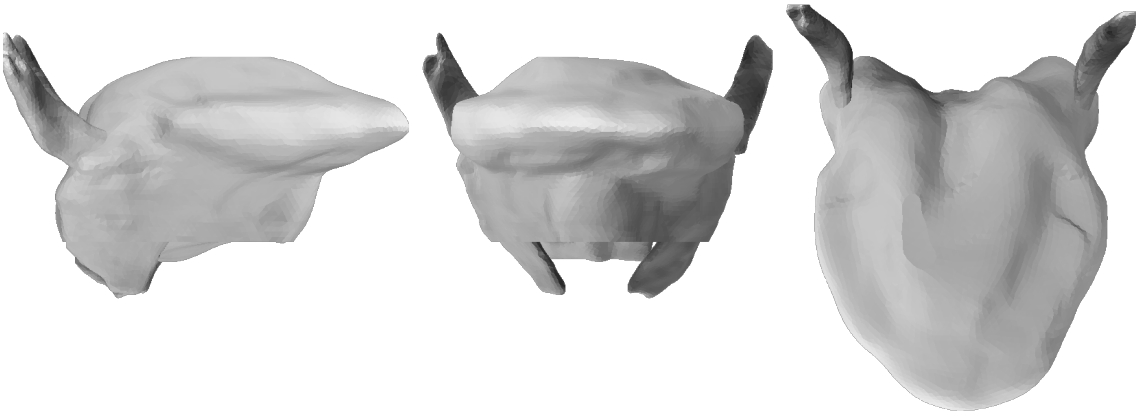

Figure 6.57: Position computed based on Halle's prediction for [o]. (This is in fact the rest position.)

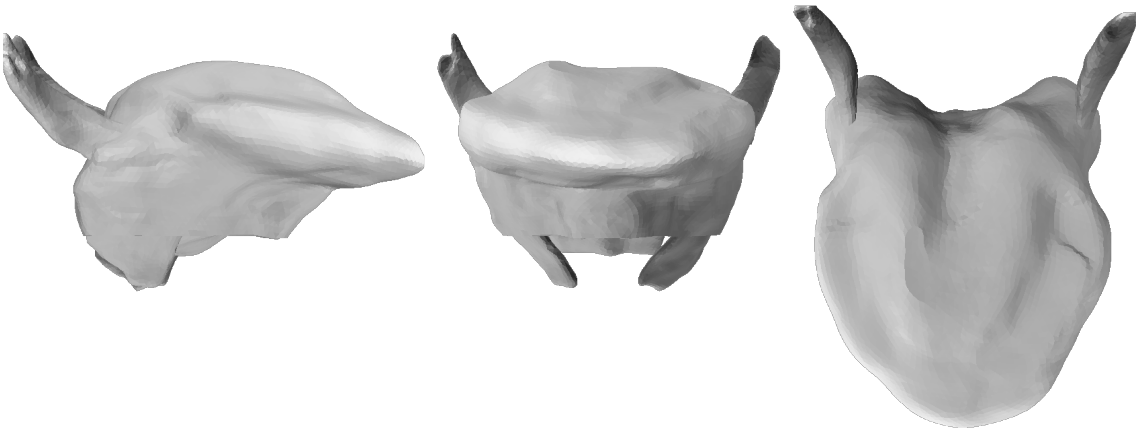

Figure 6.58: Position computed based on Halle's prediction for [e].

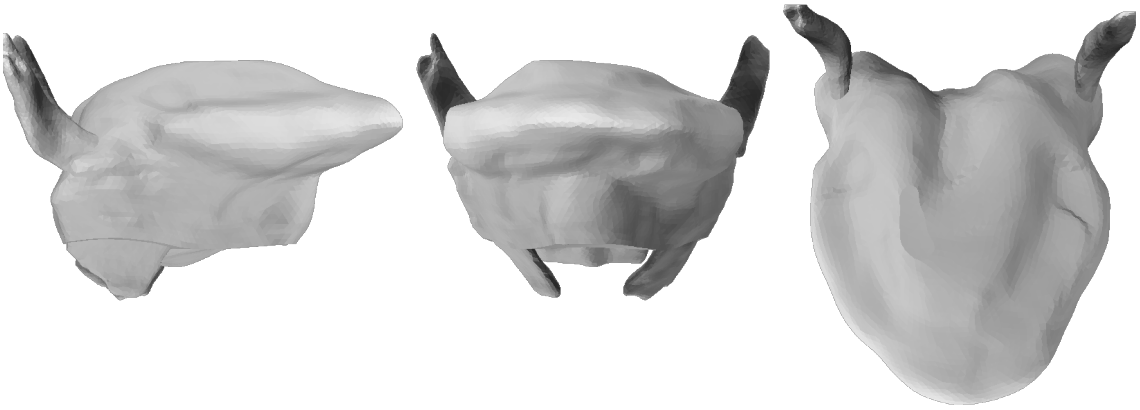

Figure 6.59: Position computed based on Halle's prediction for [a].

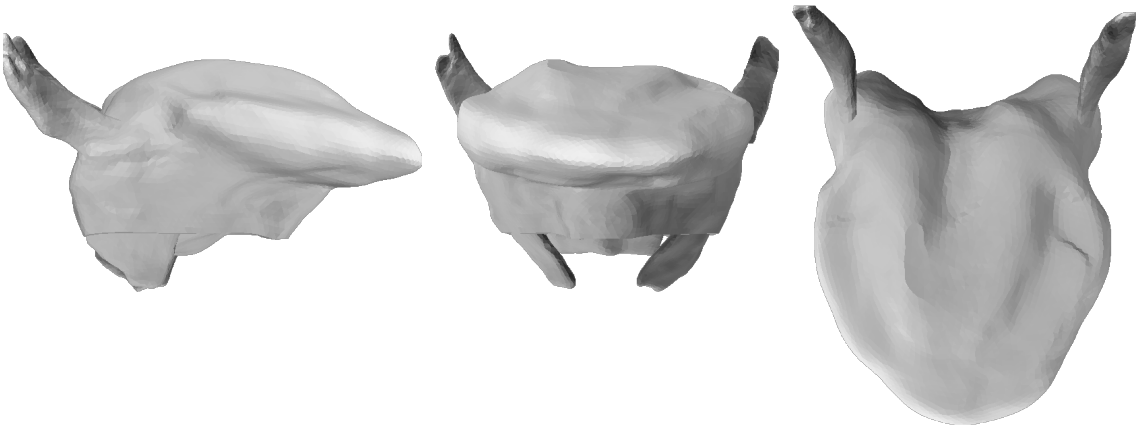

Figure 6.60: Position computed based on Halle's prediction for [æ].

intrinsic muscles in Halle's system is likely to be a real cause of inadequacy in his predictions.

6.2.4 Investigation of an intrinsic/extrinsic split

As noted in Section 2.1.2.4, Perkell (1969) has suggested that there is a functional division between the intrinsic and extrinsic muscles of the tongue, which is approximately that the intrinsic muscles are responsible for shaping the tongue, and the extrinsic muscles are responsible for positioning it. Perkell's prediction is fairly specific, but it might be hypothesized more generally that there is *some* distinction in the mechanical activities of the intrinsic and extrinsic muscles. This hypothesis receives empirical support if non-trivial distinctions can be made between the intrinsic and extrinsic muscles, i.e., ones which do not consist in enumerating the individual consequences of contracting genioglossus, styloglossus, and hyoglossus on the one hand, and transversus, verticalis, inferior longitudinalis, and superior longitudinalis on the other. The hypothesis is tested here by comparing subsets of $D_2^{\mathbb{P}_1}$. Specifically, the set of tongue postures that involve contraction of only the intrinsic muscles and the set of tongue postures that involve only contraction of the extrinsic muscles are compared.

Postures from these data sets were projected onto $\mathbb{P}_1$. A comparison of ranges of the parameters of the data is shown in Figure 6.61. The postures associated with the minimum and maximum parameter values of the extrinsic muscles are given in Figures 6.62 through 6.75; the corresponding figures for the intrinsic muscles are given in Figures 6.76 through 6.89.[8] If the intrinsic and extrinsic muscles have functional distinctions, it might be expected that one subset of muscles would dominate

[8]The reader is again referred to the Tongue Viewer program, described in Appendix C, for demonstration of these facts in a way that is perhaps easier to take in. For the graphics in this subsection, the parameters which were *not* being illustrated were set to a default value, which was the mean of the minimum and maximum values (the center of the range). This is because the parameters for the rest position did not always fall within the range of the data. This is not thought to have introduced any bias into any of the figures.

one parameter, or that the sets of muscles would cover distinct ranges in the parameter space. For the present model, however, Figure 6.61 shows that the intrinsic and extrinsic muscles both contribute to variation in the same range in six of the seven parameters of $\mathbb{P}_1$. The one exception is p_5, but since this parameter accounts for so little of the variance in the data that it is not really strong evidence of a functional division. The findings as a whole do not provide supporting evidence for the hypothesis that the intrinsic/extrinsic split has functional significance.

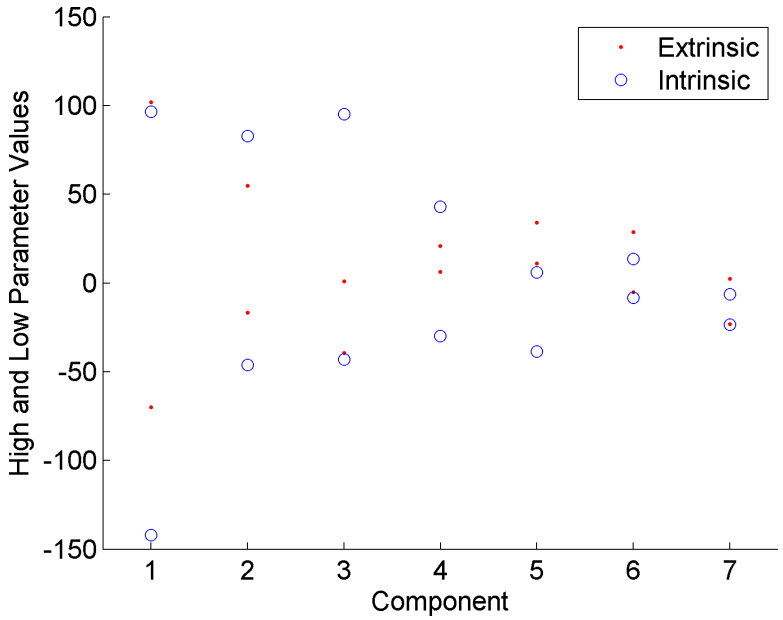

Figure 6.61: Comparison of the parameter ranges ($\mathbb{P}_1$) of tongue postures in which only extrinsic and only intrinsic muscles are contracted.

It is interesting to note that, for all seven parameters, the intrinsic-only postures demonstrate a greater range of movement. This is a little surprising, since one might otherwise expect that a muscle with a bony attachment would exert a greater mechanical influence than a muscle without one. It is, however, consistent with the muscular hydrostat theory of tongue movement since that theory predicts that the tongue deforms primarily by coactivating intrinsic muscles.

It is possible that the differences between the intrinsic and extrinsic muscles are not captured by considering variation along the parameters of $\mathbb{P}_1$. An alternate test is to derive parameter spaces for the intrinsic and extrinsic subsets separately

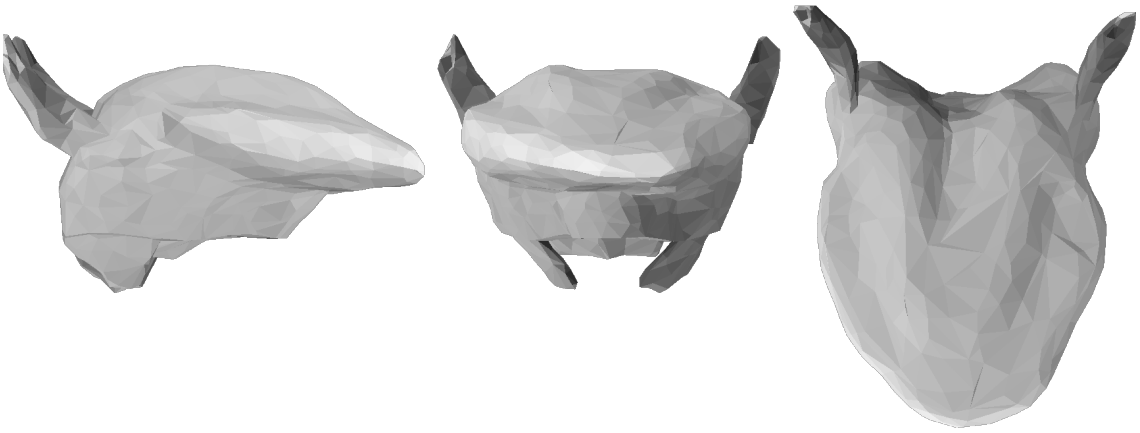

Figure 6.62: Maximum value of p_1, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

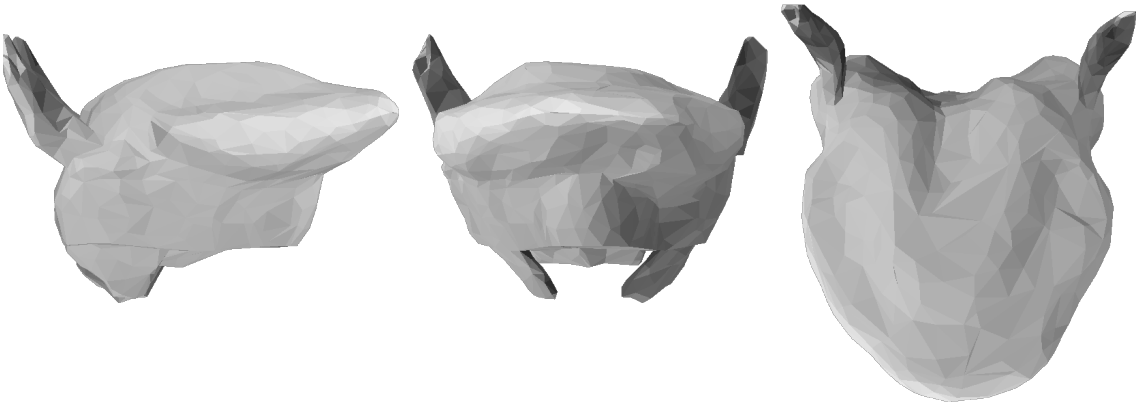

Figure 6.63: Minimum value of p_1, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

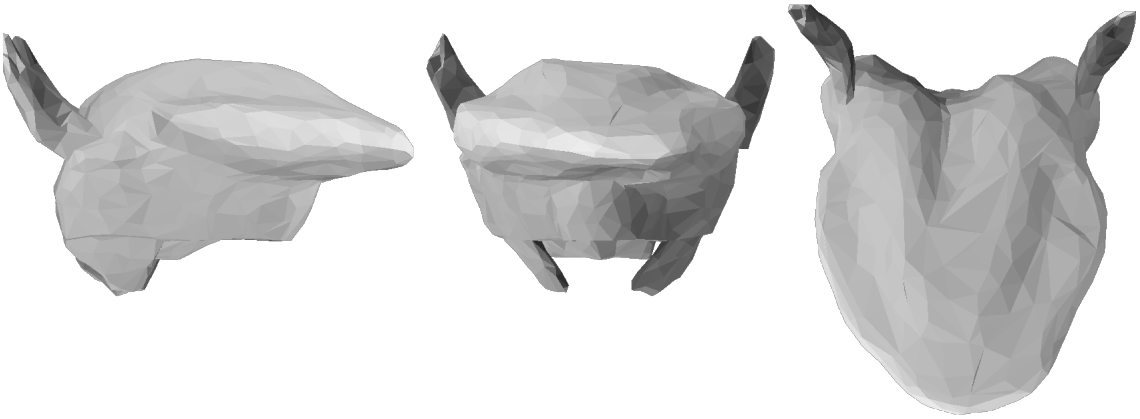

Figure 6.64: Maximum value of p_2, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

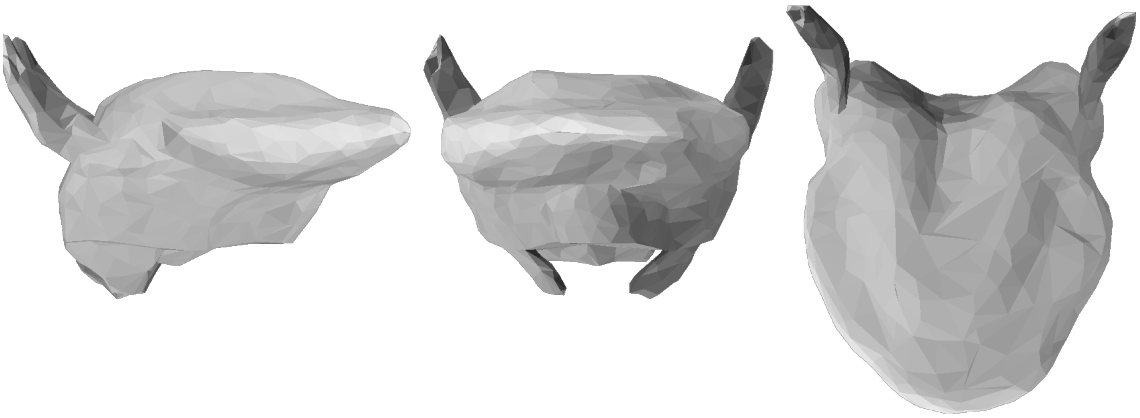

Figure 6.65: Minimum value of p_2, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

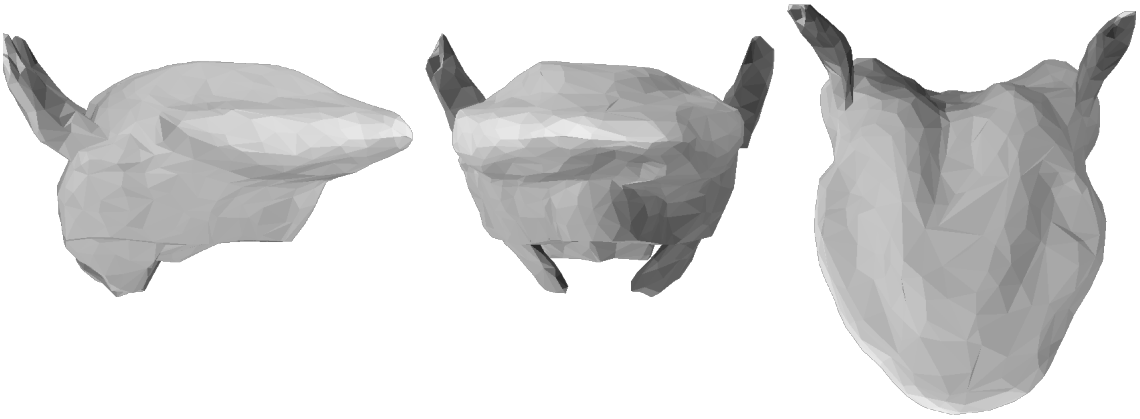

Figure 6.66: Maximum value of p_3, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

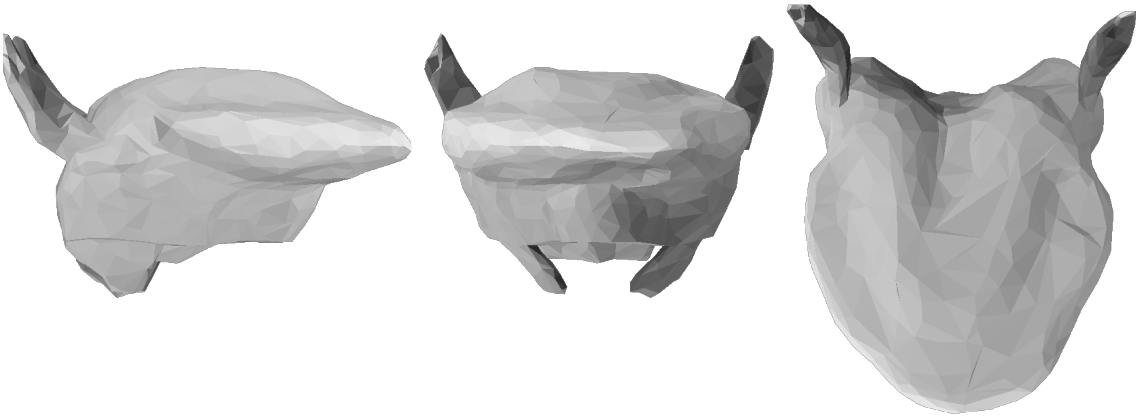

Figure 6.67: Minimum value of p_3, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

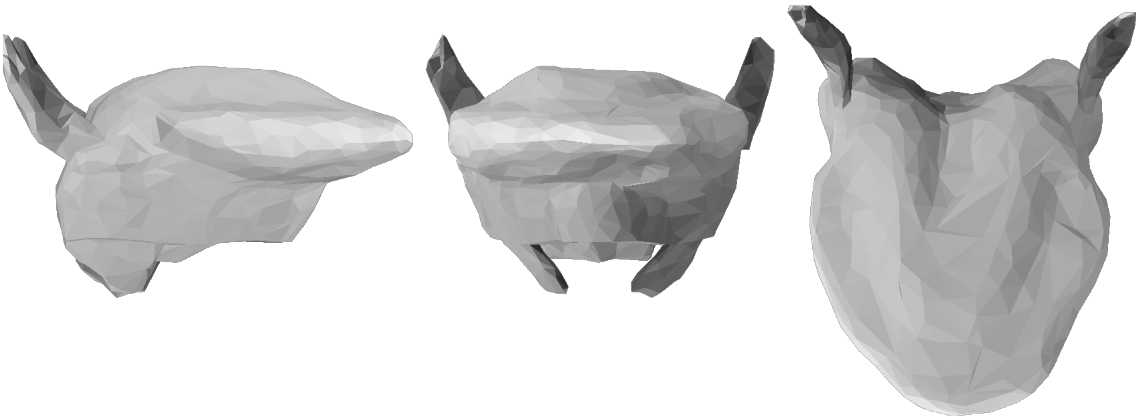

Figure 6.68: Maximum value of p_4, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

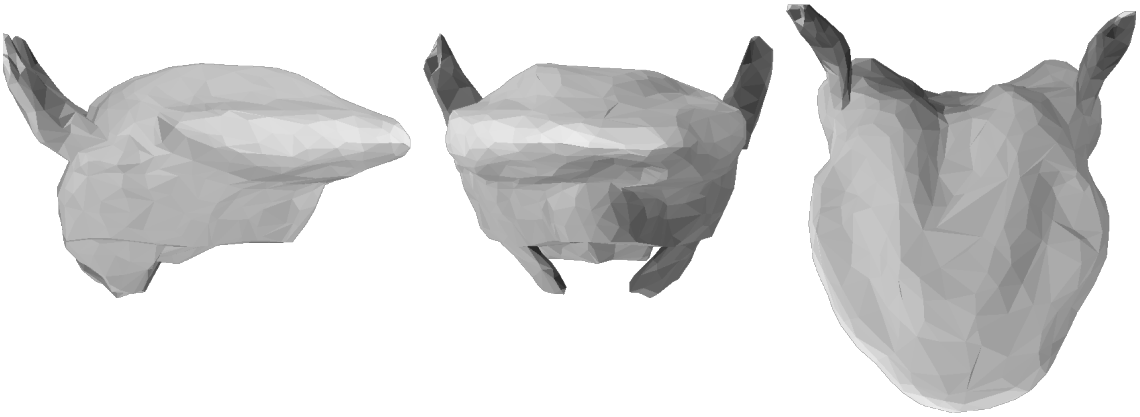

Figure 6.69: Minimum value of p_4, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

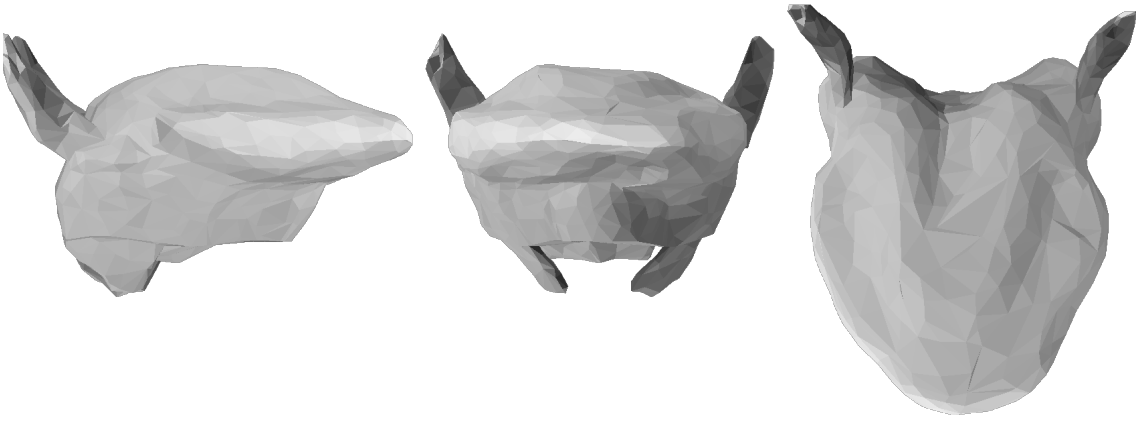

Figure 6.70: Maximum value of p_5, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

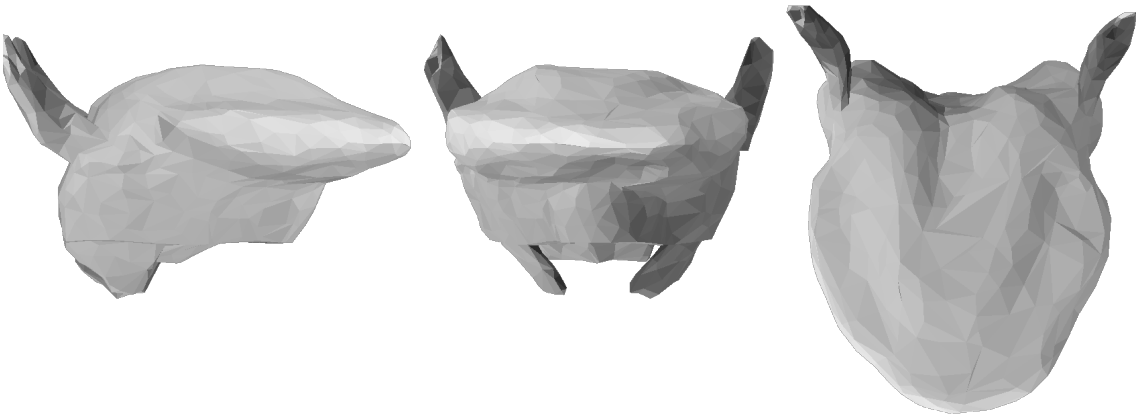

Figure 6.71: Minimum value of p_5, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

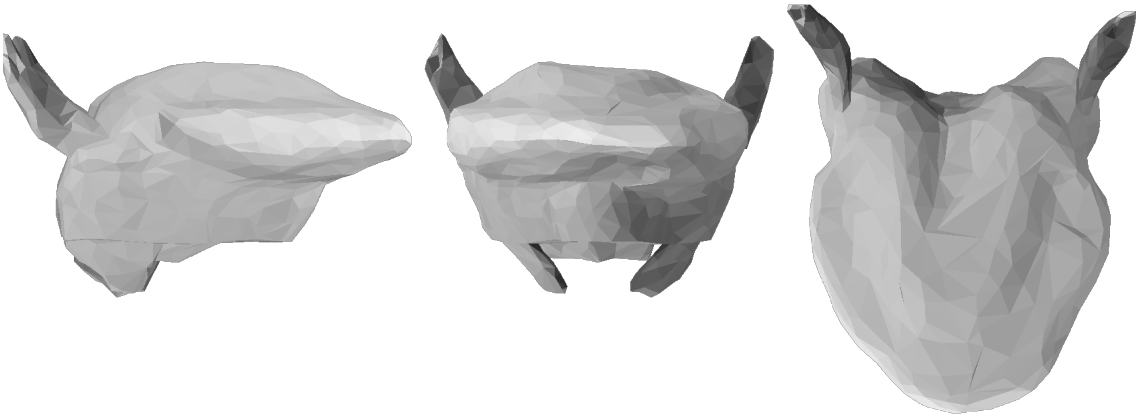

Figure 6.72: Maximum value of p_6, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

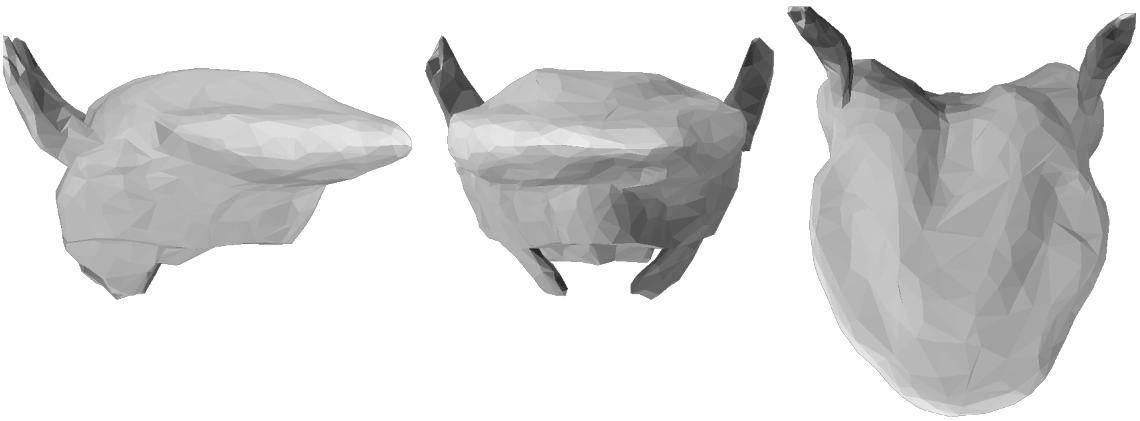

Figure 6.73: Minimum value of p_6, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

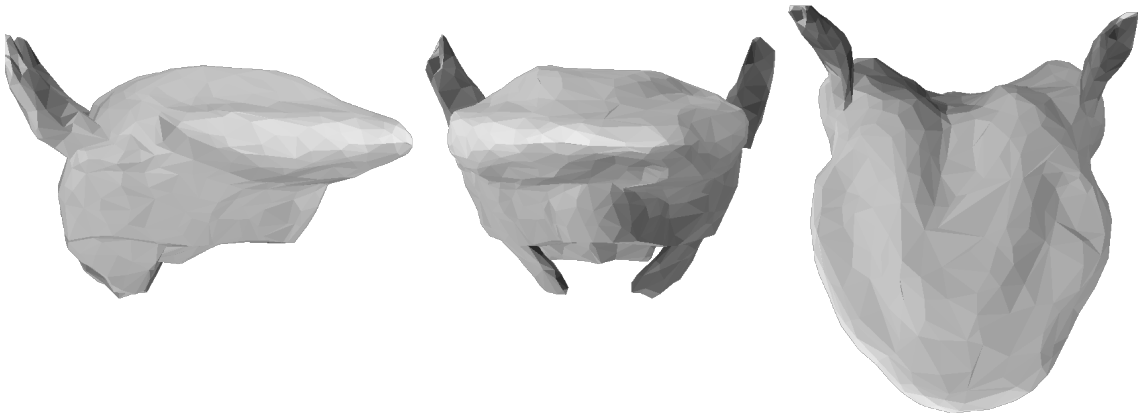

Figure 6.74: Maximum value of p_7, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

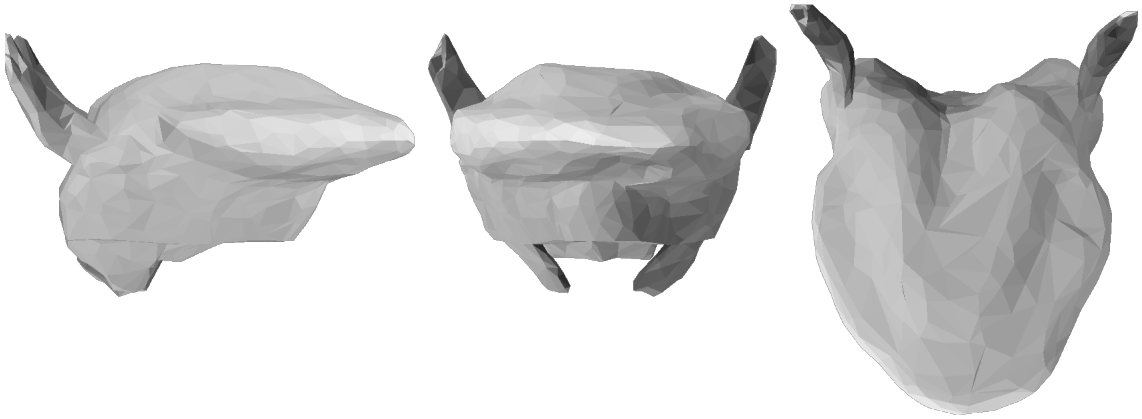

Figure 6.75: Minimum value of p_7, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only extrinsic muscles.

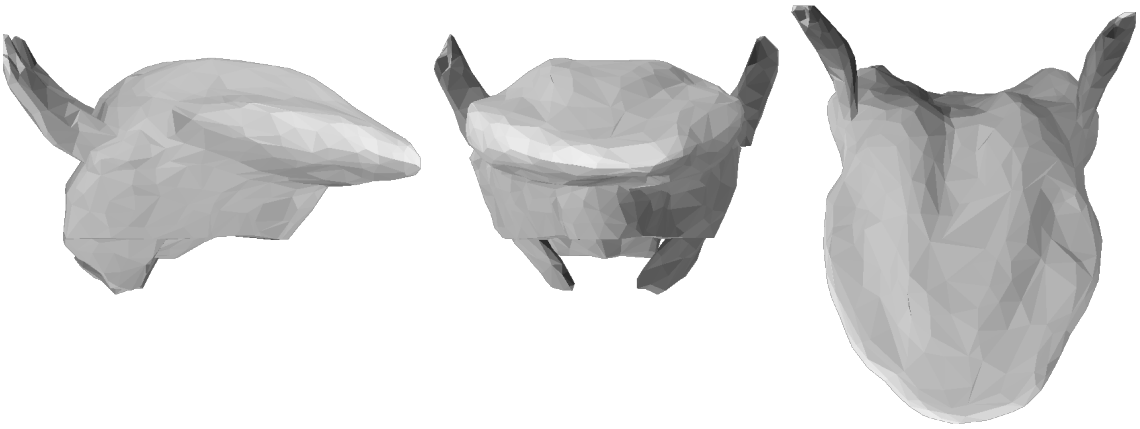

Figure 6.76: Maximum value of p_1, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

Figure 6.77: Minimum value of p_1, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

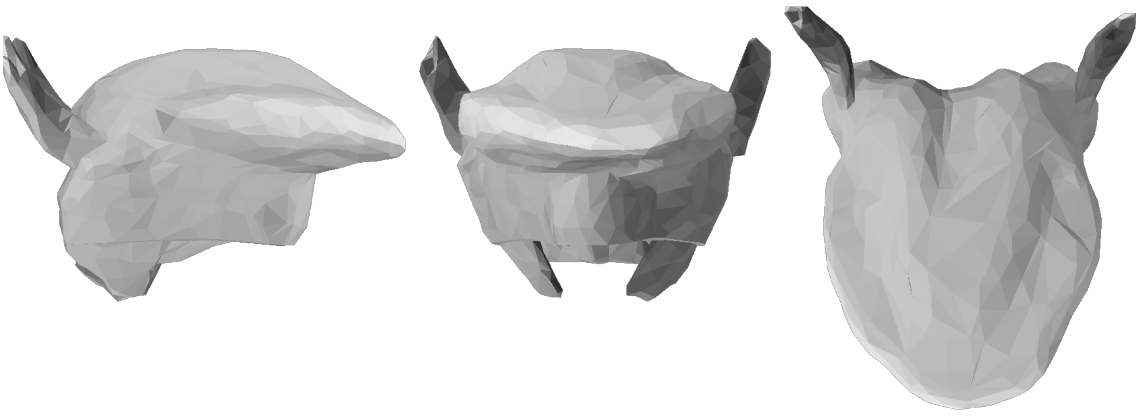

Figure 6.78: Maximum value of p_2, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

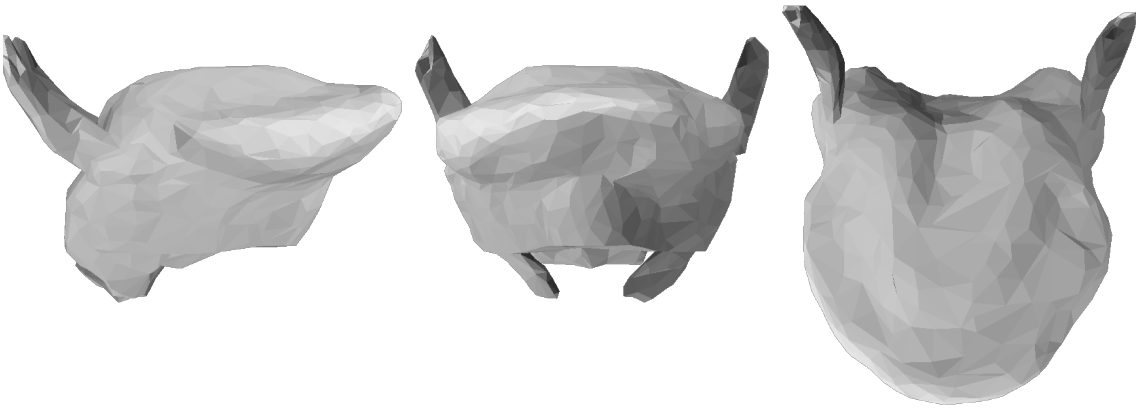

Figure 6.79: Minimum value of p_2, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

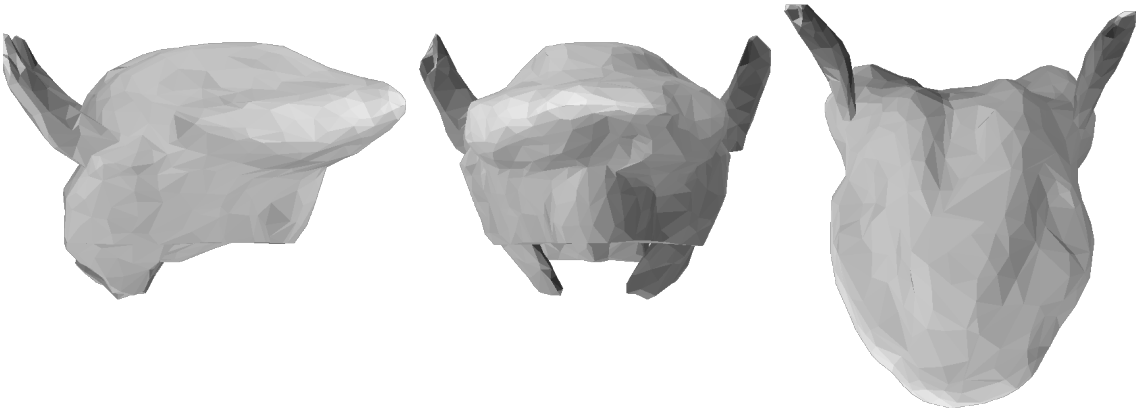

Figure 6.80: Maximum value of p_3, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

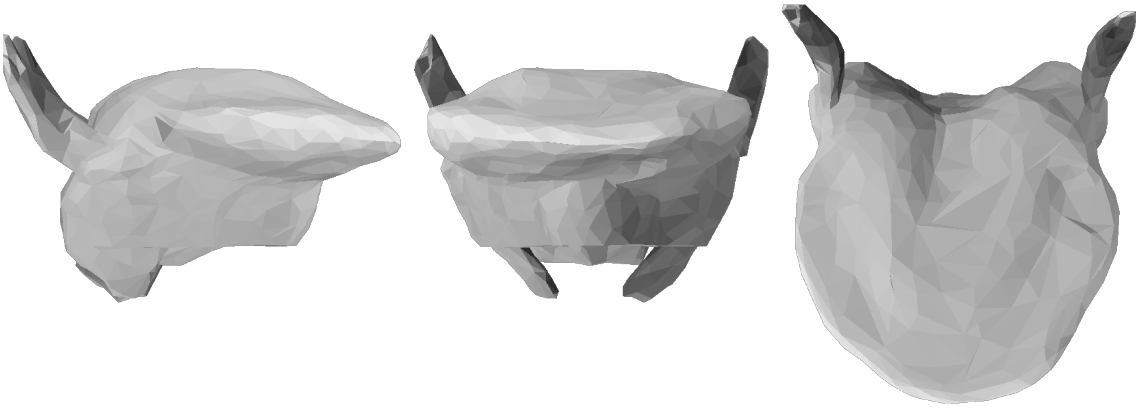

Figure 6.81: Minimum value of p_3, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

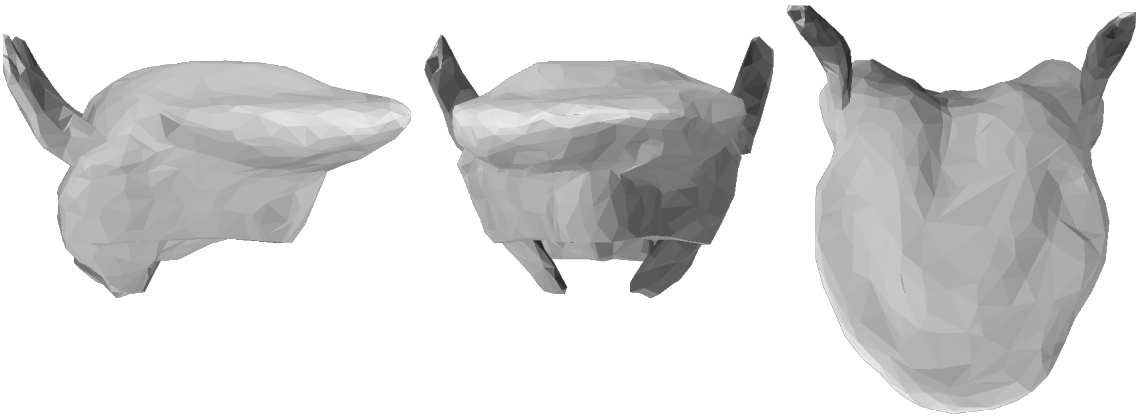

Figure 6.82: Maximum value of p_4, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

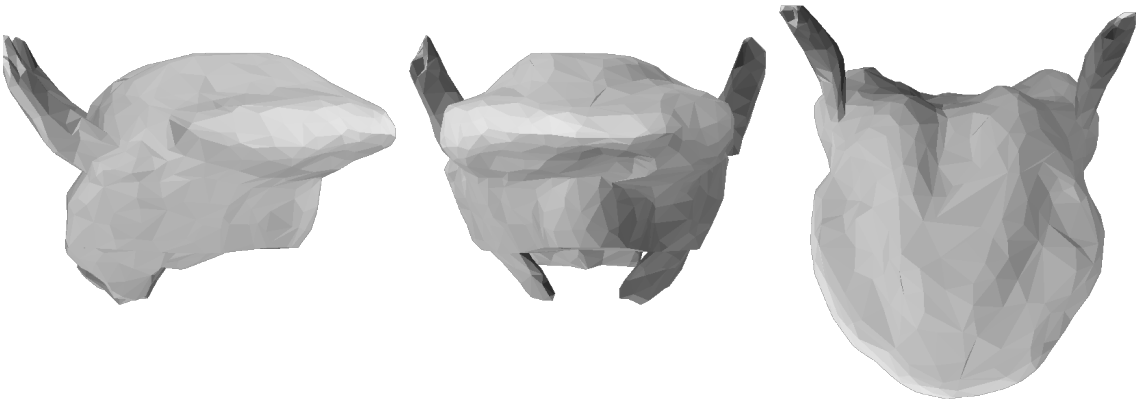

Figure 6.83: Minimum value of p_4, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

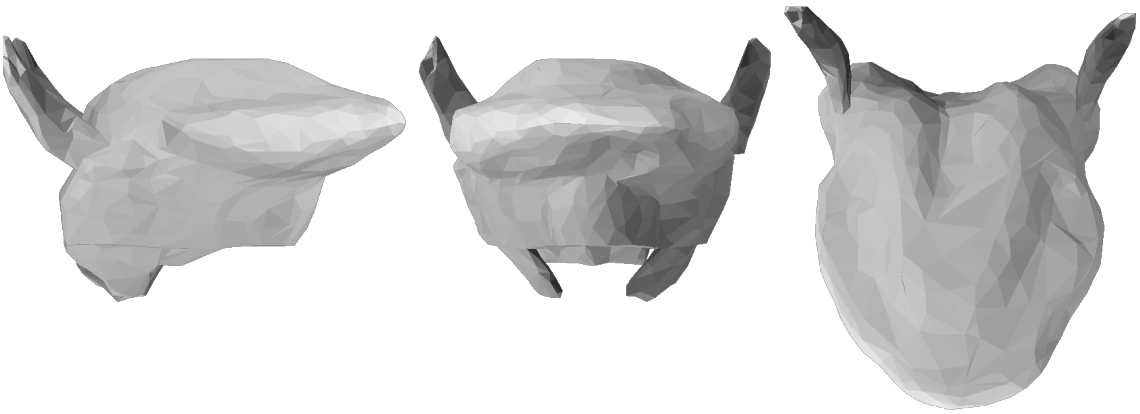

Figure 6.84: Maximum value of p_5, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

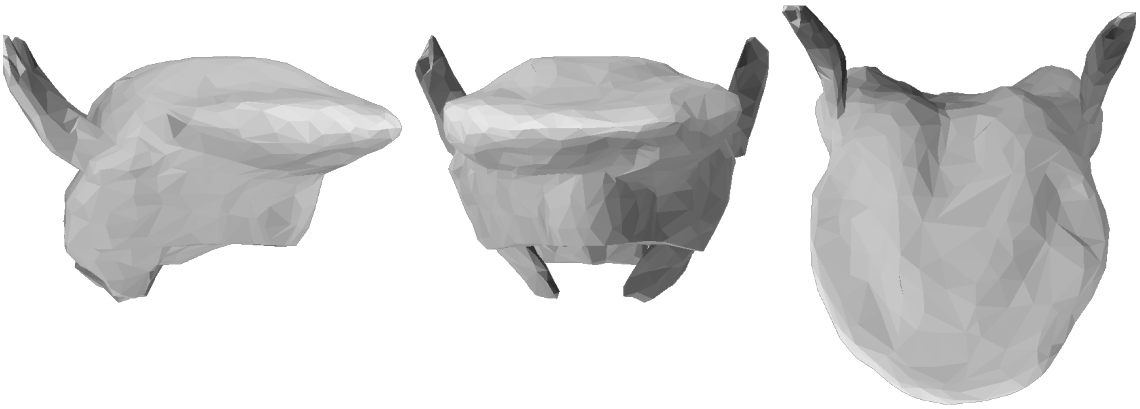

Figure 6.85: Minimum value of p_5, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

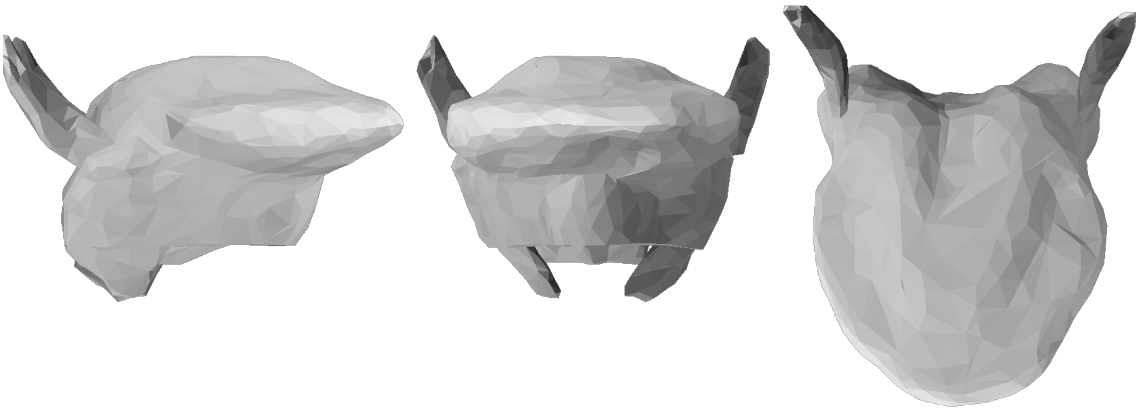

Figure 6.86: Maximum value of p_6, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

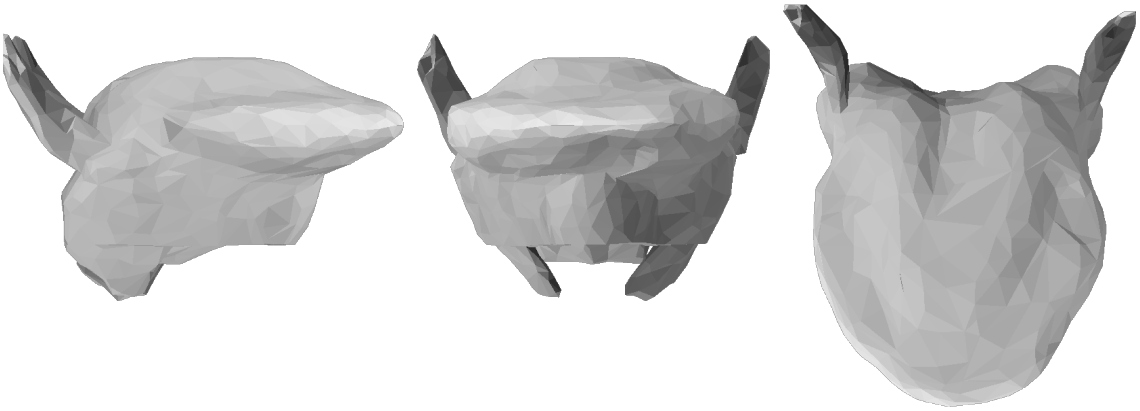

Figure 6.87: Minimum value of p_6, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

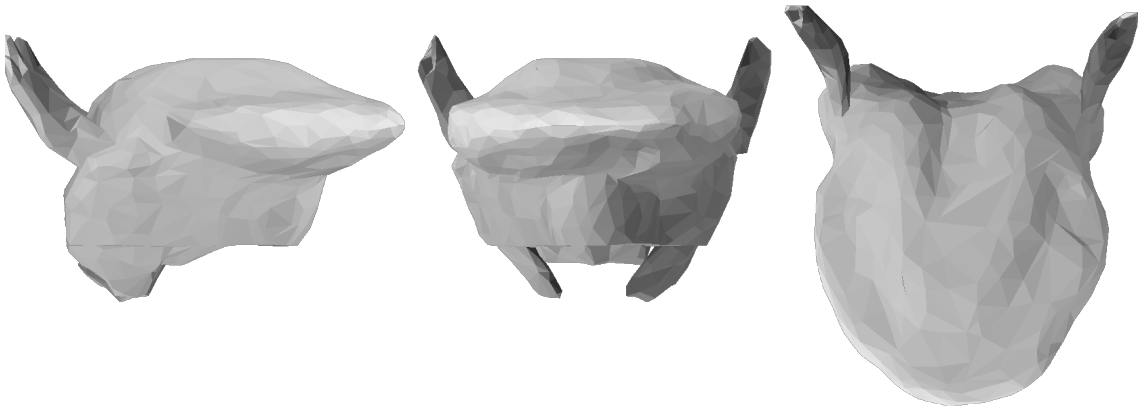

Figure 6.88: Maximum value of p_7, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

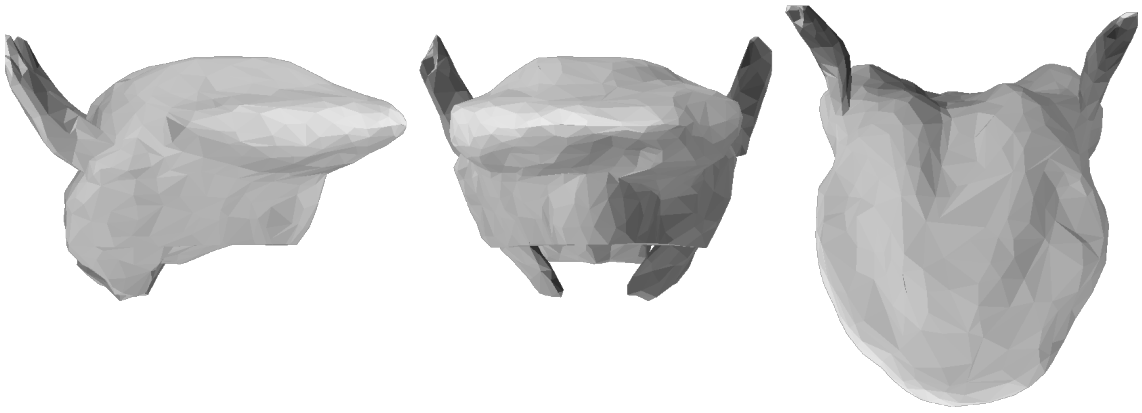

Figure 6.89: Minimum value of p_7, for tongue postures in $D_2^{\mathbb{P}_1}$ which involved contraction of only intrinsic muscles.

(following the procedure described in Section 5.2.3 and Section 5.2.4), and then to see if the resulting parameters have distinct and mutually exclusive physical interpretations. This was done. The results are given in Figures 6.90 through 6.103, which provide the postures corresponding to maximum and minimum parameter values for the extrinsic data, and in Figures 6.104 through 6.117, which do the same for the intrinsic data.

The parameter space calculated for the intrinsic postures was very similar to $\mathbb{P}_1$. As in $\mathbb{P}_1$, variation along p_1 has to do with tongue tip elevation and depression. For the intrinsic muscles, p_2 corresponds to tongue tip movement, as in $\mathbb{P}_1$. p_3 involves bunching the tongue body. p_4 is protrusion and elongation of the tongue. Movement along p_5 is the difficult-to-describe raising of the lateral portion of the tongue. p_6 is a slight movement in the tongue root. In contradistinction to $\mathbb{P}_1$, variation along p_7 produces a slight movement near the rear of the tongue where styloglossus inserts.

For the extrinsic muscles, the principal components tend to recapitulate single muscle contractions. Thus p_1 corresponds to (or is very similar to) contraction of genioglossus. p_2 corresponds to contraction of hyoglossus. p_3 and p_4 correspond to contraction of styloglossus. The remaining components produced small, non-specific postural adjustments.

In this approach as well, then, predictions from the present model do not provide evidence that supports the hypothesis that the intrinsic and extrinsic muscles have fundamentally distinct biomechanical roles.

6.3 Model of flaccid dysarthria

This section contains a simulation of a tongue with a pathological condition, flaccid dysarthria. This illustrates the potential for using the model to simulate disease conditions more generally. The tongue is innervated by the hypoglossal nucleus, the axons of which run through the hypoglossal nerve. If one of the hypoglossal nuclei or the hypoglossal nerves is lesioned, motor input to the ipsilateral half of the tongue is eliminated. Such lesions can arise from a multitude of sources: infections,

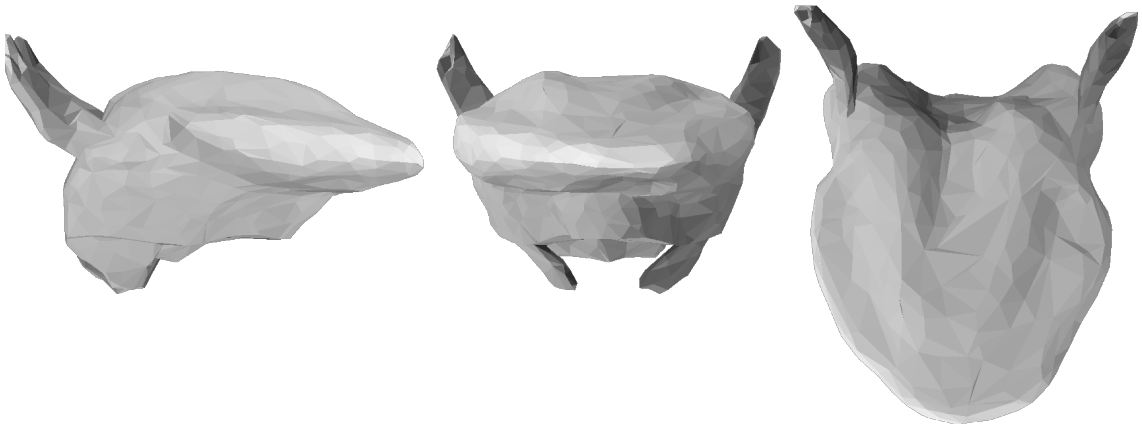

Figure 6.90: Maximum value of p_1, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

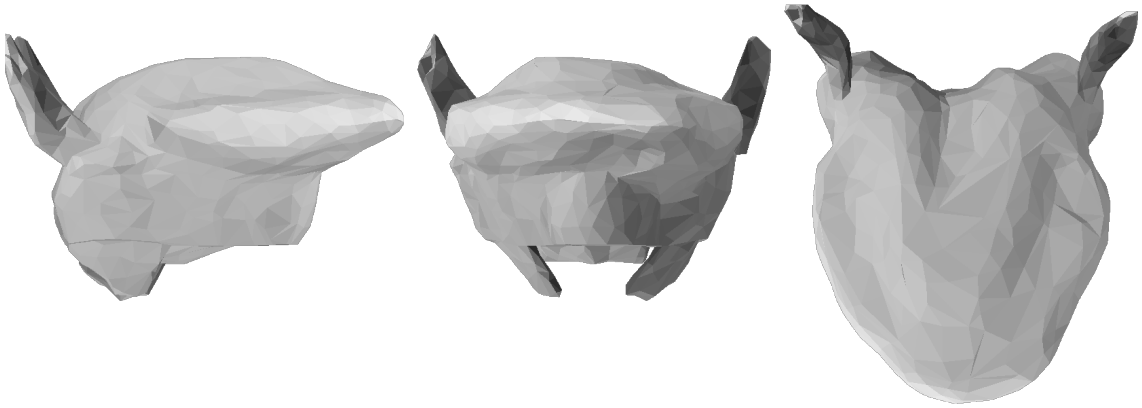

Figure 6.91: Minimum value of p_1, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

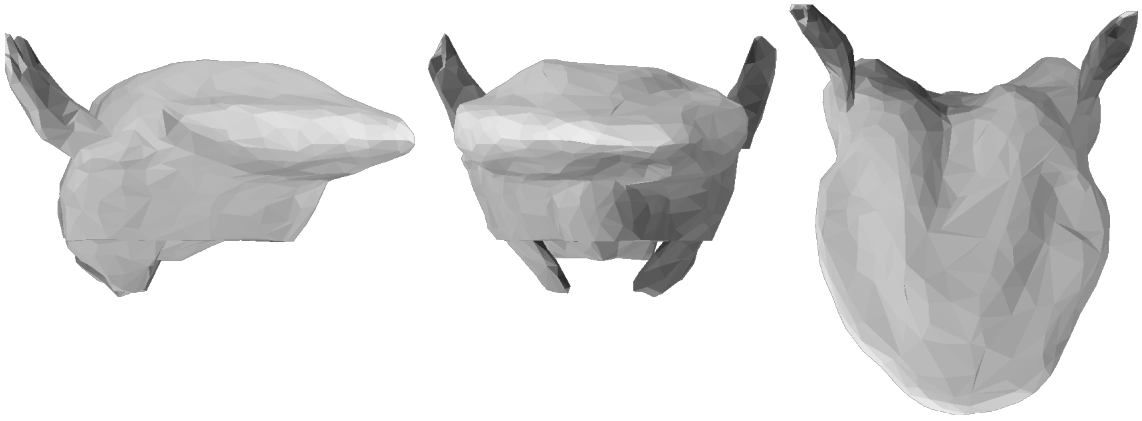

Figure 6.92: Maximum value of p_2, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

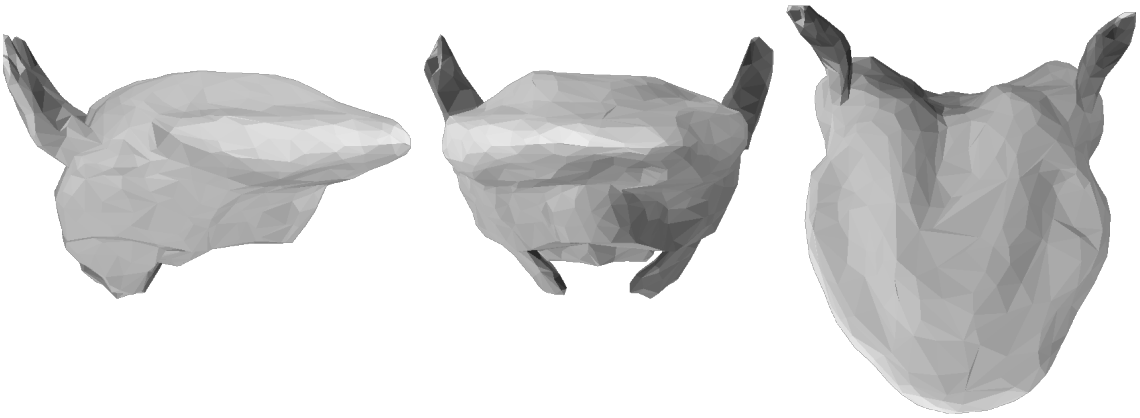

Figure 6.93: Minimum value of p_2, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

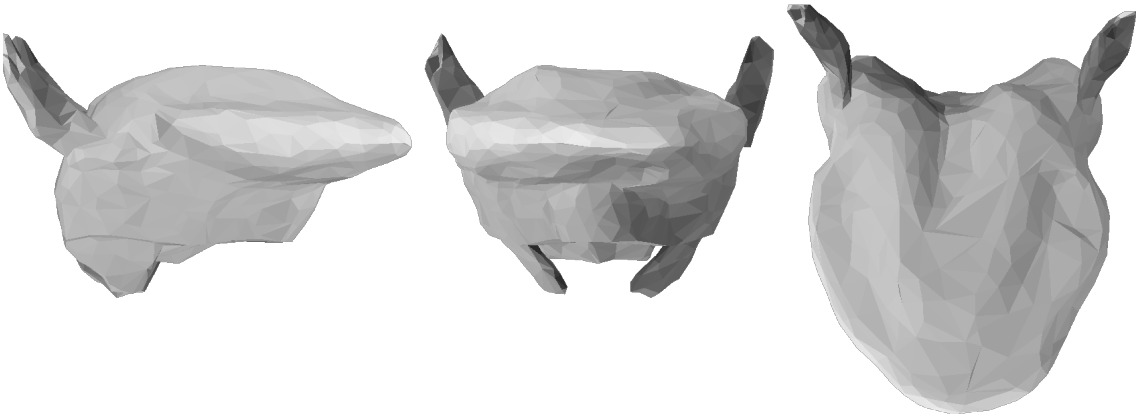

Figure 6.94: Maximum value of p_3, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

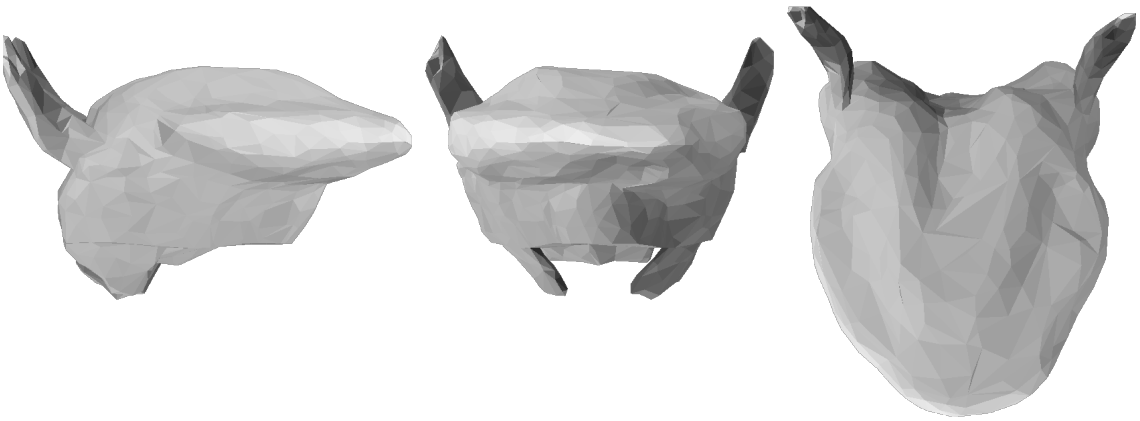

Figure 6.95: Minimum value of p_3, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

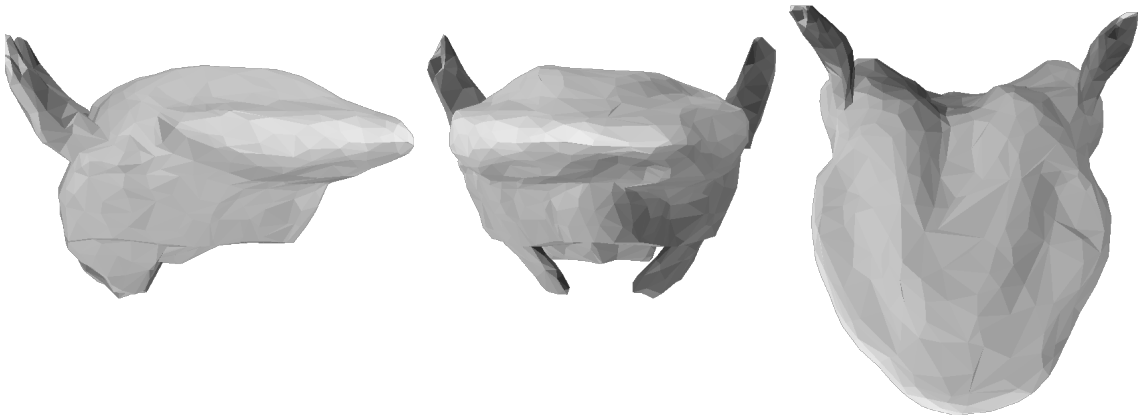

Figure 6.96: Maximum value of p_4, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

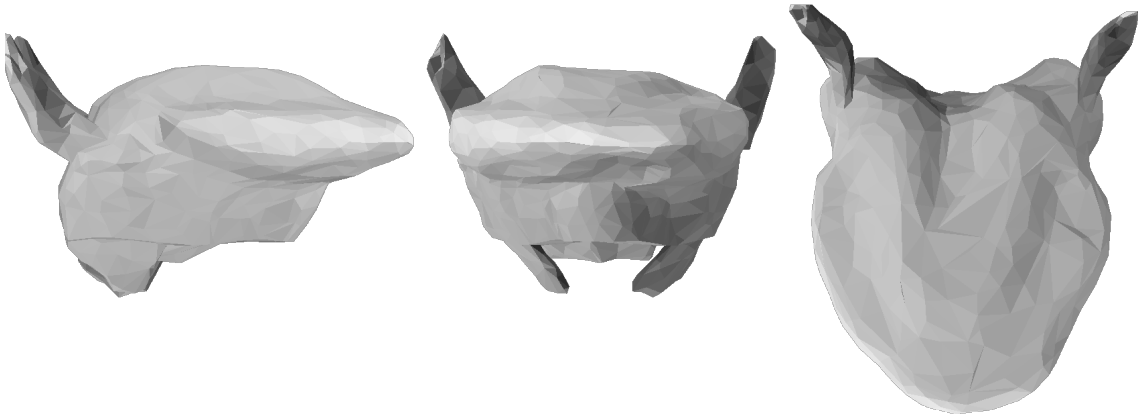

Figure 6.97: Minimum value of p_4, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

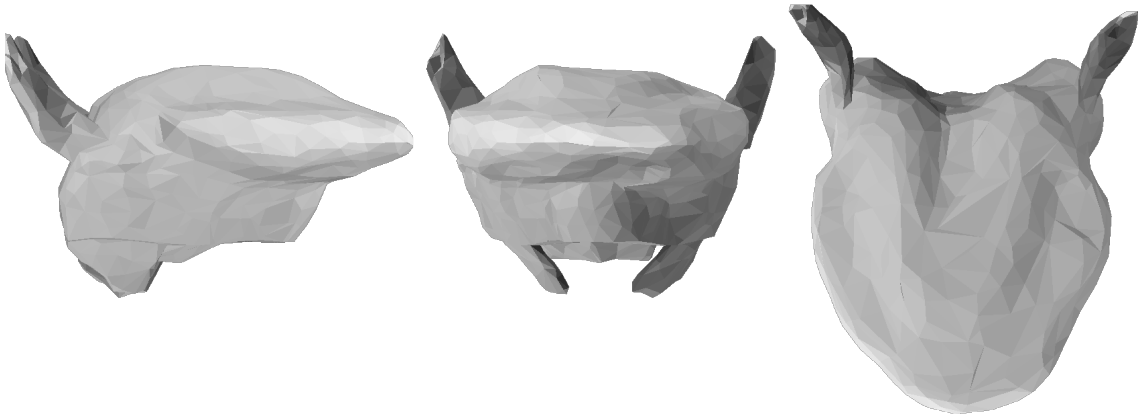

Figure 6.98: Maximum value of p_5, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

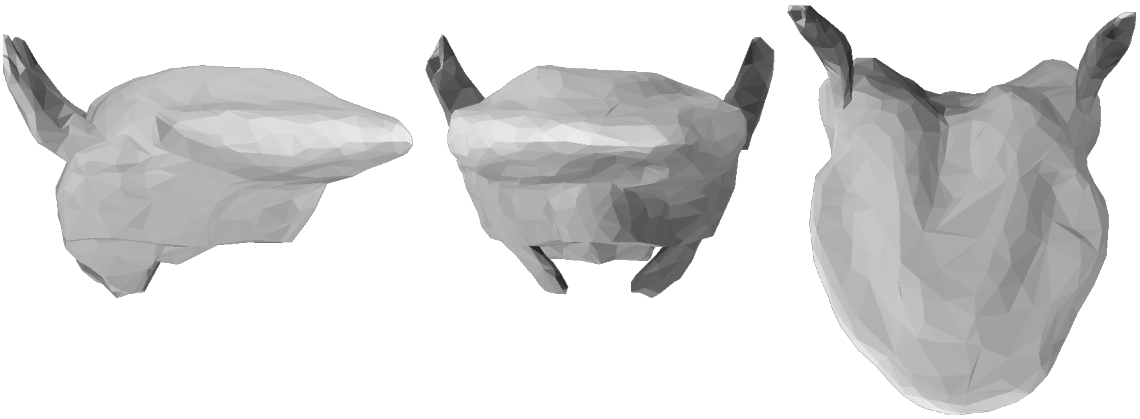

Figure 6.99: Minimum value of p_5, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

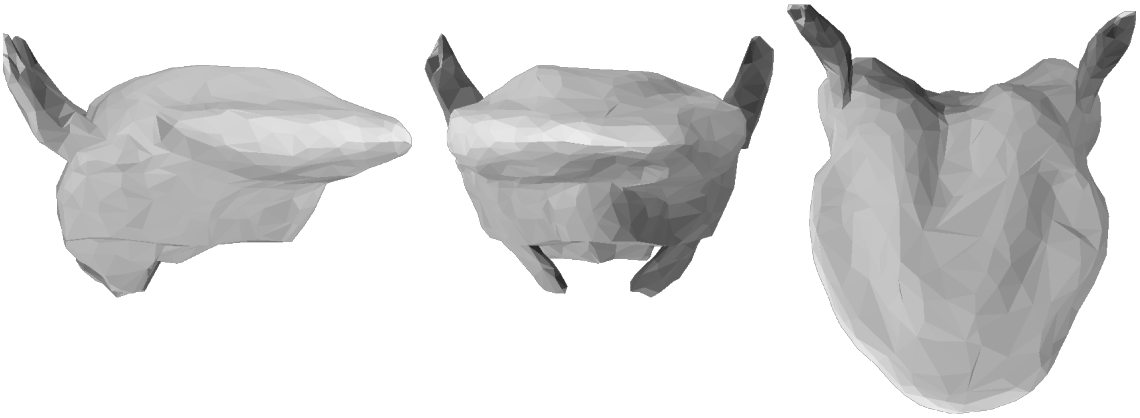

Figure 6.100: Maximum value of p_6, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

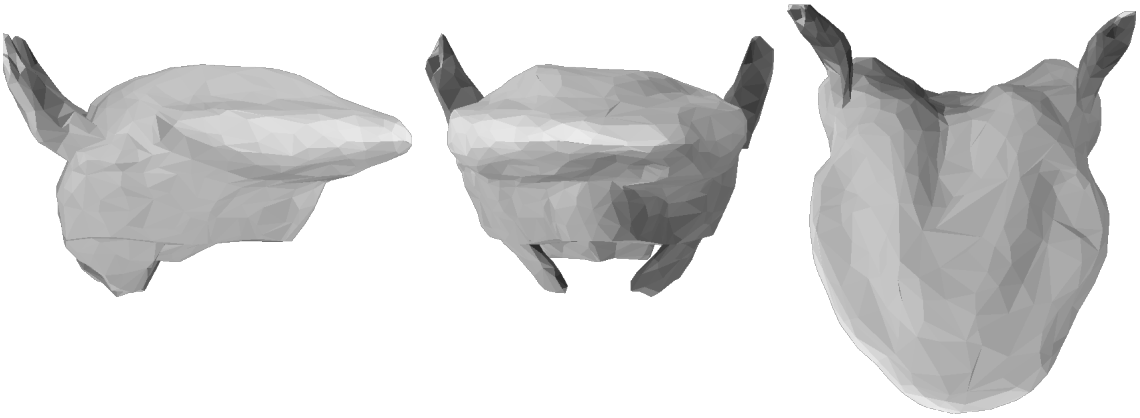

Figure 6.101: Minimum value of p_6, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

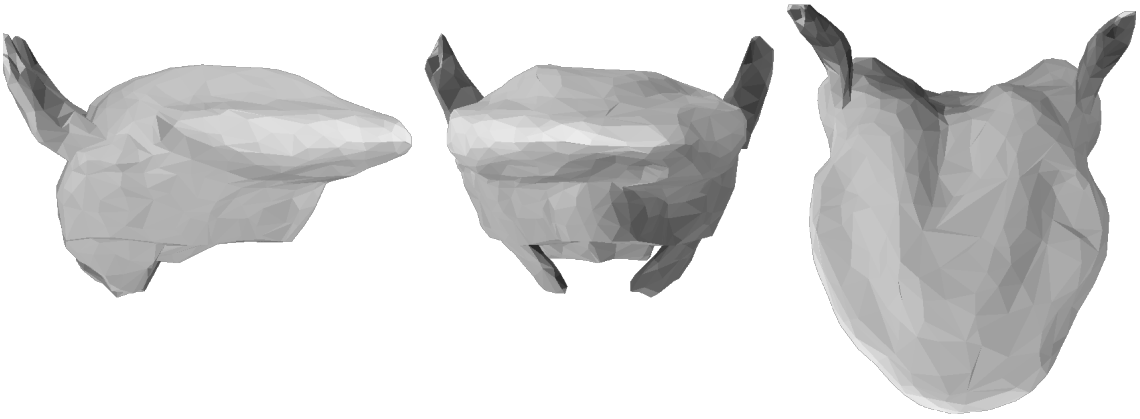

Figure 6.102: Maximum value of p_7, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

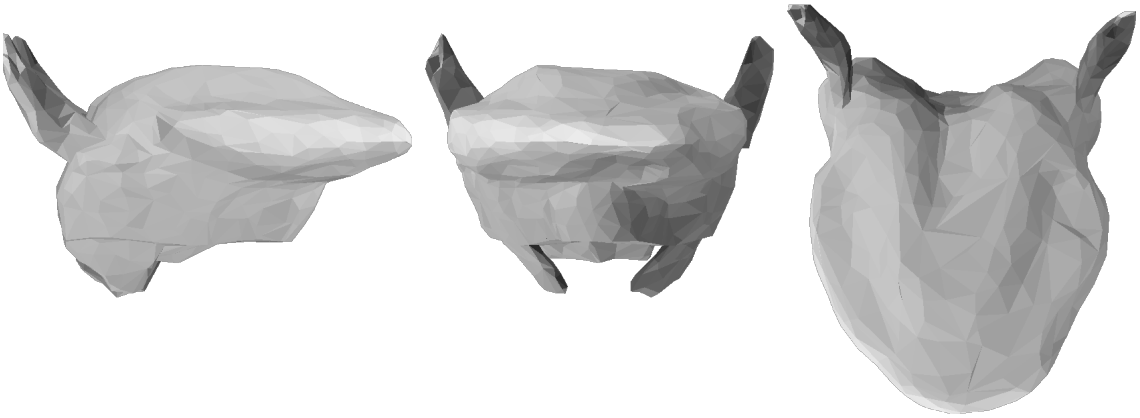

Figure 6.103: Minimum value of p_7, for tongue postures which involved contraction of only extrinsic muscles, in a parameter space derived from that data set.

Figure 6.104: Maximum value of p_1, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

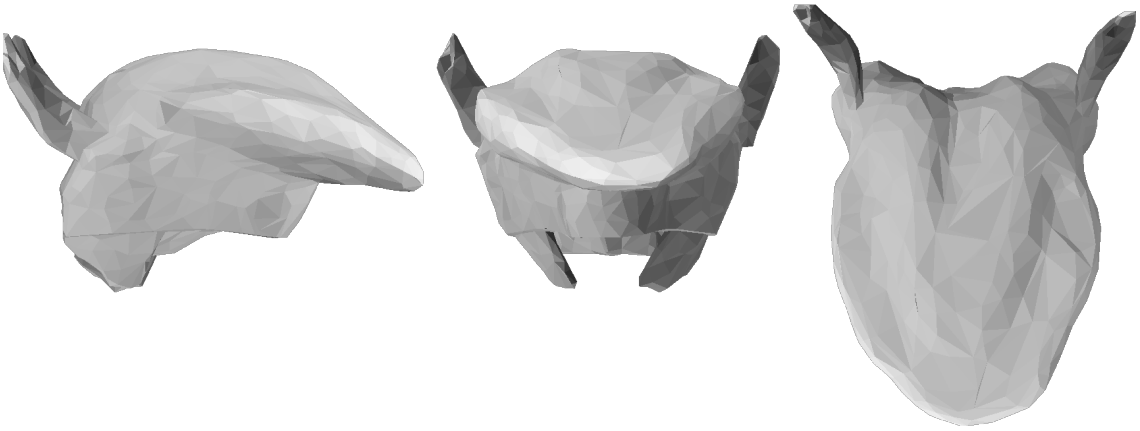

Figure 6.105: Minimum value of p_1, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

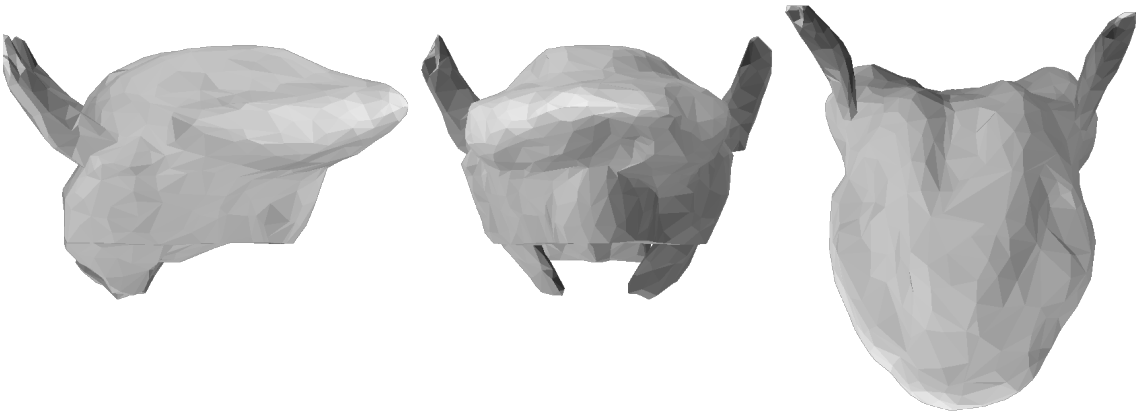

Figure 6.106: Maximum value of p_2, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

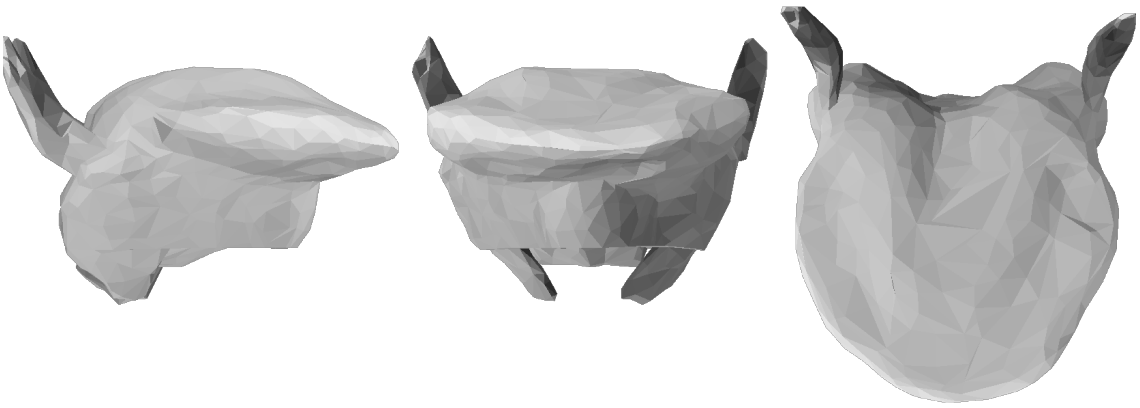

Figure 6.107: Minimum value of p_2, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

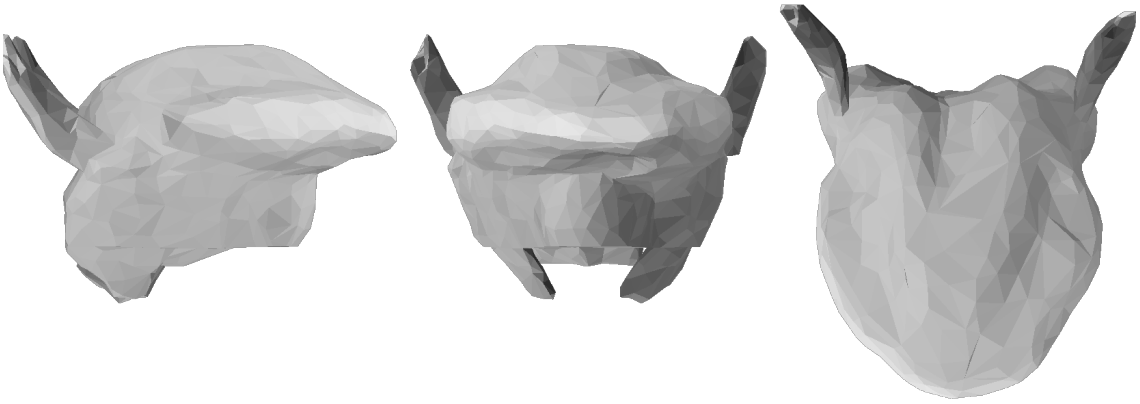

Figure 6.108: Maximum value of p_3, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

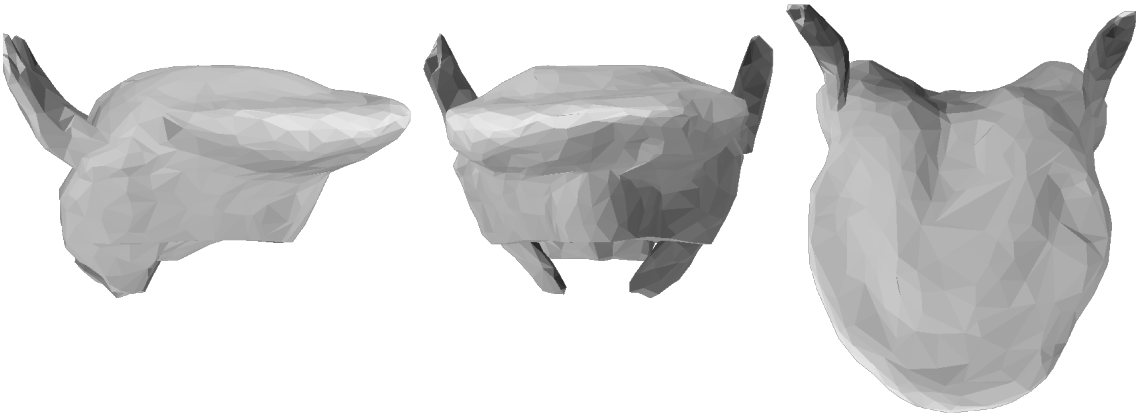

Figure 6.109: Minimum value of p_3, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

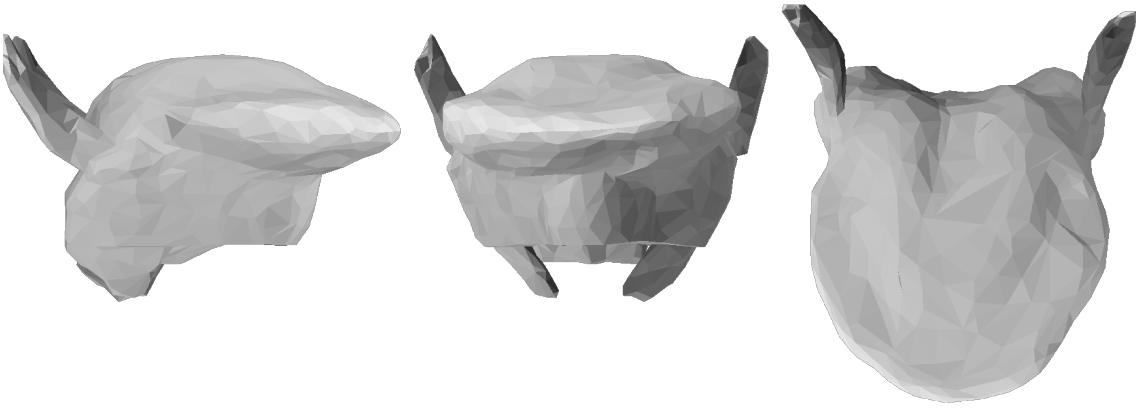

Figure 6.110: Maximum value of p_4, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

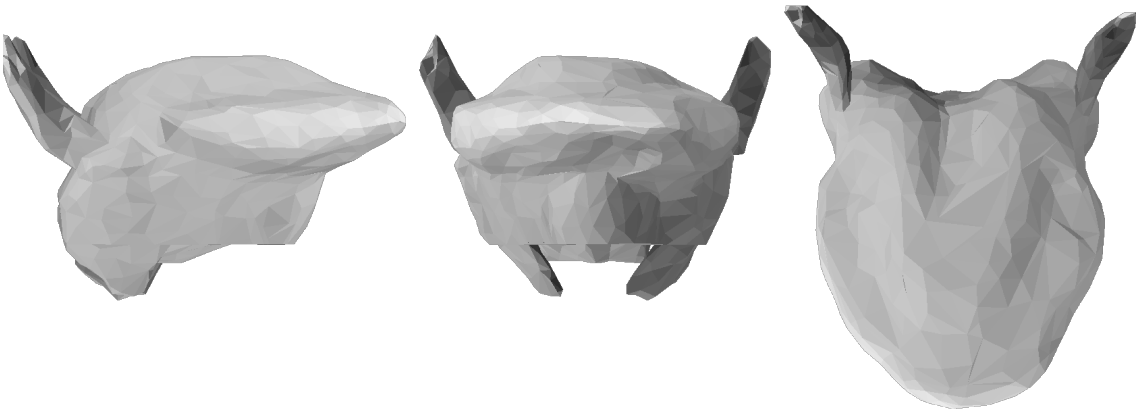

Figure 6.111: Minimum value of p_4, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

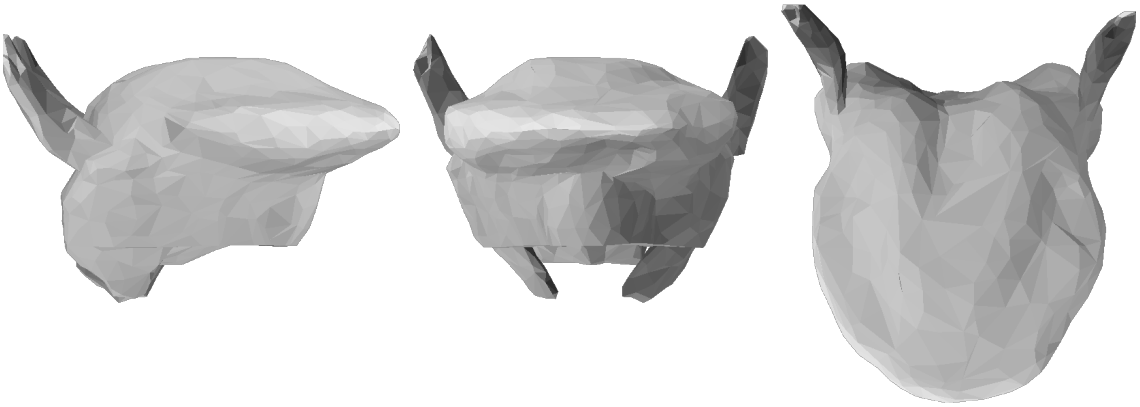

Figure 6.112: Maximum value of p_5, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

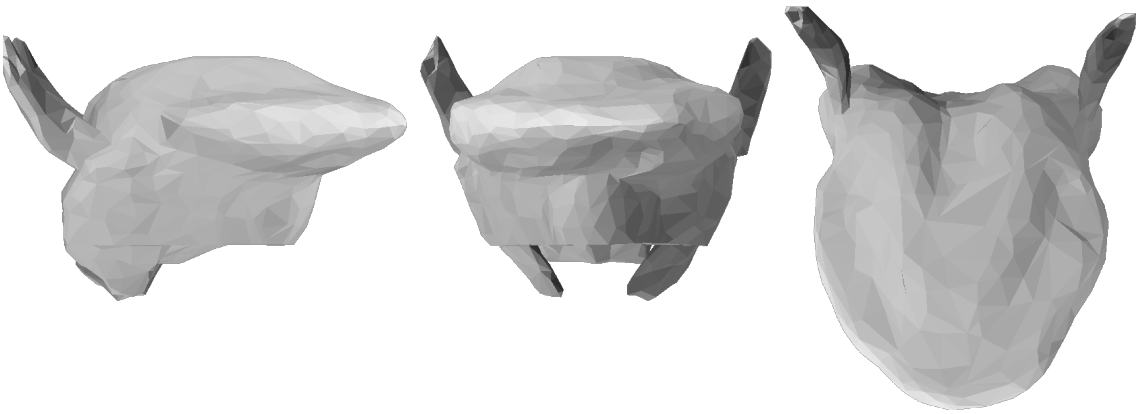

Figure 6.113: Minimum value of p_5, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

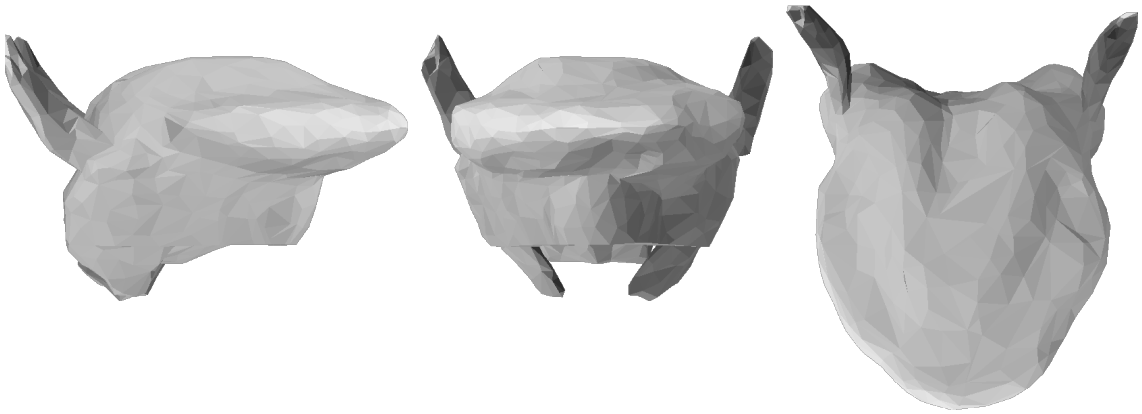

Figure 6.114: Maximum value of p_6, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

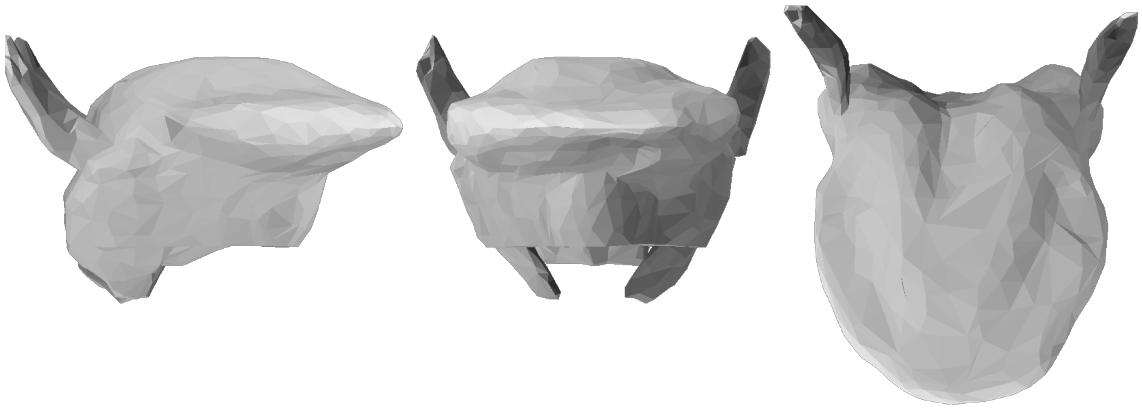

Figure 6.115: Minimum value of p_6, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

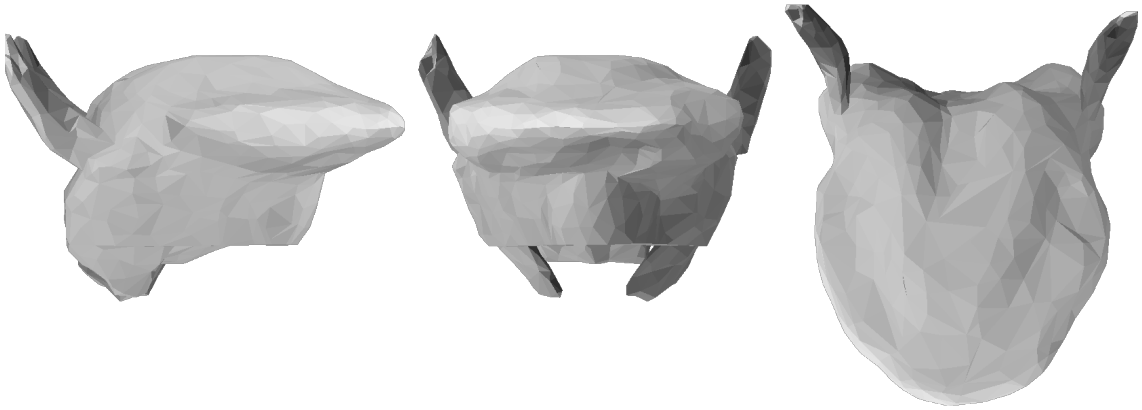

Figure 6.116: Maximum value of p_7, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

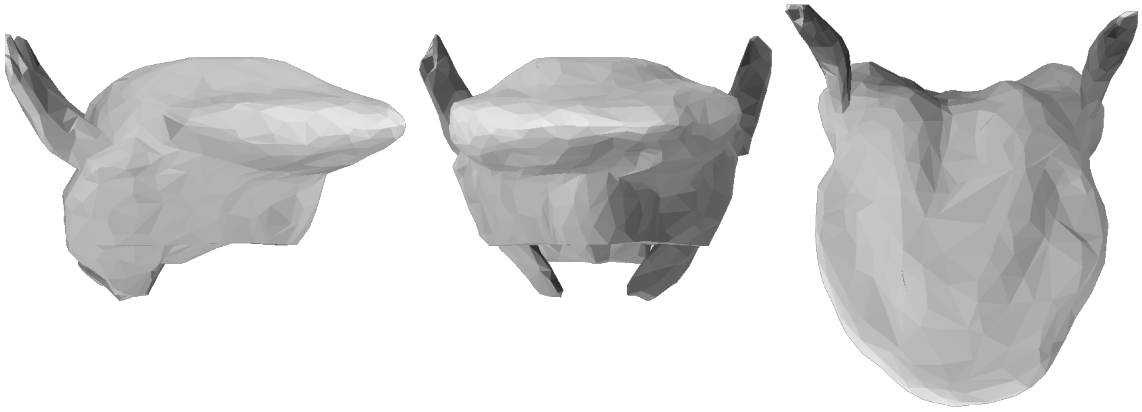

Figure 6.117: Minimum value of p_7, for tongue postures which involved contraction of only intrinsic muscles, in a parameter space derived from that data set.

tumors, trauma, and surgical error (Duffy, 2005). The primary consequence of such an injury for speech production is imprecise articulation of lingual consonants. Duffy notes that coronal consonants are effected to a greater extent than dorsal consonants. Speakers are typically able to compensate for decreased tongue mobility by using the jaw to position the tongue. This pathological condition can be modeled with a biomechanical tongue model, simply by eliminating activation in one half of the tongue.

Simulations were run with the right side of the tongue (an arbitrary choice) receiving no motor input. Two data sets were formed, on analogy with D_1 and D_2. D_4 is the hemiplegic equivalent of D_1, the set of postures generated from A_1; D_5 is the hemiplegic equivalent of D_2, the set of postures generate from A_2. Note that $A_1 \subset A_2$.[9] The data sets were analyzed with the same techniques as were used in Chapter 5; the remainder of this section assumes familiarity with those techniques.

An exhaustive study of hemiplegia was not made, since pathological tongue movement is not a major focus of the dissertation. In fact, just one hypothesis is tested: that a hemiplegic tongue would require different muscular activation from

[9]Recall from Section 5.1 that A_1 is the set of activation patterns with two levels of activation, and that A_2 is the set of activation patterns with three levels of activation.

a normal tongue in order to produce (approximately) equivalent tongue postures.[10]

6.3.1 Principal component analysis

A principal component analysis was performed on D_4, yielding a basis $\mathbb{P}_4$. As with D_1 and D_2, seven parameters were required to account for greater than 99% of the variance. The variance accounted for by each of the first ten components in shown in Figure 6.118. Parameters are illustrated at their maximum and minimum values in Figures 6.119 through 6.132. In general, the parameters in $\mathbb{P}_4$ are the hemiplegic variants of those in $\mathbb{P}_5$: the deformations are similar, except that the effect is mostly unilateral, and the gestures are less extreme. Also—apparently by coincidence—p_4 in $\mathbb{P}_4$ corresponds to p_5 in $\mathbb{P}_1$, and p_5 in $\mathbb{P}_4$ corresponds to p_4 of $\mathbb{P}_1$.

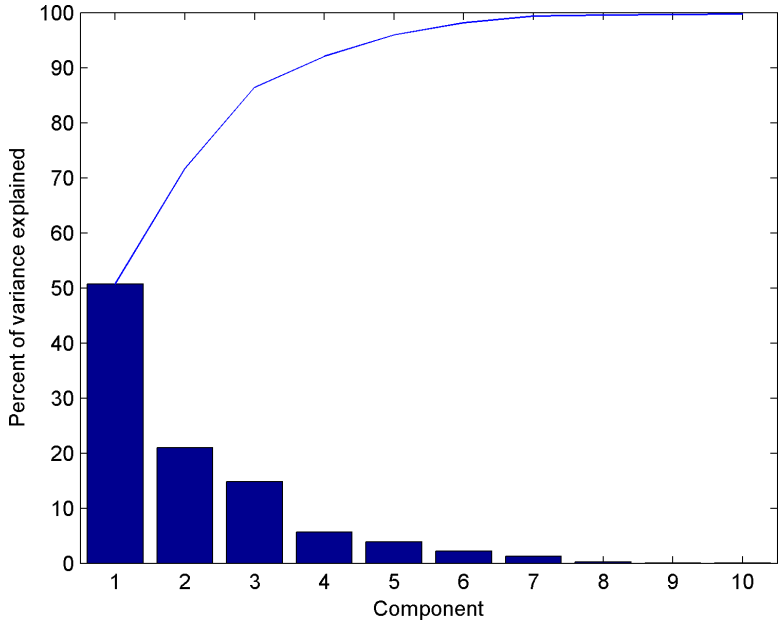

Figure 6.118: Pareto diagram of variance in D_4 accounted for by each parameter of $\mathbb{P}_4$.

[10]The intuition here is that moving a tongue that receives bilateral activation is, mechanically, a very different task from moving one that has bilateral activation, just as, say, crawling would involve very different muscular activation levels in a healthy person and a person afflicted with hemiplegia.

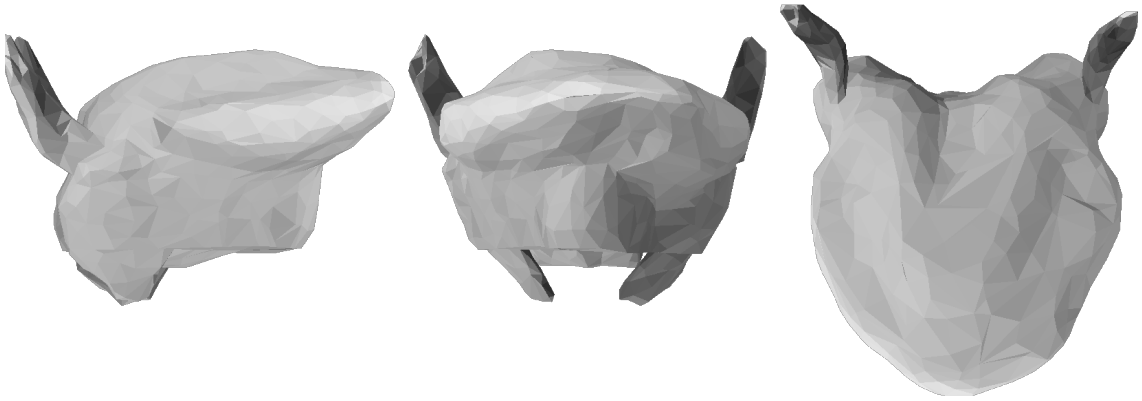

Figure 6.119: $D_4^{\mathbb{P}4}$: p_1 at minimum value.

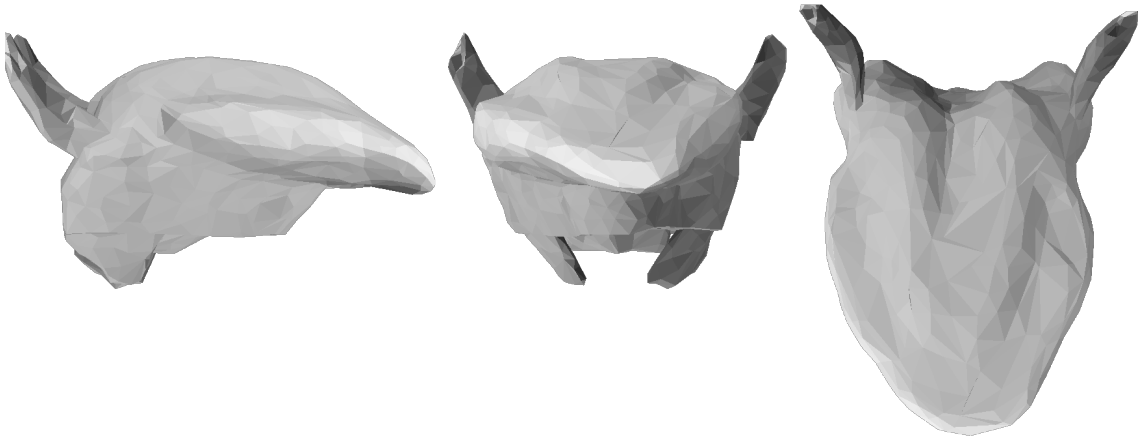

Figure 6.120: $D_4^{\mathbb{P}4}$: p_1 at maximum value.

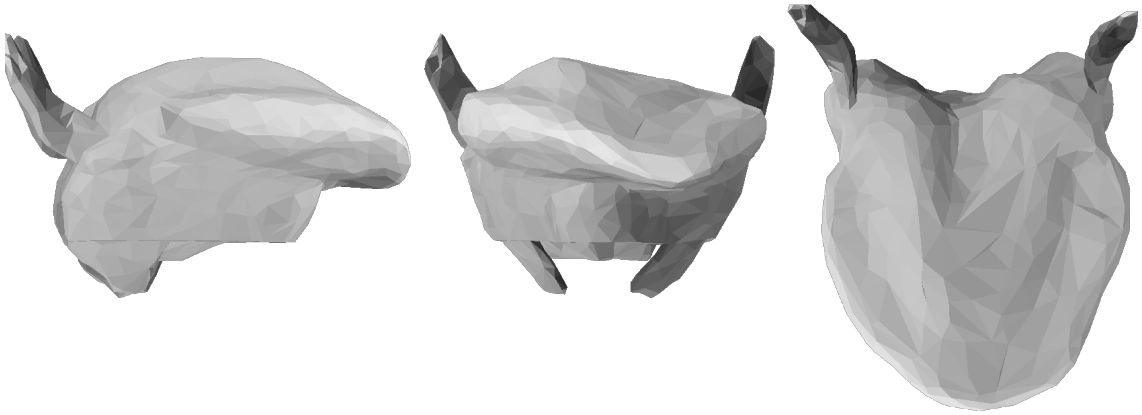

Figure 6.121: $D_4^{\mathbb{P}4}$: p_2 at minimum value.

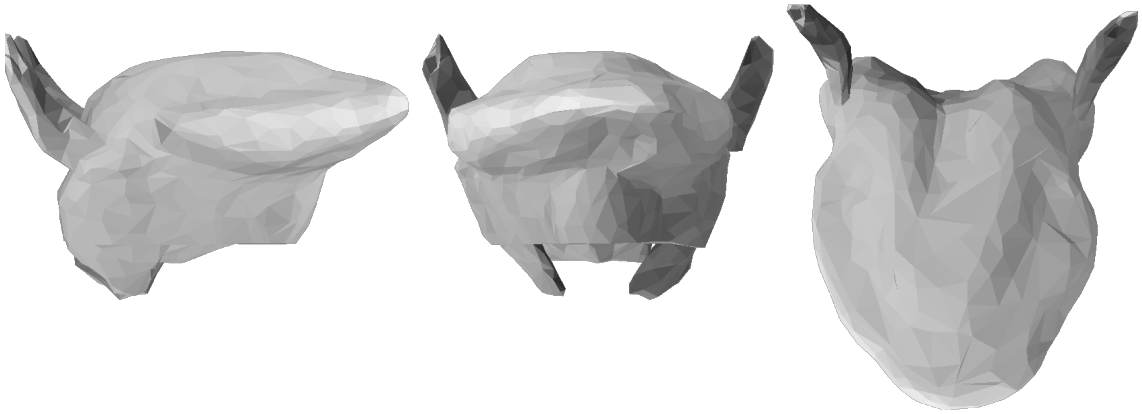

Figure 6.122: $D_4^{\mathbb{P}4}$: p_2 at maximum value.

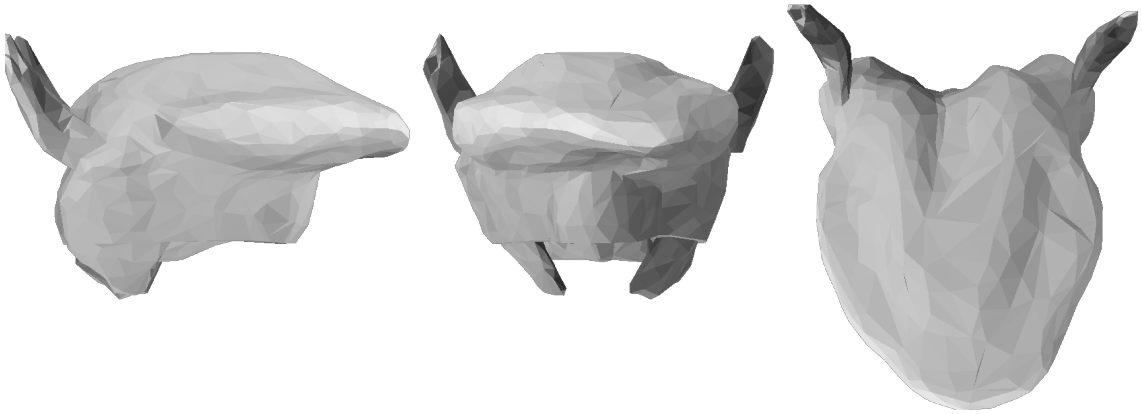

Figure 6.123: $D_4^{\mathbb{P}4}$: p_3 at minimum value.

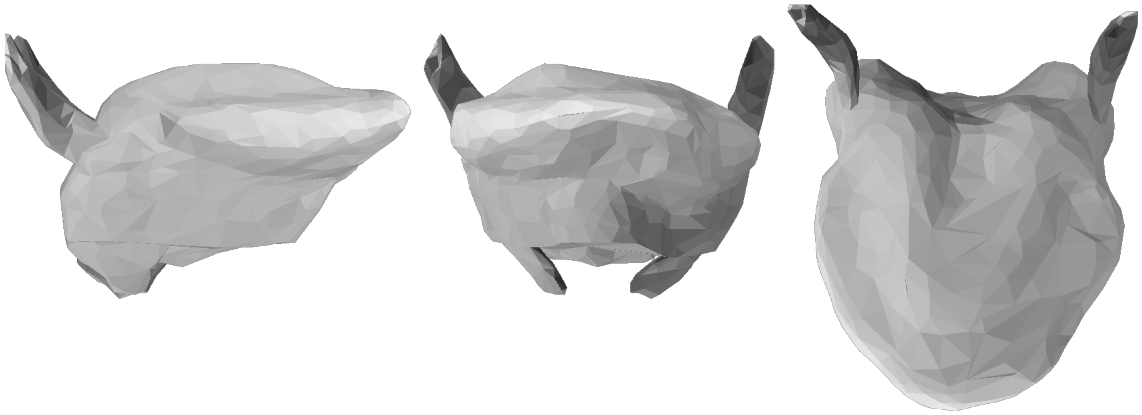

Figure 6.124: $D_4^{\mathbb{P}4}$: p_3 at maximum value.

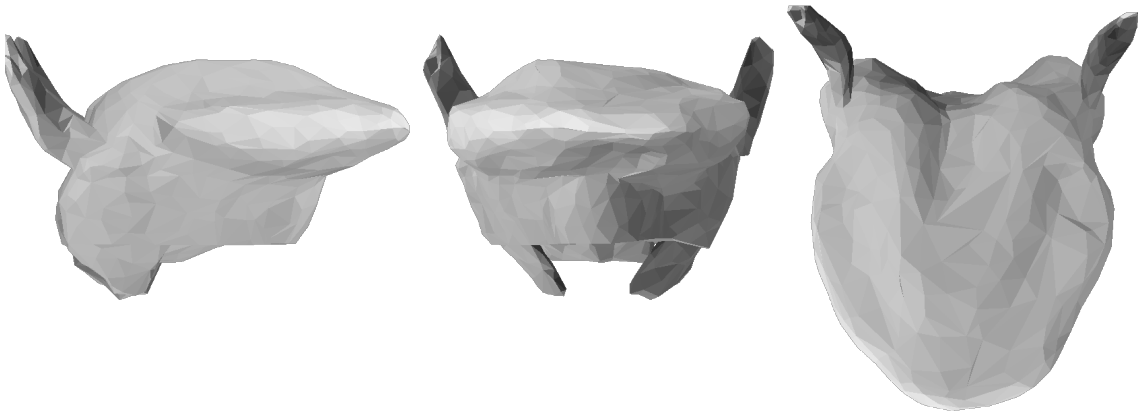

Figure 6.125: $D_4^{\mathbb{P}4}$: p_4 at minimum value.

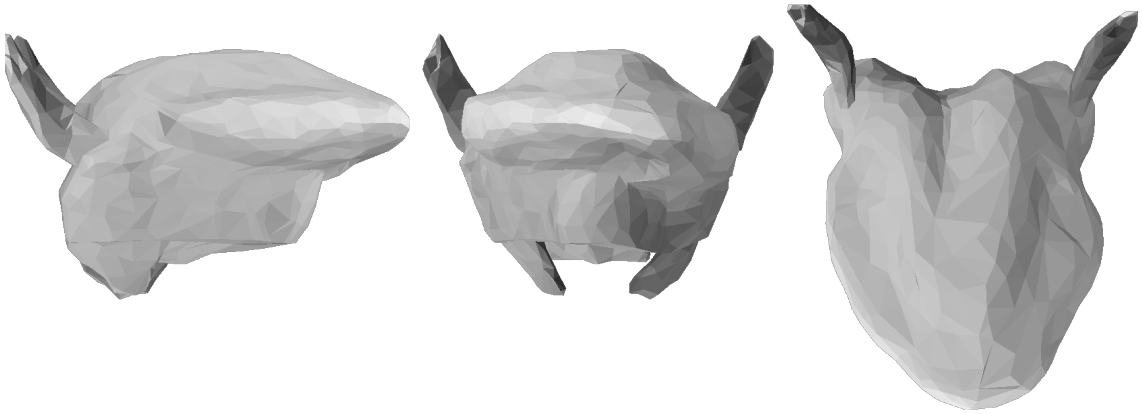

Figure 6.126: $D_4^{\mathbb{P}4}$: p_4 at maximum value.

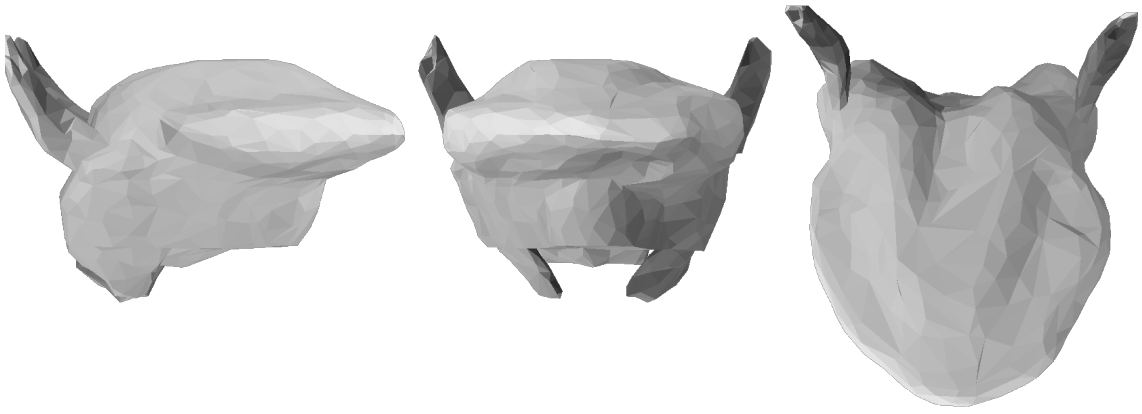

Figure 6.127: $D_4^{\mathbb{P}4}$: p_5 at minimum value.

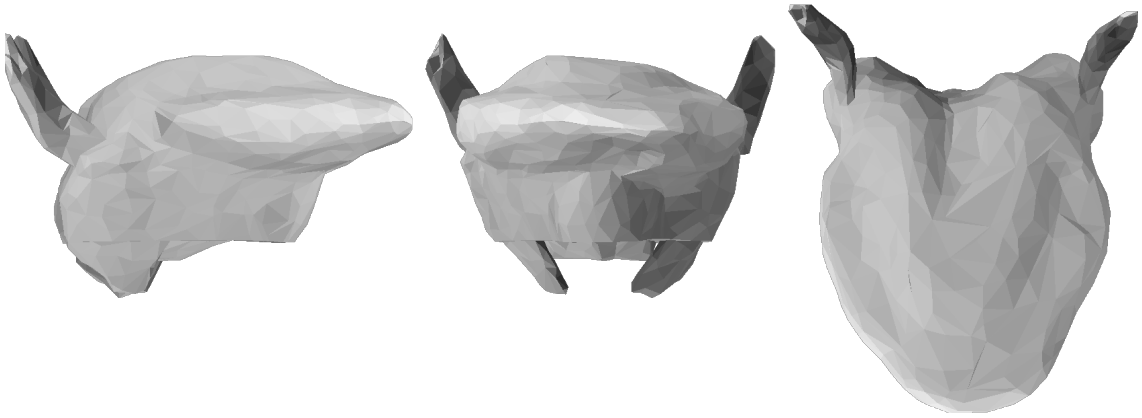

Figure 6.128: $D_4^{\mathbb{P}4}$: p_5 at maximum value.

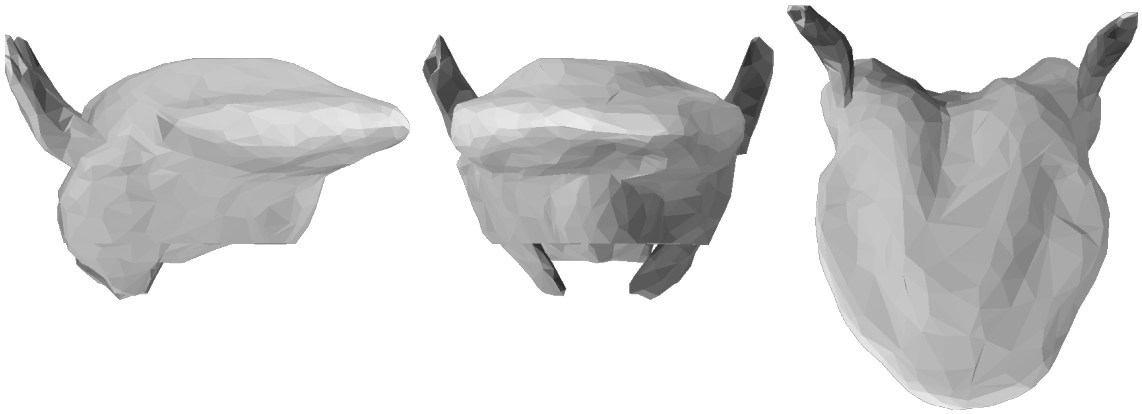

Figure 6.129: $D_4^{\mathbb{P}4}$: p_6 at minimum value.

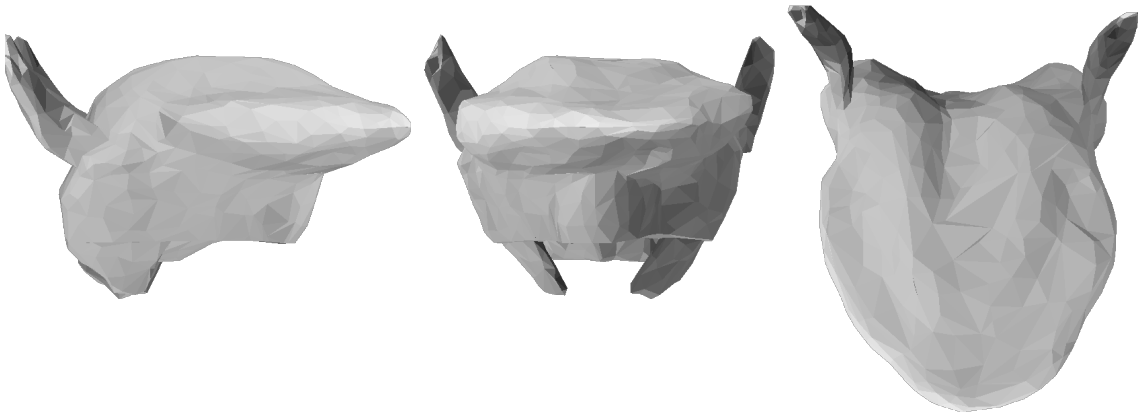

Figure 6.130: $D_4^{\mathbb{P}4}$: p_6 at maximum value.

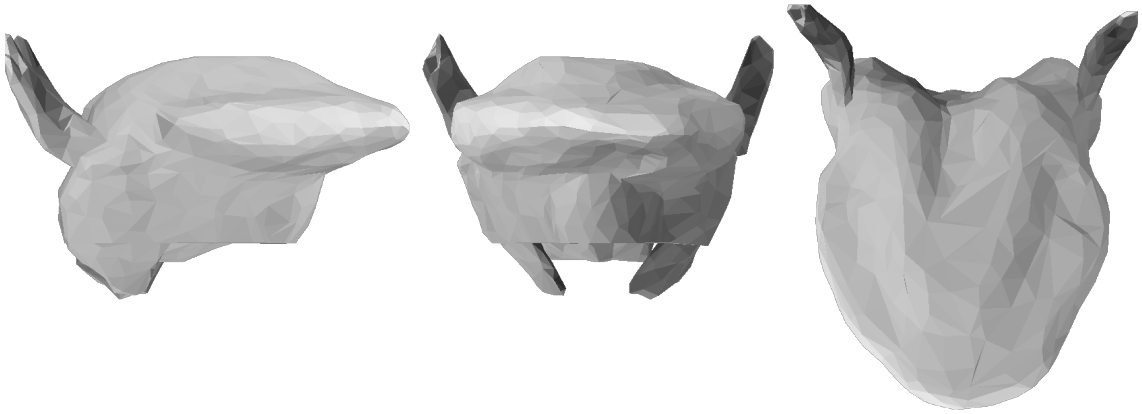

Figure 6.131: $D_4^{\mathbb{P}4}$: p_7 at minimum value.

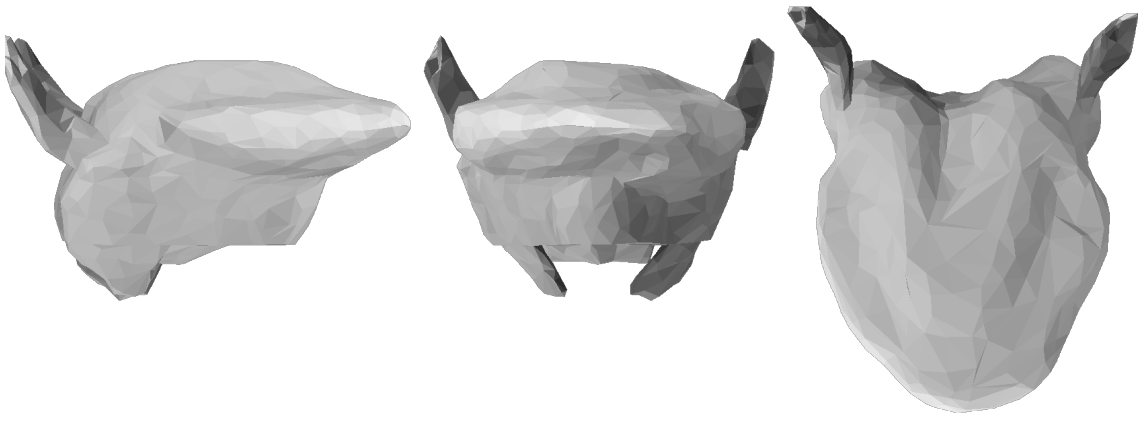

Figure 6.132: $D_4^{\mathbb{P}4}$: p_7 at maximum value.

The principal component analysis performed on D_5 yielded much the same result. The resulting basis, $\mathbb{P}_5$, had seven parameters, which accounted for greater than 99% of the variance. The variance accounted for by each of the first ten components in shown in Figure 6.133. Minimum and maximum values are shown in Figures 6.134 through 6.147.

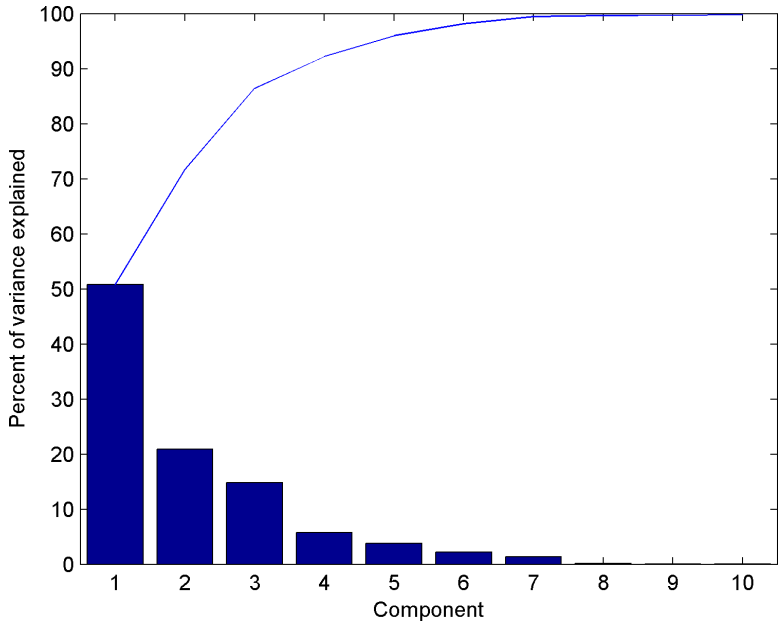

Figure 6.133: Pareto diagram of variance in D_5 accounted for by each parameter of $\mathbb{P}_5$.

As with previous analyses, D_5 was projected onto $\mathbb{P}_4$ to determine whether the smaller set of data produced a basis space sufficient for the larger set. It was found that $\mathbb{P}_4$ accounted for D_5 quite well. As shown in Figure 6.148, the first seven parameters derived from D_4 are sufficient to account for greater than 99% of the variance in D_5. Therefore a similar relation obtains for $\mathbb{P}_4$ and D_5 as exists between $\mathbb{P}_1$ and D_2. Subsequently, D_4 and D_5 will refer to $D_4^{\mathbb{P}_4}$ and $D_5^{\mathbb{P}_4}$, respectively.

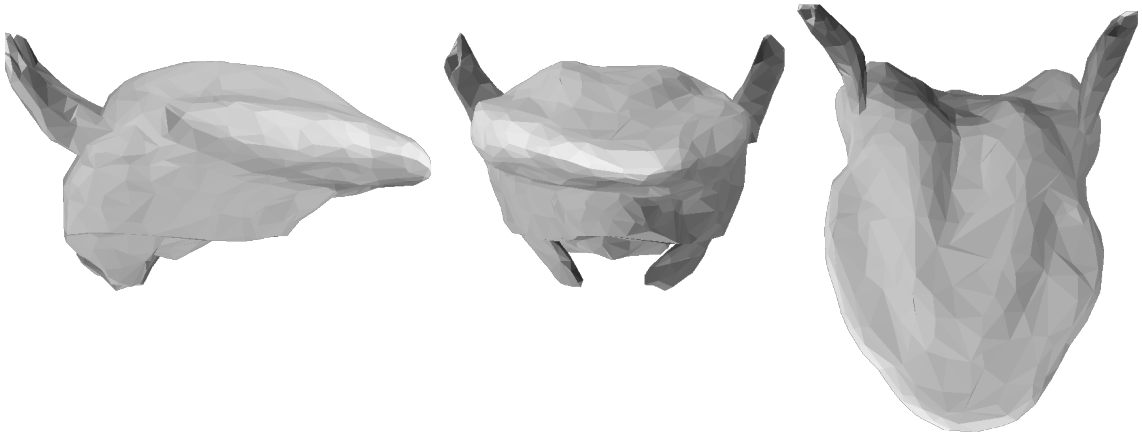

Figure 6.134: $D_5^{\mathbb{P}^5}$: p_1 at minimum value.

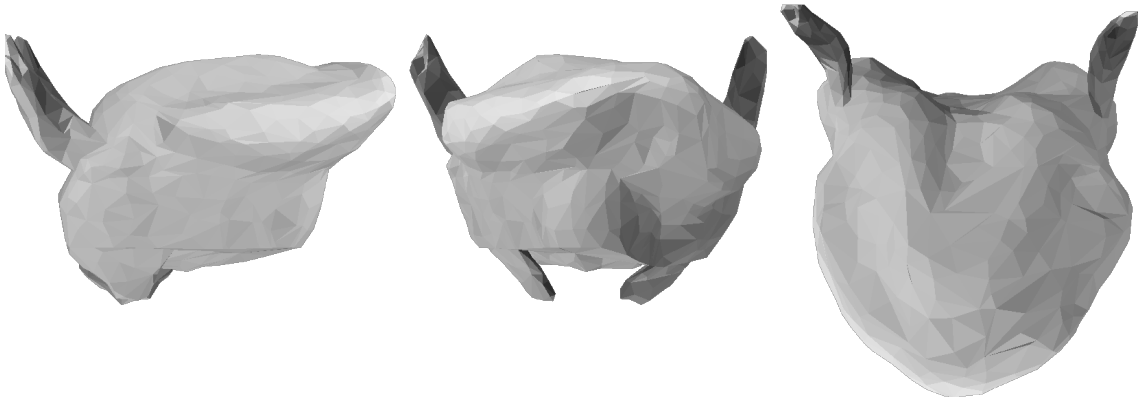

Figure 6.135: $D_5^{\mathbb{P}^5}$: p_1 at maximum value.

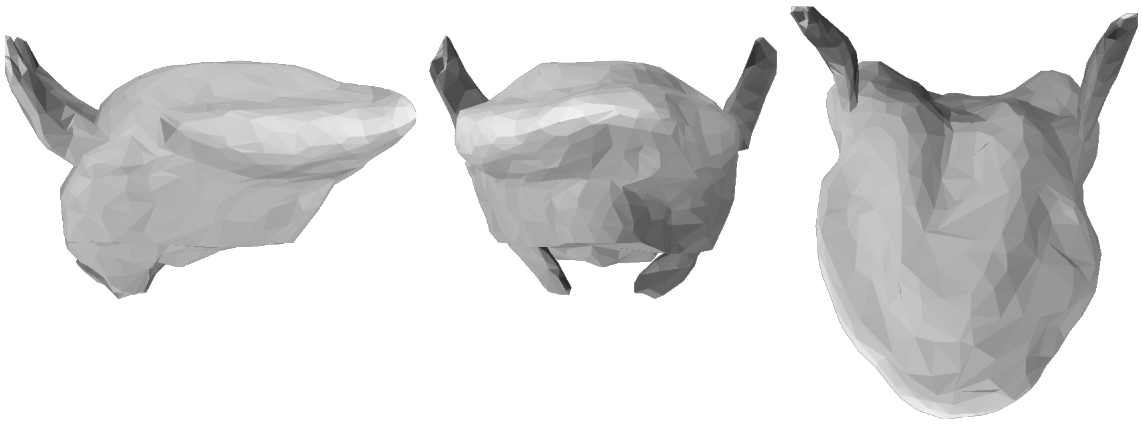

Figure 6.136: $D_5^{\mathbb{P}^5}$: p_2 at minimum value.

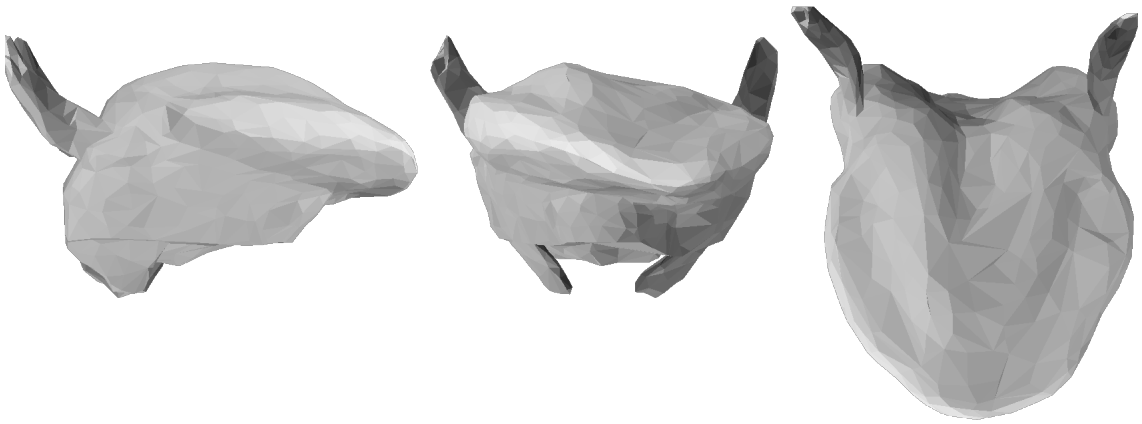

Figure 6.137: $D_5^{\mathbb{P}5}$: p_2 at maximum value.

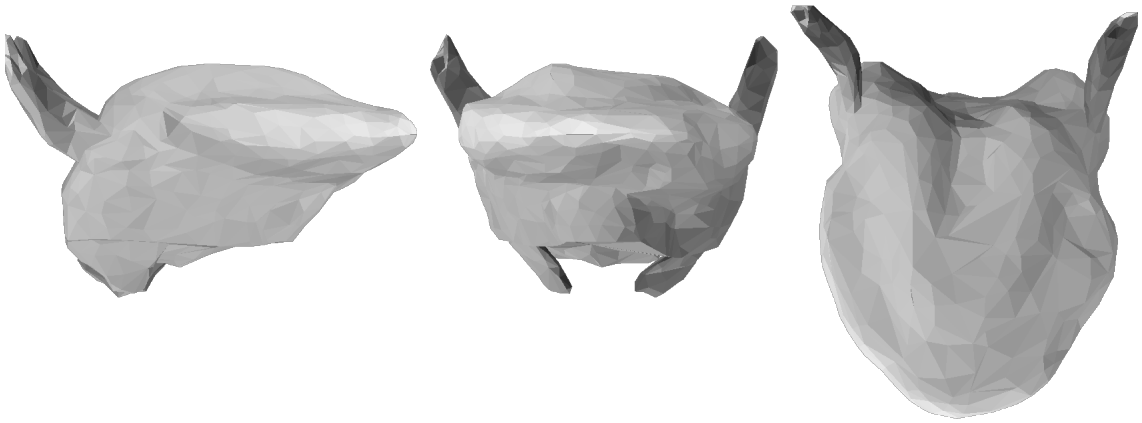

Figure 6.138: $D_5^{\mathbb{P}5}$: p_3 at minimum value.

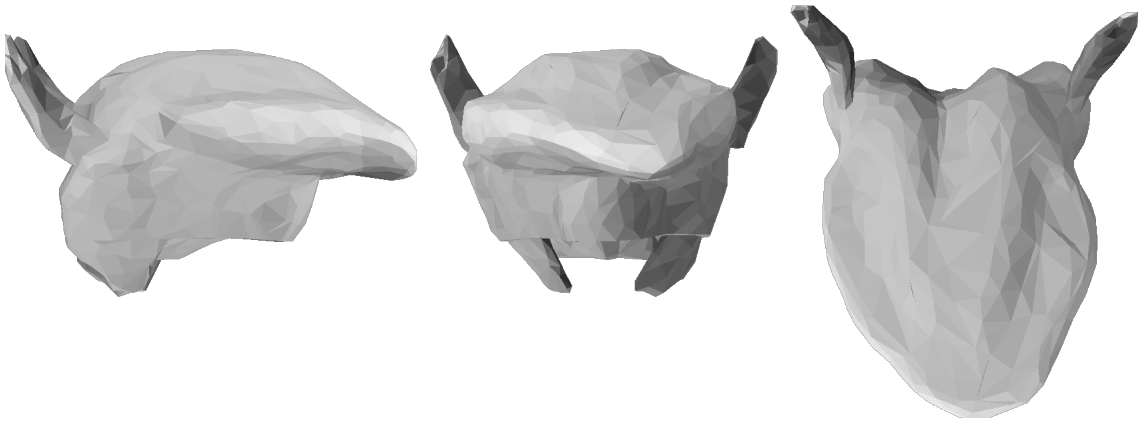

Figure 6.139: $D_5^{\mathbb{P}5}$: p_3 at maximum value.

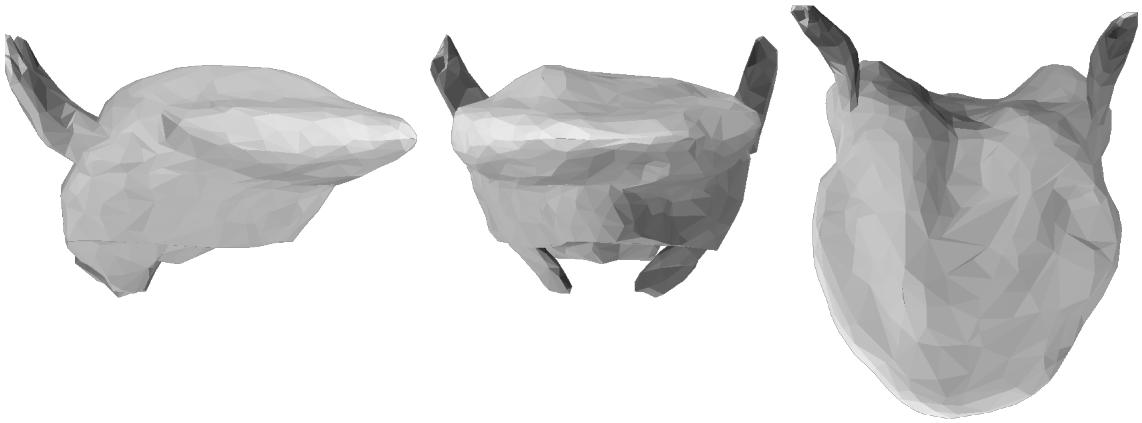

Figure 6.140: $D_5^{\mathbb{P}^5}$: p_4 at minimum value.

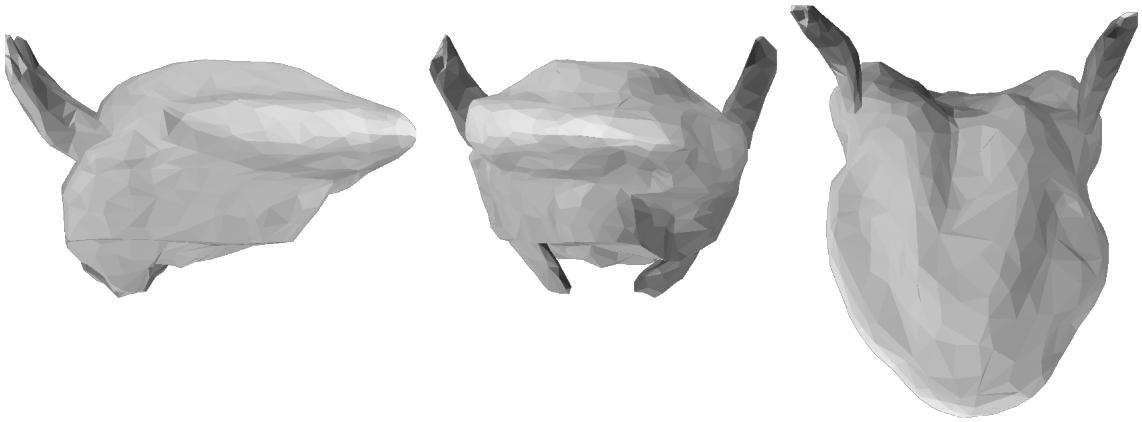

Figure 6.141: $D_5^{\mathbb{P}^5}$: p_4 at maximum value.

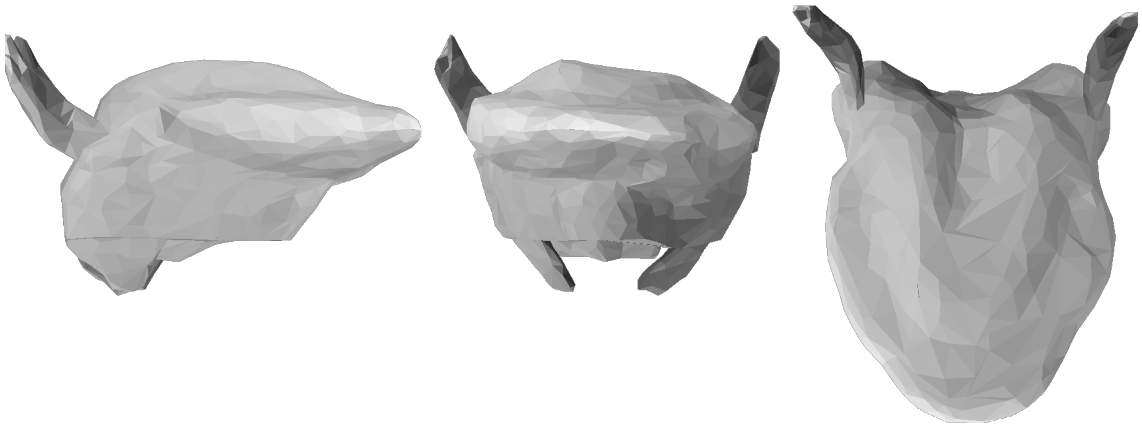

Figure 6.142: $D_5^{\mathbb{P}^5}$: p_5 at minimum value.

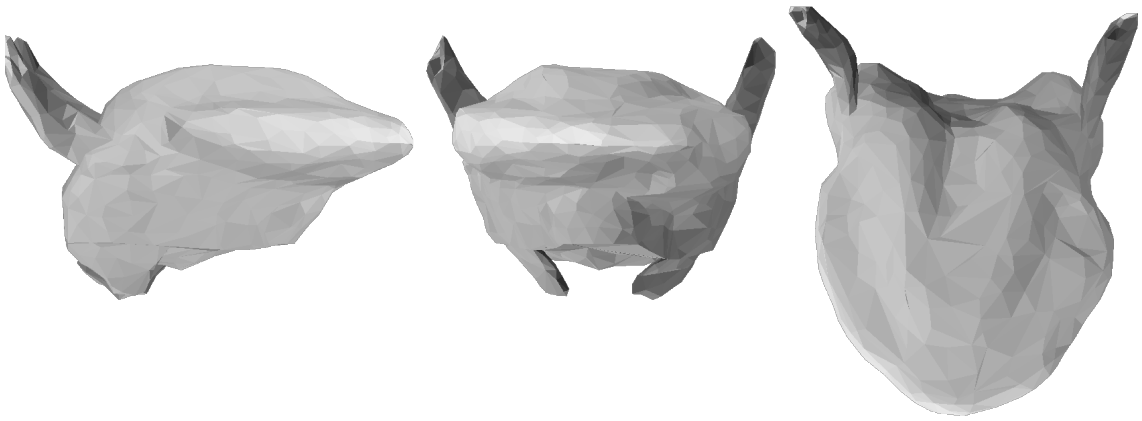

Figure 6.143: $D_5^{\mathbb{P}5}$: p_5 at maximum value.

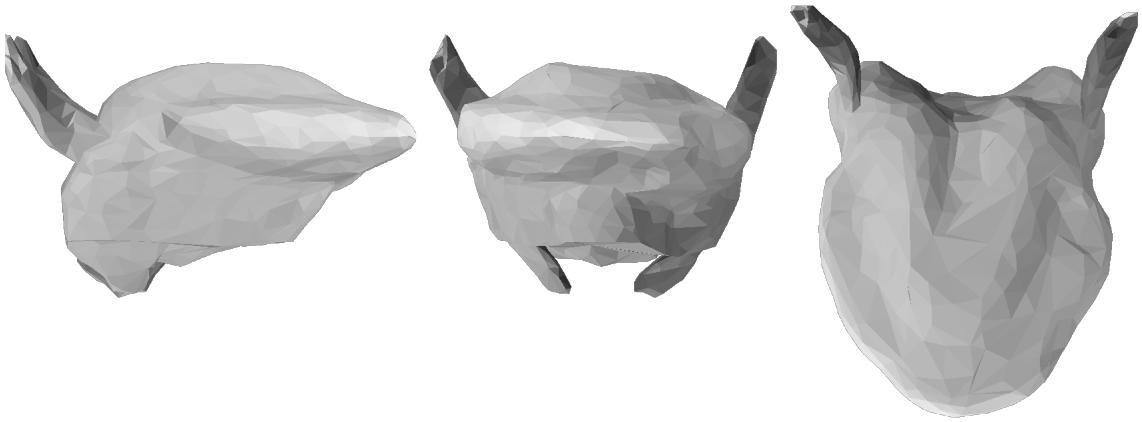

Figure 6.144: $D_5^{\mathbb{P}5}$: p_6 at minimum value.

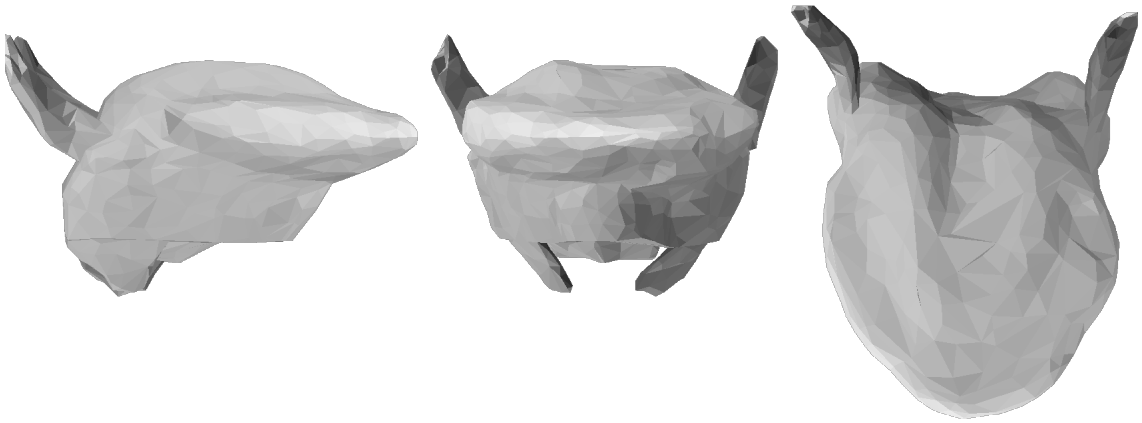

Figure 6.145: $D_5^{\mathbb{P}5}$: p_6 at maximum value.

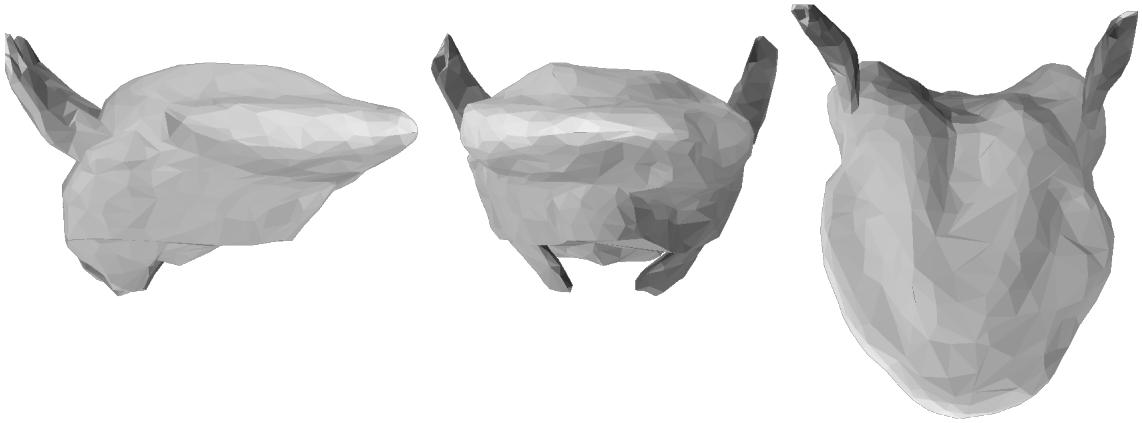

Figure 6.146: $D_5^{\mathbb{P}^5}$: p_7 at minimum value.

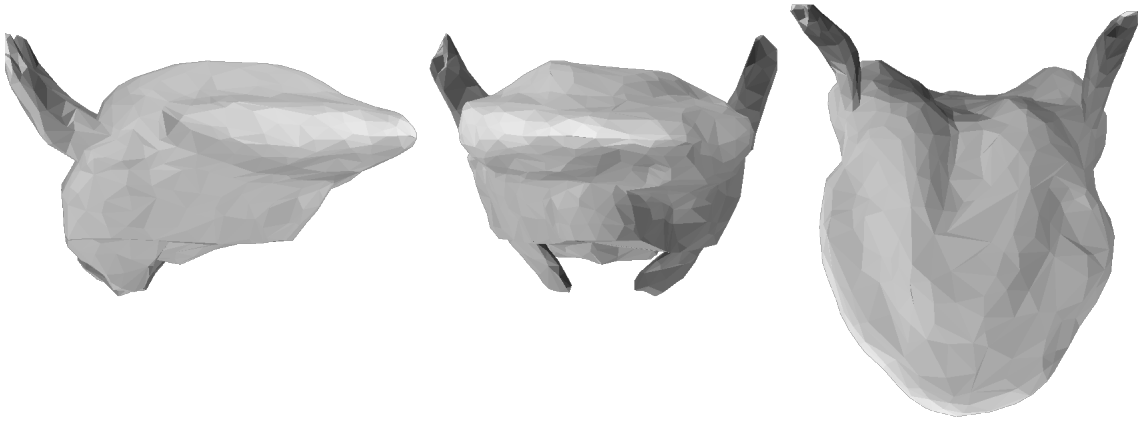

Figure 6.147: $D_5^{\mathbb{P}^5}$: p_7 at maximum value.

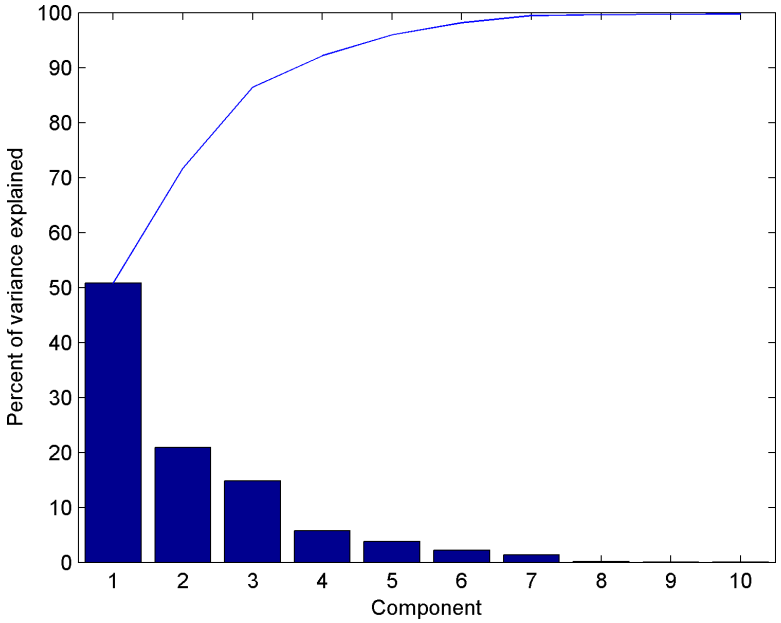

Figure 6.148: Pareto diagram of variance in D_5 accounted for by each parameter of $\mathbb{P}_4$.

6.3.2 Linear regression

Linear regression models were fit to D_4.[11] In mapping activations levels to tongue posture parameters, the model was quite successful in accounting for the data to which it was fit (D_4; adj. r^2 from 0.9359 to 0.9855), and in generalizing to the larger data set (($D_5 - D_4$); adj. r^2 from 0.8113 to 0.9554); see Table 6.4. In mapping from posture parameters to activation levels, the regression model was similarly successful (adj. r^2 from 0.9386 to 0.9885 in D_4, from 0.8973 to 0.9590 for ($D_5 - D_4$)); see Table 6.5.

[11]The procedures for the statistical models in this section are identical to those given in Section 5.3, with the exception that a testing data set was not developed for the hemiplegic tongue.

Component	D_4	$(D_5 - D_4)$
p_1	0.9778	0.9495
p_2	0.9729	0.9554
p_3	0.9855	0.9161
p_4	0.9789	0.9484
p_5	0.9359	0.9315
p_6	0.9524	0.8756
p_7	0.9650	0.8113

Table 6.4: Adjusted r^2 for different data sets, predicted from the linear regression of D_4 on A_1.

Component	D_4	$(D_5 - D_4)$
Genioglossus	0.9885	0.8973
Hyoglossus	0.9585	0.9071
Styloglossus	0.9592	0.826
Transversus	0.9821	0.9445
Verticalis	0.9830	0.9335
Superior Longitudinalis	0.9590	0.9156
Inferior Longitudinalis	0.9386	0.9445

Table 6.5: Adjusted r^2 for different data sets, predicted from the linear regression of A_1 on D_4.

6.3.3 Comparison of muscle strategies for a normal and hemiplegic tongue

This section contains a test of the hypothesis that the muscle activation pattern associated with protruding a healthy tongue is different from that for protruding a hemiplegic tongue. To test this, new linear regressions were calculated for the D_2 and D_5 data sets, so that the regression coefficients could be interpreted in order to assess the contribution of each muscle to each component of movement.

The range of parameter values for $D_2^{\mathbb{P}_1}$ and $D_5^{\mathbb{P}_4}$ was quite different. If regressions were performed on the raw parameter values then the regression coefficients would not be directly comparable. Therefore the values of $D_2^{\mathbb{P}_1}$ and $D_5^{\mathbb{P}_4}$ were scaled so that each parameter varied in the interval $[-1, 1]$. The signs of the parameters were also changed for the second and third parameters of $D_5^{\mathbb{P}_4}$ so that the physical interpretations of a positive value (tongue tip raising, or dorsum bunching, respectively) corresponded to those in $\mathbb{P}_1$. As noted in Chapter 5, the sign of the parameter values from a principal component analysis is arbitrary.

Linear regression models (including an intercept term) were fit to these data sets, with tongue parameters as dependent variables and muscle activation levels as the independent variables. In this way, the contribution of each muscle to each parameter could be assessed. Since the correspondences between the physical interpretations of the parameters of $\mathbb{P}_1$ and $\mathbb{P}_4$ were clearest for the first three components, only these are analyzed here.

Contributions of each muscle to p_1 are shown in Figure 6.149. Recall from Section 5.2.3 that a large positive value of p_1 corresponds to tongue protrusion, where a large negative value corresponds to raising of the tongue tip. Therefore, muscles that have a positive correlation coefficient for p_1 in Figure 6.149 (genioglossus, transversus, verticalis) contribute to protrusion, whereas muscle with a negative correlation coefficient (the remaining muscles, most notably superior longitudinalis) contribute to tongue tip raising. By comparing the coefficients for the normal and hemiplegic tongues, it is possible to compare the "best" strategies for tongue protrusion or tip raising, under the different stimulation conditions. For instance, relative to a

normal tongue, the best way to raise the tip with a hemiplegic tongue involves less genioglossus and more transversus and verticalis. For tip raising, the best hemiplegic strategy involves less activation of styloglossus and superior longitudinalis, and more activation of hyoglossus and inferior longitudinalis.

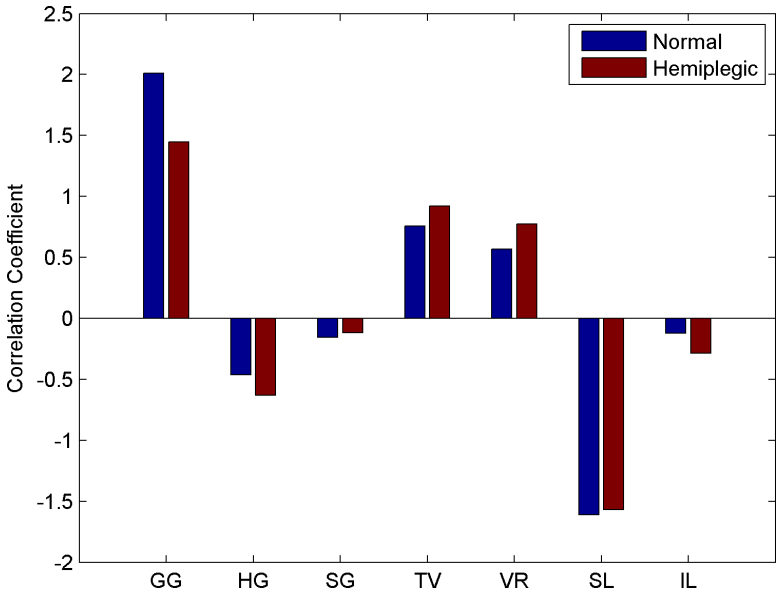

Figure 6.149: Comparison of the regression coefficients of each muscle's activation level with p_1.

Considering now p_2, which at large positive values is associated with raising and backing of the tongue body, and at large negative values with raising of the tongue tip, it is evident from Figure 6.150 that there are large differences between the normal and hemiplegic tongues. Genioglossus and transversus actually have opposite effects in the hemiplegic tongue. Correlation coefficients of the remaining muscles have very different magnitudes.

p_3 is associated with bunching at large positive values, and with flattening at large negative values. Figure 6.151 shows that in a hemiplegic tongue, transversus plays a much larger role in bunching the tongue in a hemiplegic tongue, and that genioglossus plays a smaller relative role in flattening the tongue. Styloglossus and inferior longitudinalis are associated with bunching in a normal tongue, but with flattening in a hemiplegic tongue. Superior longitudinalis is associated with flat-

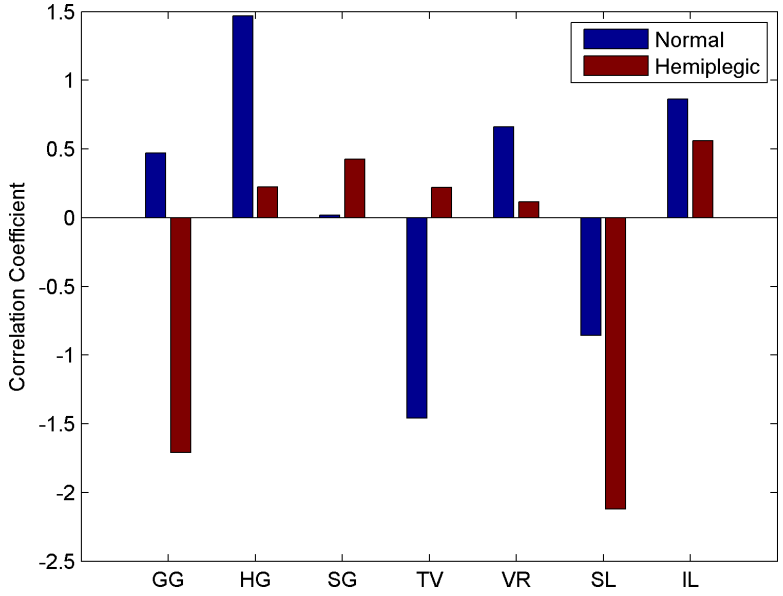

Figure 6.150: Comparison of the regression coefficients of each muscle's activation level with p_2.

tening in normal tongue, but in hemiplegic tongue it has a weak association with bunching.

6.3.4 Discussion

The results in this section indicate that the same techniques used to assess the behavior of a model of normal tongue function—principal component analysis and linear regression—are appropriate for studying a hemiplegic tongue as well. Tongue posture could be described with the same number of parameters at a similar level of precision. Also, the smaller data set generalized equally well to the larger, both in the principal components analysis and the linear regression.

It was also demonstrated that a hemiplegic tongue achieves various postures (protrusion, tip elevation, bunching, etc.) with different muscle activation levels than a normal tongue uses. This is not very surprising since, mechanically, moving the tongue that is bilaterally active is quite a different task from moving a tongue

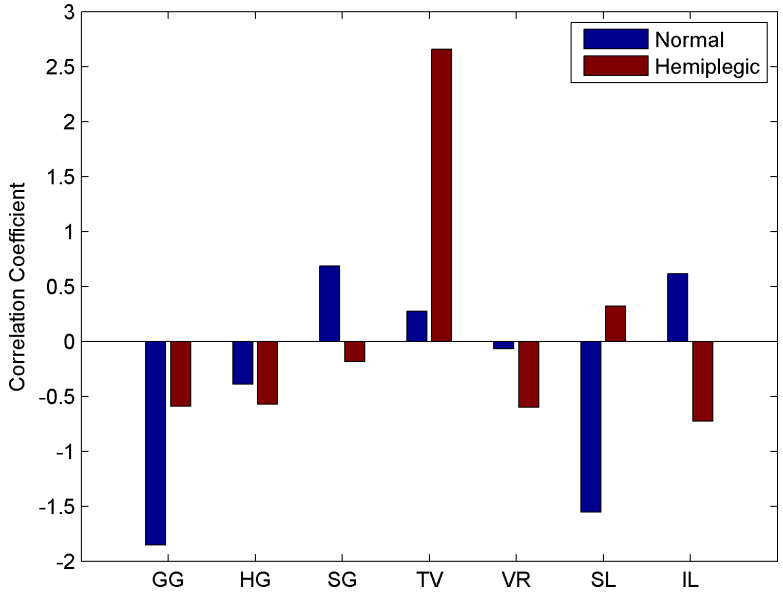

Figure 6.151: Comparison of the regression coefficients of each muscle's activation level with p_3.

that is unilaterally active. These results suggest either that speakers with a flaccid dysarthria adopt different muscular contraction strategies to deform their tongues, or that they should do so in order to regain a greater range of movement.

6.4 Summary and Discussion

This chapter has presented a selection of descriptive results from the tongue model, and some tests of predictions made in the literature. Section 6.1 examined the role of particular muscles and combinations of muscles in deforming the tongue. On the whole, the predictions made by the present model are consistent with what one would expect given the orientation of the fibers of the tongue. An exception to this was styloglossus, which in this model had very little effect on tongue posture. This result was surprising, given the role that styloglossus plays in speech production (e.g., Baer et al., 1988). Several possibilities exist for explaining this discrepancy. The effect of styloglossus could have been undermined by the inadequate specification of the

location of styloglossus, underestimation of the stress generated by styloglossus, or the position of the jaw and hyoid relative to the styloid process. The first possibility was not tested, but the latter two of were shown (by simulation) to be plausible, if not actual, explanations. Section 6.1.3 examined the role of coactivation of muscles in producing the full range of tongue movement; it was shown that muscle coactivation is crucial for achieving the full range of tongue postures.

CHAPTER 7

Conclusion

> Love never fails. But where there are prophecies, they will cease; where there are tongues, they will be stilled; where there is knowledge, it will pass away.
>
> — 1 Corinthians 13:8 (NIV)

7.1 Main results

The main result of this dissertation is a biomechanical model of the tongue. The model is based on an actual anatomical specimen, the female donor from the Visible Human Project. Digital images of the tongue were manually segmented to identify the gross shape of the tongue, and the locations of its constituent muscles. The geometry of the finite element model was determined by the anatomical data. Fiber orientations for each muscle were also determined based on the Visible Human Project data. During simulations, activated muscles generated stress, which caused the model to deform. This model is publicly available to researchers, with the hope that it will prove useful in other tongue research, and will itself be improved upon.

The focus of Chapter 5 was to test the behavior of the model systematically. This centered on the two questions of how to efficiently represent tongue posture, and how to predict tongue posture from muscle activation parameters (and vice versa). An efficient representation of tongue posture was obtained by performing a principal component analysis on the coordinates of a subset of the surface nodes of the model. The PCA produces a new coordinate space in which tongue posture could be represented; it was found that greater than 99% of the variance in the data could be captured with just seven parameters (basis vectors). Statistical models were fit using posture parameters derived from the PCA and the muscle activation levels,

alternately as dependent and independent variables. Most every statistical model performed reasonably well, and many accounted for nearly all the variance in the testing data sets.

Both of these techniques require a sample of postures from the model, but the computational expense associated with sampling becomes prohibitive the more control parameters are included, and the more levels at which they are sampled. A question that is pertinent to learning how to control the tongue—whether by a researcher or by the organism, since the problem is the same for both—is the extent to which a small sample of data can be used to generalize over larger samples. To test this, principal components analysis and statistical models were determined using 127-sample data set (D_1), which were then tested on their ability to generalize to a larger 2186-sample data set (D_2, $D_1 \subset D_2$). The basis determined from the PCA on D_1 generalized very well to D_2. The predictions of linear regression models fit to D_1 also generalized quite well to D_2 (in fact, to $D_2 - D_1$).

Since sampling all possible antisymmetrical combinations of muscle activations would be computationally prohibitive as of this writing, a novel technique was used to examine the range of antisymmetrical tongue postures. Tongue postures were identified that represented extreme positions in posture parameter space ($\mathbb{P}_1$). Then postures were calculated for all combinations of these activations in the left and right sides of the tongue. With qualifications, this can be taken to be an indicator of the range of antisymmetrical tongue movement.

These results are the first evidence that the range of possible tongue postures can be described with a small number of parameters. Previously researchers have demonstrated with principal components analysis that the shape of the vocal tract can be described with a small number of parameters (Harshman et al., 1977; Story et al., 1996b), but the sample set in previous studies was the set of English vowels. In the present study the sample set was the set of tongue postures corresponding to all possible symmetrical muscle contractions (sampled at two or three levels of activation).[1] The principal component analysis performed on antisymmetrical tongue

[1]Perrier et al. (2000) performed a principal component analysis on a large sample of tongue

postures is the first of its kind.

The results also demonstrate that efficient statistical models can be fit to predict tongue posture from muscle activation, or muscle activation from tongue posture. This finding is again significant for a researcher creating a model or for understanding how humans control their tongues. For the researcher, it means that the model can be easily added to other models of the vocal tract, for instance one that included other anatomical structures, or that included a motor learning algorithm. Since simulations of a single tongue posture can require hours of processor time to complete, this would not otherwise be practical. From a developmental perspective, it indicates that predicting postures from muscle activation, or vice versa, is far from intractable.

Finally in Chapter 6, a number of specific tests were performed. The consequences of single muscle contraction were reported—standard fare for reports on biomechanical tongue models—which were generally unsurprising, except for the fact that styloglossus seems to have minimal influence on tongue posture. Several reasons for this were suggested, including that styloglossus in humans is a vestigial muscle. A PCA of tongue postures that involve contraction of only one muscle revealed a quite limited range of motion; this indicates that coactivation of muscles plays a major role in producing the full range of tongue postures.

A number of specific predictions from the literature were tested. The predictions of the muscular hydrostat theory of tongue movement (Kier and Smith, 1985) were largely confirmed. Predictions from other researchers were not born out by the data. A systematic investigation of a possible functional division between the intrinsic and extrinsic muscles of the tongue failed to yield any evidence to support such a division.

Finally, a model of flaccid dysarthria was created by eliminating the motor input to one half of the tongue. The postures of the hemiplegic tongue were studied with the same analytical tools as the normal tongue. By comparing regression coefficients for the normal and hemiplegic tongue, it was determined that control of a hemiplegic

---

postures, but the range of motion of their tongue model was so small that their results are incommensurate with the present study.

tongue generally requires different muscle activation patterns from those of a normal tongue.

7.2 Directions for future research

Here a few directions for future research are considered. The present work has advanced the state of the art in tongue modeling in several respects, but improvements can be made, both in the design of the finite element model and in analysis of the model.

7.2.1 Model design

The major limitation of the present model is the control model, which is limited to whole muscle contraction. Since different compartments of the tongue can be controlled independently, it is of considerable importance that this level of detail be incorporated into a tongue model, as the data become available. With this done, the range of motion of the model will become reflective to a much greater degree of the range of motion of the actual tongue.

There are also numerous other improvements that could be made to the model. After the simulations for this dissertation had been completed, the more realistic muscle material of Blemker (2004) was made available for the software package that was used, FEBio. This could be incorporated into the model with very little difficulty. It would be preferable to have material parameters—both passive and active—from tongue tissue specifically. This would require mechanical tests to be performed on biological tissues. Finally, as computational restrictions lessen, it would be nice to have a finite element mesh fine enough to represent the subtle interdigitation of tongue fibers, e.g., representing all of the lamina of transversus and verticalis.

One of the goals of biomechanical speech research is a full model of the vocal tract. A tongue model represents the realization of a substantial part of that goal. Geometrically, it provides a little less than half of the vocal tract; in terms of modeling complexity, it provides perhaps a bit more. Of the remaining vocal tract

structures, the rigid structures can be added trivially, while the lips, cheeks, soft palate, and pharyngeal constrictors can be added with a fairly modest increase in modeling complexity (see Lucero and Munhall, 1999, for an extant model of the lips and cheeks). The jaw and hyoid bones, particularly the hyoid, are the remaining challenges.

The present model of the tongue is clearly limited by the fixed position of the hyoid. Motion of the hyoid is clearly important in determining the horizontal and vertical position of the tongue. It was remarked previously (Section 6.2.1) that the tongue does not, by its own deformation, lower very much. When the tongue seems to lower, it is probably actually the lowering of the tongue's main inferior attachment, the hyoid bone. Allowing for a user-specified translation and rotation of the hyoid bone is not difficult; this was already done in a limited way in Section 6.2.3. Systematic examination of the effect of hyoid motion is more challenging, since $N_H \cdot L^P$ samples are required, where N_H is the number of translations and rotations of the hyoid bone. N_H would have to be fairly large for the sample to be convincing.

Producing a model that plausibly takes into account the mechanical factors that influence the position of the hyoid bone will be more challenging still. The hyoid has superior attachments to the mandible (itself moveable) and the skull. The hyoid forms the superior portion of the larynx, itself connected to the lungs by the trachea. The hyoid also connects to inferior skeletal structures, such as the scapula (via the omohyoid muscle).

Another step in validating the model will be to conform the anatomy of the model to the anatomy of an actual speaker with a morphometric mapping. This has already been done with a finite element model by Gérard et al. (2006). The technique will allow direct comparisons between the output of the model and the behavior of an actual speaker. This would not only serve to validate (or invalidate) the output of the model, but also might make the model applicable to clinical situations.

7.2.2 Model analysis

Readers familiar with the statistical techniques employed in Chapter 5 will recognize that the author has merely plucked the low-hanging fruit as far as systematic generalizations about the model are concerned. Abundant work remains for understanding the output of the model (and by extension, the behavior of the tongue).

Much of the difficulty in analyzing the output of the model lies in sampling the behavior of the model. As the control model of the tongue becomes increasingly complex, this problem with become acute. In Section 5.5 a workaround was used in order to create a parameter space for antisymmetrical tongue postures. Some similar technique—or a massive commitment of computing time—will be required to sample the activation spaces of more complex models.

The present analysis has focused entirely on the coordinates of the surface nodes, though the model actually produces many other variables of potential interest. For surgical planning, it may be interesting to consider the position of interior nodes. In studying deglutition it would be necessary to assess the stiffness of the tongue surface, or the force that the tongue is able to exert on the hard palate.

7.3 Concluding Remarks

The tongue is a unique and fascinating organ, which is central to many social and survival-related human behaviors. Previous understanding of the control of the human tongue has been fairly schematic, having been based on researchers' impressionistic judgments, and on biomechanical models that almost certainly suffered artifacts as a result of having a coarse mesh.

In the present work, a finite element model was created based on an actual anatomical specimen, and which had a mesh that was sufficiently fine to guarantee that any errors attributable to the mesh were very small. A number of systematic explorations of the model have been performed. It has been shown that the posture of the tongue can be described with a relatively small number of independent parameters. It has also been shown that these parameters can be computed efficiently

from muscle activation levels, and vice versa, using simple statistical models. Moreover, it has been shown that these techniques can be applied to fairly small data sets, and still make good generalizations to much larger data sets. These results have numerous applications in human tongue research.

APPENDIX A

Guide to the Model Implementation

The purpose of this chapter is to assist a researcher in reproducing the model(s) reported in the dissertation. It does not describe the format of the raw anatomical data or the program used to view the results, Tongue Viewer. The reader is referred respectively to Appendix B and to Appendix C for a discussion of these topics.

A.1 Software implementation

The finite element model was implemented with FEBio, a finite element analysis program that is currently available under a no-cost license (Maas and Weiss, 2008). FEBio has good documentation, so the discussion below focus on the particular approach used, rather than recapitulating the documentation of the software. Except where specified, FEBio's default settings are assumed.

The tetrahedral mesh was created with a separate program and imported into FEBio via another piece of software, AssignProperties (see Section A.2 below). Each control parameter corresponded to a single material in FEBio. In a typical simulation, there was one material for the mucosa, and one for each of the seven muscles. Each muscle was associated with a loadcurve, which was used to specify the magnitude of that muscle's contraction for the particular simulation. More materials were used when different subportions of the tongue were controlled independently (in the antisymmetrical contraction case, for example).

The conjugate gradient solver was used, with a maximum of 10^6 iterations allowed.

The material parameters are given in Table 4.2.

Though this work was concerned only with static tongue postures, for computational reasons all simulations had a time component. A typical simulation was run

with one time step of 0.1 s, or with ten time steps of 0.01 s. Activation of a muscle was set to ramp up from no activation at $t = 0$, to full activation at $t = 0.1$. In this way, if a solution could not be found for full activation, the solver's auto-stepper could find a solution at a time with a less extreme deformation. These settings produced acceptable results in most cases, requiring between 2–6 hours, depending on the run. For some simulations, the auto-stepper was not efficient in finding solutions. In these cases the auto-stepper was disabled, and 20 time steps of 0.005 s each were used. This produced successful results, typically with a 4 hour simulation time.

A.2 AssignProperties

AssignProperties is a command-line program for assigning muscle properties to mesh geometry and converting the data into FEBio format.

AssignProperties is written in C. Assuming that the GNU GCC compiler is available, the following command will create an executable named "ap.exe" or "ap" (depending on the operating system).

(A.1) `gcc AssignProperties.c -lm -o ap`

A sample run of AssignProperties is shown in Figure A.1. Text after the > symbol is entered by the user. AssignProperties first prompts the user for a filename. The user types "Coarse," which indicates that AssignProperties should look for `Coarse_node.ascii` and `Coarse_element.ascii`, which contain nodal coordinates and element coordinates, respectively.[1,2] AssignProperties requests the number of nodes in the mesh (offering some reminders); the user enters "1337." Next AssignProperties requests a muscle scheme. Schemes 2, 3, and 4 are the only ones used in this dissertation. Finally AssignProperties asks whether to include the hyoid bone or not (see Section 6.2.3).

[1]These file formats are transparent. Two notes about the element file are that (1) the node indices are 1-indexed, and (2) the element indices are expected to be in appropriate order (so as not to produce a negative Jacobian), which should be true of the output of any meshing program.

[2]Coarse is a coarse mesh used only for testing software.

```
Enter the base of the filename:
> Coarse
How many nodes are there?
        Coarse: 1337
        ValidationMesh0: 14830
        ValidationMesh01: 2628
        ValidationMesh02: 1673
        ValidationMesh03: 1243
        ValidationMesh04: 988
        ValidationMesh1: 15788
        ValidationMesh11: 24237
        ValidationMesh2: 29072
        ValidationMesh21: 46490
        ValidationMesh3: 106449
        ValidationMesh31: 181758
        ValidationMesh32: 211814
        ValidationMesh33: 347298
        ValidationMesh4: 507955
> 1337
What muscle activation scheme would you like?
        1. The idiotic testing model
        2. Whole muscle symmetrical contraction
        3. Whole muscle antisymmetrical contraction
        4. Halle test
        5. Suckle test
        6. Suckle test II
        7. Suckle test III
> 2
Should the hyoid be added? (y/n)
> n
```

Figure A.1: Sample run of AssignProperties

AssignProperties's output consists of four text files. `Coarse_aux.txt` can be used for debugging purposes, but is generally empty. The remaining files contain data in FEBio's XML format, for inclusion in FEBio input files. `Coarse_geometry.txt` contains the model geometry (nodes and elements). The fiber orientations are given in `Coarse_elementdata.txt`. `Coarse_boundaries.txt` contains the boundary conditions.

APPENDIX B

Guide to the Anatomical Data

The results of the anatomical segmentation of the tongue and the determination of the direction of the tongue muscle fibers, as described in Chapter 3, are freely available to researchers. This dissertation should be cited as the source of the data.

The anatomical images themselves are available to researchers through a no-cost license from the Visible Human Project at the National Libraries of Medicine:

(B.1) `http://www.nlm.nih.gov/research/visible/visible_human.html`

B.1 Image Segmentation

B.1.1 Region of interest

The image segmentation was performed on a certain region of interest (ROI) of the female dataset. It is computationally more efficient to work with a subset of the dataset, rather than the entire data set. The 475^{th} to 751^{st} axial slices were used (inclusive). In the file scheme of the Visible Human set, these are images 1159b through 1251b (inclusive). Only a sub-portion of each axial image was used, the rectangle defined by the coordinates $(874, 553)$ and $(1307, 891)$ (with zero-indexed coordinates). The dimensions of the ROI, then, were $434 \times 399 \times 277$ (width $\times$ depth $\times$ height, or left-right $\times$ antero-posterior $\times$ superior-inferior).

To convert the images from the raw binary format the VHP provides, the images were processed with the Visible Human Preprocessing Toolkit (v1.0) for ImageJ. The blue gel removal option was used.

(B.2) `http://ciar.rcm.upr.edu/projects/imageJ`

B.1.2 Uncompressed File Format

The image segmentation resulted in a description of eight regions of the tongue:

- Genioglossus
- Hyoglossus
- Styloglossus
- Transversus/Verticalis
- Superior Longitudinalis
- Inferior Longitudinalis
- Bone[1]
- Mucosa

The description of the location of each label is a set of binary images ("segmentation images") that correspond to each anatomical image. White pixels indicate that the pixel belongs to the label, while black pixels indicate that the pixel does not. This is illustrated in Figure B.1 for an axial image that contains genioglossus. The original image (a) is matched by a black-and-white image that indicates where genioglossus is (b). From this, the region corresponding to genioglossus on the original image can be easily identified (c).

B.1.3 Compressed File Format

Since each segmentation image is black and white, each pixel can be described with a single bit. Since there are eight regions of the tongue, each pixel can be completely label with eight bits, or one byte. The segmentation data are therefore coded with one byte per pixel, which each label encoded in a different place in the byte.

[1]The description of "Bone" includes the inferior portion of the mandible, and the entire hyoid. These regions can be easily differentiated from one another, so they have been represented with a single label.

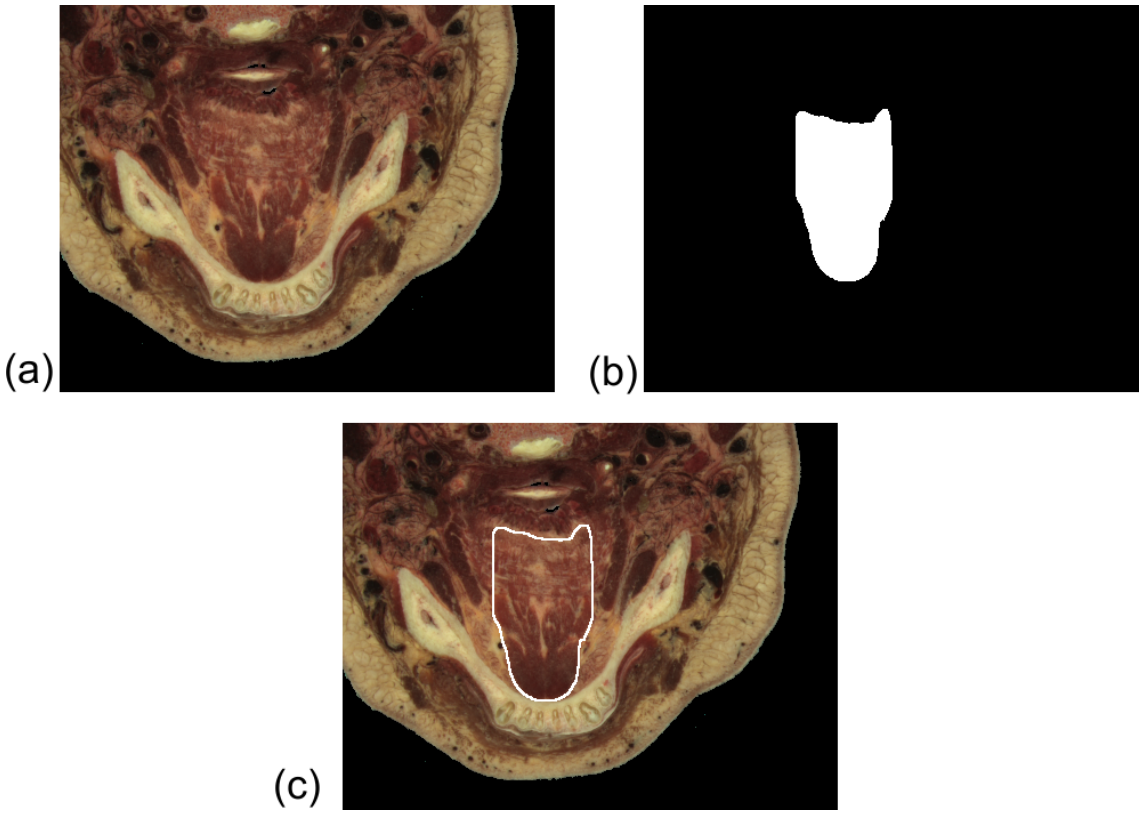

Figure B.1: Representation of the image segmentation. (a) Original anatomical image. (b) Bitmap representing which regions of the image contain fibers from genioglossus. (c) Outline of genioglossus.

Row 0	0	1	2	$\cdots$	431	432	433
Row 1	434	435	436	$\cdots$	865	866	867
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
Row 337	145921	145922	145923	$\cdots$	146352	146353	146354
Row 338	146692	146693	146694	$\cdots$	147123	147124	147125

Table B.1: Bytes (zero-indexed) corresponding to image positions, for the first slice.

The region of interest is represented $434 \times 339 \times 277 = 40,753,902$ bytes. The data are stored as sequences of axial images (superior to inferior order). Each image is stored in row-major order, which means that the data are serialized one row at a time ('row' here corresponds to the right-to-left axis). The byte locations for each pixel of the first image are illustrated in Table B.1. The byte locations for the second image could be obtained by adding 147126 to each value. The general expression for the byte location of the pixel in the i^{th} column of the j^{th} row of the k^{th} axial image (where i, j, and k are zero-indexed) is:

$$(k \cdot 147126) + (j \cdot 434) + i \tag{B.3}$$

Each byte contains the labeling information for each of the eight regions. Each region is associated with one bit of the byte. The key is given in Table B.2. Bit 1 is the least significant bit (the 1's place); bit 8 is the most significant bit (the 128's place). For instance, if the byte for a given pixel has a 1 in the 5th bit, then that voxel was assigned to Superior Longitudinalis. Note that regions can be overlapping (which reflects the interdigitation of muscle fibers in the tongue).

For a given pixel with a byte value b, C expressions for determining the value of particular labels are given in Table B.3. & is a the bitwise AND operator, while $x \gg y$ means "x bitwise-right-shifted by y bits."

Bit	Region
1	Genioglossus
2	Hyoglossus
3	Styloglossus
4	Transversus/Verticalis
5	Superior Longitudinalis
6	Inferior Longitudinalis
7	Bone
8	Mucosa

Table B.2: Bit coding of consolidated data.

Genioglossus	`b&1`
Hyoglossus	`(b&2)≫1`
Styloglossus	`(b&2)≫ 2`
Transversus/Verticalis	`(b&2)≫3`
Superior Longitudinalis	`(b&2)≫4`
Inferior Longitudinalis	`(b&2)≫5`
Bone	`(b&2)≫6`
Mucosa	`(b&2)≫7`

Table B.3: Expressions for extracting a particular label (1 or 0) from a byte b.

Muscle	File with $\vec{p}_i$	File with $\vec{v}_i$
Genioglossus	gg_loc.txt	gg_dir.txt
Hyoglossus	hg_loc.txt	hg_dir.txt
Styloglossus	sg_loc.txt	sg_dir.txt
Transversus	tr_loc.txt	tr_dir.txt
Verticalis	ve_loc.txt	ve_dir.txt
Superior Longitudinalis	sl_loc.txt	sl_dir.txt
Inferior Longitudinalis	il_loc.txt	il_dir.txt

Table B.4: Filenames associated with vectors that represent muscle fiber orientations

B.2 Fiber Orientations

As described in Section 3.2.3, the muscle fiber orientations of each muscle determined by interpolating between a small number of vectors. These vectors are provided in plain text files. The locations of the vectors ($\vec{p}_i$) for a given muscle are `XX_loc.txt` (where `XX` is a mnemonic for the muscle; see Table B.4). The directions vectors ($\vec{v}_i$) are in `XX_dir.txt`. For genioglossus, for instance, the location vector $\vec{p}_i$ is found in the i^{th} row of `gg_loc.txt`. The corresponding direction vector is found in the i^{th} row of `gg_dir.txt`. The direction vectors are relative from the origin. To draw a line originating from $\vec{p}_i$ and indicating the fiber orientation at that point, for instance, one would draw a line from $\vec{p}_i$ to $(\vec{p}_i + \vec{v}_i)$.

The location and direction vectors all have units of pixels. They must be scaled by 0.33 mm/pixel to convert them to millimeters. The coordinates are otherwise in the same space as the finite-element mesh.

APPENDIX C

Auxiliary Software

Most of the results of this dissertation are viewable with the Tongue Viewer application, which is included as supplementary material. Tongue Viewer has the capability to visualize the exterior nodes of an unreduced mesh, a tongue posture reconstructed from posture parameters, and animations produced by varying posture parameters or muscle activation levels. Descriptions of file formats are included in this appendix, but files are also provided in Section C.5 that should permit the program to be used without knowledge of these formats.

C.1 User interaction

When the program starts, the user is prompted to select a settings file, which contains references to filenames and other important program information. Only one settings file can be used in each run of the application; to switch files, launch a new instance of Tongue Viewer. The format of the setting file is described below in Section C.2.

The top of the "Settings & Information" bar indicates the current settings file. Below this, a drop-down menu indicates the mode that Tongue Viewer is in. There are four modes: Dynamic, Static, Parameter Animation, and Control Animation. The mode is changed with the drop-down menu, or if the user clicks a button that necessitates a change in mode.

In any mode, the tongue model can be rotated by clicking and dragging with the mouse, or by changing the three sliders in the "Orientation" toolbar.

C.1.1 Static Mode

In Static mode, the program displays static tongue postures in binary format. A file is selected with the "Open Static" button. The format of the file is given below in Section C.3.

C.1.2 Dynamic Mode

In Dynamic mode, the user changes the posture of the tongue by using the posture parameter sliders (in the "Parameters" toolbar) or the control parameter sliders (in the "Controls" toolbar; control parameters typically whole muscles) . If control parameters are specified, then the corresponding posture parameters are calculated using a neural network (see Section 5.4). (The neural network to use must be specified by the user.)

If the "Keep control activations synchronized with parameters with a neural net" box in the "Settings & Information" toolbar is checked, then whenever the parameter values are changed the control settings are updated automatically with a neural network. (The neural network to use must be specified by the user.)

If the "Check that coordinate is within the hull" box in the "Settings & Information" toolbar is checked, then whenever the parameter values are changed the software will display a message in the "Settings & Information" toolbar indicating whether the current parameter settings are within the convex hull. (The convex hull must be specified by the user.)

The "Reset Parameters" button resets the tongue to the user-specified initial parameter settings. The "Take Pictures" button produces a Portable Network Graphics (PNG) image of the current model from the right, front, and top perspectives; the user is prompted for a filename.[1] The "Dump Binary" button writes the nodal coordinates of the current model in binary format.

[1] Incoveniently, the software produces garbled results if the toolbars are covering the rectangular bounding box of the tongue. These should be moved away from the display window—or if in full screen mode, to the right side of the screen—to avoid this problem.

C.1.3 Parameter Animation Mode and Control Animation Mode

These modes produce animations of time-varying posture parameter or control parameter values. When an animation file is loaded, playback of the animation can be controlled with the usual controls on the "Animation" toolbar. The animation will replay continuously until playback is paused. The slider indicates the frame position, but moving it has no effect.

The file format for animation files, which is identical for both parameter scores and control scores, is given in Section C.4.

C.2 Settings file

The settings for Tongue Viewer are stored in an XML file, an example of which is given below. Each tag is explained, along with the file format. If filenames are given, as below, without a full path, Tongue Viewer will look for the file in its directory of execution. It may be more convenient to specify a full path.

```
<?xml version="1.0" encoding="utf-8"?>
<settings>
<controls>
   <control>GG</control>
   <control>HG</control>
   <control>SG</control>
   <control>TV</control>
   <control>VR</control>
   <control>SL</control>
   <control>IL</control>
</controls>
<parameters>
   <parameter>P1</parameter>
   <parameter>P2</parameter>
   <parameter>P3</parameter>
```

```
   <parameter>P4</parameter>
   <parameter>P5</parameter>
   <parameter>P6</parameter>
   <parameter>P7</parameter>
</parameters>
<filenames>
   <static>
      <sNPoints>11331</sNPoints>
      <sNTriangles>23000</sNTriangles>
      <sConnectivity>VM21vert.bin</sConnectivity>
   </static>
   <dynamic>
      <dNPoints>960</dNPoints>
      <dNTriangles>23000</dNTriangles>
      <dConnectivity>VM21_2k_el.bin</dConnectivity>
      <dEigen>b19eigen.bin</dEigen>
      <dMeans>b19means.bin</dMeans>
      <dRange>b20b19range.bin</dRange>
      <dInitial>b19initial.bin</dInitial>
      <dNHull>929834</dNHull>
      <dHull>b20b19hull7.bin</dHull>
      <annControlsToParameters>ann_coord_act.bin</annControlsToParameters>
      <annParametersToControls>ann_act_coord.bin</annParametersToControls>
   </dynamic>
</filenames>
</settings>
```

The `<controls>` element contains `<control>` elements. The text in a `<control>` element is the label for the control (in this example, muscle labels). The number of `<control>` tags tells Tongue Viewer how many control parameters to expect; this number must be consistent with the other settings.

The <parameters> element likewise contains the <parameter> element. Each of these is a label for a parameter value. The number of <parameter> elements must be the number of parameters that is assumed in other data files.

The <static> tag has information required to display static tongue postures. <sNPoints> has the number of nodes in the mesh. <sNTriangles> has the number of triangles in the mesh. <sConnectivity> is a filename for a file that specifies the nodes of each triangle.[2] The nodes are zero-indexed indices from the static data file, so the largest acceptable value is (<sNPoints>−1). Each index is an unsigned 2-byte integer. The first three 2-byte blocks are the nodes of the first triangle, the second three are the nodes of the second triangle, and so forth. The file is 6×<sNTriangles> bytes in length. For rendering purposes, it is necessary to have appropriate triangle normals. For nodes $\vec{n}_1, \vec{n}_2, \vec{n}_3$, the normal is calculated as $(\vec{n}_3 - \vec{n}_1) \otimes (\vec{n}_2 - \vec{n}_1)$, where $\otimes$ is the cross product. The normal should point away from the tongue. If the normal is pointing in the wrong direction, the triangles will be rendered as black regions.

The <dynamic> tag has information required to display dynamic tongue postures. The element <dNPoints> has the number of nodes in the dynamic mesh. <dNTriangles> is the number of triangles in the dynamic mesh. <dConnectivity> is the filename of a file that specifies the triangles; it has the same format as the file specified in <sConnectivity>; the file is 6×<dNTriangles> bytes in length. <dEigen> is the matrix of eigenvectors from the principal components analysis. The matrix is stored in column-major order; each entry is an 8-byte floating point variable. This file is 72×<dNPoints>×<dNPoints> bytes in length. <dMeans> indicates a file with the vector of means from the principal component analysis, stored as 8-byte floating point variables; the file is 24×<dNPoints> bytes in length.[3] <dRange> has the ranges of each parameter in the principal component analysis. All of the mini-

[2]Note that in Static mode, the user will supply a file containing the coordinates of each node.

[3]Note that this file must have the mean for *every* parameter, not just the parameters that are being used.

mum values are stored, followed by the maximum values.[4] Each entry is an 8-byte floating point variable, so the file is 48*<dNPoints> bytes in length. <dInitial> indicates a filename with the initial parameter values of the model, stored as 8-byte floating point variables; the file is 24×<dNPoints> bytes in length.

<dNHull> has the number of faces on the convex hull of the data set. <dHull> indicates a file with a specification of the hull. The nature of the hull depends on the number of parameters being manipulated in the data set, so the number of <parameter> tags (for short, $N_{parameter}$) must be consistent with this file. The file is a series of <dNHull> blocks of $N_{parameter}$ numbers. The program computes the dot product of each of these blocks with the vector of current parameter settings in the model. If any of the dot products is greater than 1.0001,[5] the parameter settings are judged to be outside of the hull. Each number is an 8-byte floating point number. The file indicated by <dHull> is $8 \times N_{parameter} \times$<dNHull> bytes in length.

<annControlsToParameters> and <annParametersToControls> both indicate files with specifications of neural networks. The formats of these files is somewhat obscure, since the file structure depends on the number of layers and number of neurons. The structure is shown in Table C.1. The first block is the number of layers in the network. The second is the number of inputs to the network. Next, for each layer, is the number of neurons in the layer. Next are the weights from the input to the first layer. The length of this field will be the number of inputs times the number of neurons in the first layer. The order is: the weights for each input to the first neuron, the weights for each input to the second neuron, etc. Next the weights between each layer follow. Next, the bias values for each layer are provided. Finally, the maximum and minimum values of the input and output, respectively, are given.

By way of a more concrete example, the file format for a particular network—in fact, the one that maps posture parameters to muscle activation levels—is shown in

[4]Note that this file must have the range every parameter, *not* just the parameters that are being used.

[5]The ten-thousandth here compensates for numerical imprecisions in calculation.

Field	Size
Number of layers	unsigned 4-byte integer
Number of inputs	unsigned 4-byte integer
Number of neurons in the first layer	unsigned 4-byte integer
...	
Number of neurons in the final layer	unsigned 4-byte integer
Weights from the input to the first layer	8-byte floating point
Weights from the first layer to the next layer	8-byte floating point
...	
Weights from the next-to-last layer to the last layer	8-byte floating point
Bias values for the first layer	8-byte floating point
...	
Bias values for the last layer	8-byte floating point
Minimum values of the input dimensions	8-byte floating point
Maximum values of the input dimensions	8-byte floating point
Minimum values of the output dimensions	8-byte floating point
Maximum values of the output dimensions	8-byte floating point

Table C.1: File format for neural network specification

Table C.2.

C.3 Static file format

Static files are lists of coordinates of nodes. The number of nodes must be consistent with the value of `<sNPoints>` in the settings file. Each coordinate is a 4-byte floating point value. The file is then 12×`<sNPoints>` bytes in length. The nodes are listed in xyz order, thus: $x_1, y_1, z_1,\ x_2, y_2, z_2,\ x_3, y_3, z_3,$ ...

C.4 Animation file format

Animation files specify an animation to play. A "parameter score" is an animation in which the user specifies changing posture parameters. A "control score" is an animation in which the user specifies changing control values (i.e., muscle activations). The file format is identical in either case. The user must ensure that the score

Field	**Value (Size)**
Number of layers	2
Number of inputs	7
Number of neurons in the first layer	20
Number of neurons in the second layer	7
Weights from the input to the first layer	(7×20 8-byte floating point values)
Weights from the first layer to the second layer	(20×7 8-byte floating point values)
Bias values for the first layer	(20 8-byte floating point values)
Bias values for the second layer	(7 8-byte floating point values)
Minimum values of the input dimensions	(7 8-byte floating point values)
Maximum values of the input dimensions	(7 8-byte floating point values)
Minimum values of the output dimensions	(7 8-byte floating point values)
Maximum values of the output dimensions	(7 8-byte floating point values)

Table C.2: File format of a particular neural network

is opened with whichever of "Open Parameter Score" or "Open Control Score" is appropriate. The user specifies parameter or control values at specific time increments. If a frame is required at a time with no specification, values are calculated with linear interpolation.

The animation file is plain text. A sample is shown in Figure C.1, which is a parameter score that demonstrates the effect of varying p_1. The first three lines are general settings for the animation. The first line tells the program to expect 3 sets of parameter values (i.e., parameter values at 3 times). The second line tells the program to expect seven parameter values at each time step. This number must be consistent with the number of parameters (or controls, as the case may be) in the settings file; if it is not Tongue Viewer will give an error. The third line tells the program to render the animation at 30 frames per second. Following this are 3 lines with eight values apiece. The first number is the time; the remaining seven are the seven parameter values. The separation of the numbers can be either a tab or a space character. The time does not need to be specified in equal intervals, as it is in Figure C.1.

```
3
7
30
0 -169.7476 37.0751 -9.3266 -1.1950 21.9618 28.4139 -10.7615
1 226.5673 37.0751 -9.3266 -1.1950 21.9618 28.4139 -10.7615
2 -169.7476 37.0751 -9.3266 -1.1950 21.9618 28.4139 -10.7615
```

Figure C.1: Sample animation format

C.5 Included files

The following settings files are included, as well as the binary files that they reference. Please note that it may be necessary to specify a full path to each file for Tongue Viewer to work properly.

D1P1.xml For viewing $D_1^{\mathbb{P}_1}$

D2P2.xml For viewing $D_2^{\mathbb{P}_2}$

D2P1.xml For viewing $D_2^{\mathbb{P}_1}$

D4P4.xml For viewing $D_4^{\mathbb{P}_4}$

D5P5.xml For viewing $D_5^{\mathbb{P}_5}$

Extreme.xml For viewing the antisymmetrical tongue postures from Section 5.5

ExtrinsicNative.xml For viewing the PCA of tongue postures involving contraction of only extrinsic muscles (Section 6.2.4)

ExtrinsicP1.xml Tongue postures involving contraction of only extrinsic muscles, projected onto $\mathbb{P}_1$ (Section 6.2.4)

IntrinsicNative.xml For viewing the PCA of tongue postures involving contraction of only intrinsic muscles (Section 6.2.4)

IntrinsicP1.xml Tongue postures involving contraction of only intrinsic muscles, projected onto $\mathbb{P}_1$ (Section 6.2.4)

The following animation files are included:

P1_illustration.txt For illustrating the range of p_1 in the $D_2^{\mathbb{P}_1}$ data.

P2_illustration.txt For illustrating the range of p_2 in the $D_2^{\mathbb{P}_1}$ data.

P3_illustration.txt For illustrating the range of p_3 in the $D_2^{\mathbb{P}_1}$ data.

REFERENCES

Abd-El-Malek, S. (1938). A contribution of the study of the movements of the tongue in animals, with special reference to the cat. *Journal of Anatomy*, **73**.

Abd-El-Malek, S. (1939). Observations on the morphology of the human tongue. *Journal of Anatomy*, **73**.

Abd-El-Malek, S. (1955). The part played by the tongue in mastication and deglutition. *Journal of Anatomy*, **89**(2), pp. 250–4.

Anderson, C. and A. Fuglevand (2008). Probability-Based Prediction of Activity in Multiple Arm Muscles: Implications for Functional Electrical Stimulation. *Journal of Neurophysiology*, **100**(1), p. 482.

Atal, S., J. Chang, J. Mathews, and W. Tukey (1978). Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. *Journal of the Acoustical Society of America*, **63**, pp. 1535–1555.

Atsumi, T. and T. Miyatake (1987). Morphometry of the degenerative process in the hypoglossal nerves in amyotrophic lateral schlerosis. *Acta Neuropathologica*, **73**.

Baer, T., P. Alfonso, and K. Honda (1988). Electromyography of the tongue muscles during vowels in/əpVp/environment. *Ann. Bull. RILP, Univ. Tokyo*, **7**, pp. 7–18.

Bailey, E. F. and R. F. Fregosi (2004). Coordination of intrinsic and extrinsic tongue muscles during spontaneous breathing in the rat. *Journal of Applied Physiology*, **96**, pp. 440–449.

Bailey, E. F., Y.-H. Huang, and R. F. Fregosi (2006). Anatomic consequences of intrinsic tongue muscle activation. *Journal of Applied Physiology*, **101**, pp. 1377–1385.

Bailey, E. F., A. D. Rice, and A. J. Fuglevand (2007). Firing Patterns of Human Genioglossus Motor Units During Voluntary Tongue Movement. *Journal of Neurophysiology*, **97**, pp. 933–936.

Bailly, G. (1997). Learning to speak. Sensori-motor control of speech movements. *Speech Communication*, **22**, pp. 251–267.

Barron, G. (1993). History and Development. In r. Brauer, J. (ed.) *What Every Engineer Should Know About Finite Element Analysis, Second Edition, Revised and Expanded*. Marcel Dekker, Inc.

Blemker, S., P. Pinsky, and S. Delp (2005). A 3D model of muscle reveals the causes of nonuniform strains in the biceps brachii. *Journal of Biomechanics*, **38**(4), pp. 657–665.

Blemker, S. S. (2004). *3D Modeling of Complex Muscle Architecture Geometry*. Ph.D. thesis, Stanford University.

Boersma, P. (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. Ph.D. thesis, University of Amsterdam.

Bonet, J. and R. Wood (1997). *Nonlinear Continuum Mechanics for Finite Element Analysis*. Cambridge University Press.

Bowman, J. (1971). *The Muscle Spindle and Neural Control of the Tongue*. Thomas, Springfield, Ill.

Browman, C. P. and L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology*, **6**.

Chiel, H., P. Crago, J. Mansour, and K. Hathi (1992). Biomechanics of a muscular hydrostat: a model of lapping by a reptilian tongue. *Biological Cybernetics*, **67**.

Chomsky, N. and M. Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, **2**.

Cooper, S. (1953). Muscle spindles in the intrinsic muscles of the human tongue. *Journal of Physiology*, **122**, pp. 192–202.

Dang, J. and K. Honda (1998). Speech Production of Vowel Sequences Using a Physiological Articulatory Model. In *Proceedings of ICSLP 98*, pp. XX–XX.

Dang, J. and K. Honda (2001). A physiological articulatory model for simulating speech production process. *Acoustic Science & Technology*, **22**(6), pp. 415–425.

Dang, J. and K. Honda (2002). Estimation of vocal tract shapes from speech sounds with a physiological articulatory model. *Journal of Phonetics*, **30**.

Doran, G. and H. Baggett (1971). A structural and functional classification of mammalian tongues. *Journal of Mammology*, **52**(2), pp. 427–429.

Duffy, J. R. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management, Second Edition*. Elsevier Mosby.

English, A., S. Wolf, and R. Segal (1993). Compartmentalization of muscles and their motor nuclei: the partitioning hypothesis. *Physical Therapy*, **73**(12), p. 857.

English, A. W. and W. D. Letbetter (1982a). Anatomy and Innervation Patterns of Cat Lateral Gastrocnemius and Plantaris Muscles. *The Anatomical Record*, **164**, pp. 67–77.

English, A. W. and W. D. Letbetter (1982b). A Histochemical Analysis of Identified Compartments of Cat Lateral Gastrocnemius Muscle. *The Anatomical Record*, **204**, pp. 123–130.

Engwall, O. (2003). Combining MRI, EMA and EPG measurements in a three-dimensional tongue model. *Speech Communication*, **41**, pp. 303–329.

Enoka, R. M. and A. J. Fuglevand (2001). Motor unit physiology: some unresolved issues. *Muscle & Nerve*, **24**, pp. 4–17.

Fant, G. (1970). *Acoustic Theory of Speech Production with Calculations Based on X-ray Studies of Russian Articulations*. Mouton De Gruyter.

Feldman, A. (1986). Once more on the equilibrium-point hypothesis (lambda model) for motor control. *Journal of Motor Behavior*, **18**(1), pp. 17–54.

Felton, S. M., T. A. Gaige, T. G. Reese, V. J. Wedeen, and R. J. Gilbert (2007). Mechanical basis for lingual deformation during the propulsive phase of swallowing as determined by phase contrast magnetic resonance imaging. *Journal of Applied Physiology*, **103**(1), pp. 255–365.

Forster, L. (1894). Zur Kenntnis der Muskelspindeln. *Virchows Arch*, **137**, pp. 121–154.

Fujita, S., J. Dang, N. Suzuki, and K. Honda (2007). A Computational Tongue Model and its Clinical Application. *Oral Science International*, **4**(2), pp. 97–109.

Fung, Y.-C. (1993). *Biomechanics: Mechanical Properties of Living Tissues*. Springer-Verlag.

Garland, M. and P. S. Heckbert (1997). Surface simplification using quadric error metrics. *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pp. 209–216.

Gerárd, J.-M., J. Ohayon, V. Luboz, P. Perrier, and Y. Payan (2005). Non-linear elastic properties of the lingual and facial tissues assessed by indentation technique: Application to the biomechanics of speech production. *Medical Engineering & Physics*, **27**, pp. 884–892.

Gérard, J.-M., P. Perrier, and Y. Payan (2004). Modélisation biomécanique tri-dimensionnelle de l'articulateur lingual pour étudier le contrôle de la parole. In *Journées d'Etude sur la Parole 2004*.

lAFCP. Available http://www.lpl.univ-aix.fr/jep-taln04/proceed/actes/jep2004/Gerard-Perrier-Payan.pdf.

Gérard, J.-M., P. Perrier, and Y. Payan (2006). 3D biomechanical tongue modeling to study speech production. In Harrington, J. and M. Tabain (eds.) *Speech Production: Models, Phonetic Processes, and Techniques*. New York: Psychology Press.

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

Gilbert, R. J., L. H. Magnusson, V. J. Napadow, T. Benner, R. Wang, and V. J. Wedeen (2006). Mapping Complex Myoarchitecture in the Bovine Tongue with Diffusion-Spectrum Magnetic Resonance Imaging. *Biophysical Journal*, **91**, pp. 1014–1022.

Gilbert, R. J. and V. J. Napadow (2005). Three-Dimensional Muscular Architecture of the Human Tongue Determined In Vivo With Diffusion Tensor Magnetic Resonance Imaging. *Dysphagia*, **20**, pp. 1–7.

Gray, H. (1918). *Anatomy of the Human Body, Twentieth Edition*. Philadelphia: Lea & Febiger.

Guccione, J. and A. McCulloch (1993). Mechanics of active contraction in cardiac muscle: Part I–Constitutive relations for fiber stress that describe deactivation. *Journal of Biomechanical Engineering*, **115**(1), pp. 72–81.

Guccione, J., L. Waldman, and A. McCulloch (1993). Mechanics of active contraction in cardiac muscle: Part II–Cylindrical Models of the Systolic Left Ventricle. *Journal of Biomechanical Engineering*, **115**(1), pp. 82–90.

Guenther, F. H. (1992). *Neural models of adaptive sensori-motor control for flexible reaching and speaking*. Ph.D. thesis, Boston University.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, **102**(3), pp. 594–621.

Guenther, F. H. and J. S. Perkell (2003). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In Maassen, B., R. Kent, H. Peters, P. V. Lieshout, and W. Hulstijn (eds.) *Speech motor control in normal and disordered speech*, pp. 29–49. Oxford: Oxford University Press.

Guo, Y., S. Goldberg, and J. McClung (1996). Compartmental organization of styloglossus and hyoglossus motoneurons in the hypoglossal nucleus of the rat. *Brain Research*, **728**(2), pp. 277–280.

Halle, M. (1983). On distinctive features and their articulatory implementation. *Natural Language and Linguistic Theory*, **1**, pp. 91–105.

Halle, M. (1992). Phonological features. In Bright, W. (ed.) *International encyclopedia of linguistics, vol. 3*. Oxford, Oxford University Press.

Hamilton, A. F., K. E. Jones, and D. M. Wolpert (2004). The scaling of motor noise with muscle strength and motor unit number in humans. *Experimental Brain Research*, **157**.

Hardcastle, W. J. (1976). *Physiology of Speech Production: An Introduction for Speech Scientists*. Academic Press.

Harshman, R. A., P. N. Ladefoged, and L. Goldstein (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, **62**, pp. 693–707.

Hashimoto, K. and S. Suga (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.

Hayashi, Y., E. Hoashi, and T. Nara (1997). Ultrasonographic analysis of sucking behavior of newborn infants: the driving force of sucking pressure. *Early Human Development*, **49**, pp. 33–38.

Henneman, E. (1979). Functional organization of motoneurons pools: the size principle. In Asunuma, H. and V. Wilson (eds.) *Integration in the Nervous System*, pp. 13–25. Tokyo: Igaku Shoin.

Hiiemae, K. M. and J. B. Palmer (2003). Tongue movements in feeding and speech. *Critical reviews in oral biology and medicine*, **14**(6), pp. 413–429.

Hill, A. (1938). The Heat of Shortening and the Dynamic Constants of Muscle. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **126**(843), pp. 136–195.

Hirano, M., Y. Kuroiwa, S. Tanaka, H. Matsuoka, K. Sato, and T. Yoshida (1992). Dysphagia following various degrees of surgical resection for oral cancer. *Annals of Otology, Rhinology & Laryngology*, **101**(2 Pt 1), pp. 138–41.

Holzapfel, G. A. and T. C. Gasser (2000). A viscoelastic model for fiber-reinforced composites at finite strains: Continuum basis, computational aspects and applications. *Computer Methods in Applied Mechanics and Engineering*, **190**, pp. 4379–4430.

Iskander, A. and I. Sanders (2003). Morphological comparison between neonatal and adult human tongues. *Annals of Otology, Rhinology & Laryngology*, **112**(9), pp. 768–776.

Jakobson, R., C. G. Fant, and M. Halle (1951). *Preliminaries to Speech Analysis*. Cambridge: The MIT Press.

Johansson, T., P. Meier, and R. Blickhan (2000). A finite-element model for the mechanical analysis of skeletal muscles. *Journal of Theoretical Biology*, **206**(1), pp. 131–49.

Jolliffe, I. (2002). *Principal component analysis, second edition.* Springer Series in Statistics. Springer.

Jürgens, U. and M. Alipour (2002). A comparative study on the cortico-hypoglossal connections in primates, using biotin dextranamine. *Neuroscience Letters*, **328**(3), pp. 245–248.

Kakita, Y., O. Fujimura, and K. Honda (1985). Computation of mapping from muscular contraction patterns to formant patterns in vowel space. In Fromkin, V. A. (ed.) *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. Academic Press.

Keen, D. and A. Fuglevand (2004a). Common Input to Motor Neurons Innervating the Same and Different Compartments of the Human Extensor Digitorum Muscle. *Journal of Neurophysiology*, **91**(1), pp. 57–62.

Keen, D. and A. Fuglevand (2004b). Distribution of Motor Unit Force in Human Extensor Digitorum Assessed By Spike-Triggered Averaging and Intraneural Microstimulation. *Journal of Neurophysiology*, **91**(6), pp. 2515–2523.

Kent, R. D. (2004). Motor control in disorders. Motor control perspectives on motor speech disorders. In Maassen, B. (ed.) *Speech motor control in normal and disordered speech*. Oxford University Press.

Kier, W. M. and K. K. Smith (1985). Tongues, tentacles, and trunks: The biomechanics of movement in muscular hydrostats. *Zoological Journal of the Linnean Society*, **83**.

Kim, S., A. S. Barnett, C. Pierpaoli, and G. Chi-Fishman (forthcoming). Three-dimensional mapping of lingual myoarchitecture by diffusion tensor MRI. *NMR in Biomedicine*, —, pp. —. Early view.

Kiritani, S., K. Miyawaki, O. Fujimura, and J. Miller (1976). A computational model of the tongue. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo*, **10**.

Kuypers, H. (1958). Corticobulbar connexions to the pons and lower brain-stem in man. *Brain*, **81**, pp. 364–388.

Laboissière, R., D. Ostry, and A. Feldman (1996). The control of multi-muscle systems: human jaw and hyoid movements. *Biological Cybernetics*, **74**(4), pp. 373–384.

Lai, W., D. Rubin, and E. Krempl (1993). *Introduction to Continuum Mechanics, 3rd ed.)*. Elsevier.

Legrice, I., P. Hunter, and B. Smaill (1997). Laminar structure of the heart: a mathematical model. *American Journal of Physiology- Heart and Circulatory Physiology*, **272**(5), pp. 2466–2476.

Lemon, R. N. and J. Griffiths (2005). Comparing the function of the corticospinal system in different species: organizational differences for motor specialization? *Muscle Nerve*, **32**, pp. 261–279.

Liu, X., M. Stone, and J. Prince (2006). Tracking tongue motion in three dimensions using tagged MR images. In *Proceedings of the IEEE ISBI*.

Logemann, J. A. (1983). *Evaluation and treatment of swallowing disorders*. College-Hill Press.

Lowe, A. (1981). The neural regulation of tongue movements. *Progress in Neurobiology*, **15**, pp. 295–344.

Lucero, J. and K. Munhall (1999). A model of facial biomechanics for speech production. *The Journal of the Acoustical Society of America*, **106**, p. 2834.

Lundberg, A. and M. Stone (1999). Three-dimensional tongue surface reconstruction: Practical considerations for ultrasound data. *The Journal of the Acoustical Society of America*, **106**, p. 2858.

Maas, S. and J. Weiss (2008). *FEBio Theory Manual*. Musculuskeletal Research Laboratories, University of Utah.

Markey, K. L. (1994). *The sensorimotor foundations of phonology: a computational model of early childhood articulatory and phonetic development*. Ph.D. thesis, University of Colorado, Boulder.

Marr, D. (1982). *Vision*. Freeman New York.

Martins, J., E. Pires, R. Salvado, and P. Dinis (1998). A numerical model of passive and active behavior of skeletal muscles. *Computer Methods in Applied Mechanics and Engineering*, **151**(3-4), pp. 419–433.

McClung, J. and S. Goldberg (2000). Functional anatomy of the hypoglossal innervated muscles of the rat tongue: A model for elongation and protrusion of the mammalian tongue. *The Anatomical Record*, **260**(4), pp. 378–386.

McConnel, F., J. Logemann, A. Rademaker, B. Pauloski, S. Baker, J. Lewin, D. Shedd, M. Heiser, S. Cardinale, S. Collins, et al. (1994). Surgical Variables Affecting Postoperative Swallowing Efficiency in Oral Cancer Patients: A Pilot Study. *Laryngoscope*, **104**(1), p. 87.

McConnel, F., B. Pauloski, J. Logemann, A. Rademaker, L. Colangelo, D. Shedd, W. Carroll, J. Lewin, and J. Johnson (1998). Functional Results of Primary Closure vs Flaps in Oropharyngeal Reconstruction A Prospective Study of Speech and Swallowing. *Archives of Otolaryngology—Head and Neck Surgery*, **124**(6), pp. 625–630.

Meier, P. and R. Blickhan (2000). FEM-Simulation of skeletal muscle: the influence of inertia during activation and deactivation. In Herzog, W. (ed.) *Skeletal Muscle Mechanics: from mechanisms to function*, pp. 207–223. Wiley.

Meyers, J. J., J. C. OReilly, J. A. Monroy, and K. C. Nishikawa (2004). Mechanism of tongue protraction in microhylid frogs. *The Journal of Experimental Biology*, **207**, pp. 21–31.

Miller, J. L., K. L. Watkin, and M. F. Chen (2002). Muscle, Adipose, and Connective Tissue Variations in Intrinsic Musculature of the Adult Human Tongue. *Journal of Speech, Language, and Hearing Research*, **45**, pp. 51–65.

Miyawaki, K. (1974). A study of the musculature of the human tongue. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo*, **8**.

Miyawaki, K., H. Hirose, T. Ushijima, and M. Sawashima (1975). A Preliminary Report on the Electromyographic Study of the Activity of Lingual Muscles. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo*, **9**.

Mu, L. and I. Sanders (1999). Neuromuscular organization of the canine tongue. *The Anatomical Record*, **256**(4), pp. 412–424.

Mu, L. and I. Sanders (2000). Neuromuscular Specializations of the Pharyngeal Dilator Muscles: II. Compartmentalization of the Canine Genioglossus Muscle. *The Anatomical Record*, **260**, pp. 308–325.

Napadow, V., R. Kamm, and R. Gilbert (2002). A Biomechanical Model of Sagittal Tongue Bending. *Journal of Biomechanical Engineering*, **124**, pp. 547–556.

National Library of Medicine (1987). The Visible Human Project. `http://www.nlm.nih.gov/research/visible/visible_human.html`.

Nishikawa, K. C. (1999). Neuromuscular control of prey capture in frogs. *Philosophical transactions of the Royal Society of London*, **354**, pp. 951–954.

Nishikawa, K. C., W. M. Kier, and K. K. Smith (1999). Morphology and mechanics of tongue movement in the African pignosed frog *hemisus marmoratum*: a muscular hydrostat model. *The Journal of Experimental Biology*, **202**, pp. 771–780.

Nolte, J. (2002). *The Human Brain: An Introduction to Its Functional Anatomy, 5th Edition*. Mosby.

O'Kusky, J. R. and M. G. Norman (1995). Sudden Infant Death Syndrome: Increased Number of Synapses in the Hypoglossal Nucleus. *Journal of Neuropathology and Experimental Neurology*, **54**(5).

Özkaya, N. and M. Nordin (1991). *Fundamentals of Biomechanics: Equilibrium, Motion, and Deformation*. Van Nostrand Reinhold.

Payan, Y. and P. Perrier (1997). Synthesis of V-V Sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis. *Speech Communication*, **22**(2/3), pp. 185–205.

Perkell, J. S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Research Monograph No. 53. The M.I.T. Press.

Perkell, J. S. (1974). *A physiologically-oriented model of tongue activity in speech production*. Ph.D. thesis, MIT.

Perrier, P., J. Perkell, Y. Payan, M. Zandipour, F. Guenther, and A. Khalighi (2000). Degrees of freedom of tongue movements in speech may be constrained by biomechanics. In *Proceedings of the Sixth International Conference on Spoken Language Processing, ICSLP-2000*, volume 2.

Rodrigues, M., D. Gillies, and P. Charters (2001). Realistic Deformable Models for Simulating the Tongue during Laryngoscopy. *Proceedings of the International Workshop on Medical Imaging and Augmented Reality (MIAR'01)*.

Rogers, D. (2001). *An Introduction to NURBS: With Historical Perspective*. Morgan Kaufmann.

Saboisky, J. P., J. E. Butler, D. K. McKenzie, R. B. Gorman, J. A. Trinder, D. P. White, and S. C. Gandevia (2007). Neural drive to human genioglossus in obstructive sleep apnoea. *Journal of Physiology*, **585**.

Saito, H. and I. Itoh (2003). Three-dimensional architecture of the intrinsic tongue muscles, particularly the longitudinal muscle, by the chemical-maceration method. *Anatomical Science International*, **78**, pp. 168–176.

Saito, H. and I. Itoh (2007). The three-dimensional architecture of the human styloglossus especially its posterior muscle bundles. *Annals of Anatomy*, **189**(3), pp. 261–267.

Salter, H. (1852). Tongue. In todd, R. B. (ed.) *The Cyclopaedia of Anatomy and Physiology*, volume 4, pp. 1120–1163. London: Longman, Brown, Green, Longmans, & Roberts.

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

Sanguineti, V., R. Laboissière, and Y. Payan (1997). A control model of human tongue movements in speech. *Biological Cybernetics*, **77**, pp. 11–22.

Shinagawa, H., E. Murano, J. Zhuo, B. Landman, R. Gullapalli, J. Prince, and M. Stone (2008). Tongue muscle fiber tracking during rest and tongue protrusion with oral appliances: A preliminary study with diffusion tensor imaging. *Acoustical Science and Technology*, **29**(4), pp. 291–294.

Slaughter, K., H. Li, and A. J. Sokoloff (2005). Neuromuscular Organization of the Superior Longitudinalis Muscle in the Human Tongue. *Cells Tissues Organs*, **181**, pp. 51–64.

Smith, K. (1989). Histological Demonstration of Muscle Spindles in the Tongue of the Rat. *Archives of Oral Biology*, **34**(7), pp. 529–534.

Smith, W. L., A. Erenburg, A. Novak, and J. Edmund A Franken (1985). Physiology of Sucking in the Normal Term Infant using Real-Time US. *Radiology*, **156**, pp. 379–381.

Sokoloff, A. and T. Deacon (1992). Musculotopic organization of the hypoglossal nucleus in the cynomolgus monkey, Macaca fascicularis. *Journal of Comparative Neurology*, **324**, pp. 81–93.

Sokoloff, A. J. (2000). Localization and Contractile Properties of Intrinsic Longitudinal Motor Units of the Rat Tongue. *Journal of Physiology*, **84**, pp. 827–835.

Sokoloff, A. J. (2004). Activity of tongue muscles during respiration: it takes a village? *Journal of Applied Physiology*, **96**, pp. 438–439.

Stål, P., S. Marklund, L.-E. Thornell, R. D. Paul, and P.-O. Eriksson (2003). Fibre Composition of Human Intrinsic Tongue Muscles. *Cells Tissues Organs*, **173**, pp. 147–161.

Stachler, R., S. Hamlet, R. Mathog, L. Jones, L. Heilbrun, L. Manov, and J. O'Campo (1994). Swallowing of bolus types by postsurgical head and neck cancer patients. *Head Neck*, **16**(5), pp. 413–9.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, **17**.

Stone, M. and A. Lundberg (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *The Journal of the Acoustical Society of America*, **99**, pp. 3728–3737.

Story, B. H. and I. R. Titze (1998). Parametrization of vocal tract area functions by empirical orthogonal modes. *Journal of Phonetics*, **26**(3), pp. 223–260.

Story, B. H., I. R. Titze, and E. Hoffman (1996a). Vocal tract area functions from magnetic resonance imaging. *Journal of Acoustical Society of America*, **100**(1).

Story, B. H., I. R. Titze, and E. A. Hoffman (1996b). Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America*, **100**(1), pp. 537–554.

Strong, L. H. (1956). Muscle fibers of the tongue functional in consonant production. *The Anatomical Record*, **126**(1), pp. 61–79.

Takemoto, H. (2001). Morphological Analysis of the Human Tongue Musculature for Three-Dimensional Modeling. *Journal of Speech, Language, and Hearing Research*, **44**.

Teran, J., E. Sifakis, S. Blemker, V. Ng-Thow-Hing, C. Lau, and R. Fedkiw (2005). Creating and simulating skeletal muscle from the visible human data set. *IEEE Transactions on Visualization and Computer Graphics*, **11**(3), pp. 317–328.

Touré, G., L. Bicchieray, J. Selva, and C. Vacher (2005). The intra-lingual course of the nerves of the tongue. *Surgical and Radiologic Anatomy*, **27**(4), pp. 297–302.

Tsuiki, S., T. Ono, Y. Ishiwata, and T. Kuroda (2000). Functional divergence of human genioglossus motor units with respiratory-related activity. *European Respiratory Journal*, **15**, pp. 906–910.

Uemura-Sumi, M., M. Itoh, and N. Mizuno (1988). The distribution of hypoglossal motoneurons in the dog, rabbit and rat. *Anatomy and Embryology*, **177**(5), pp. 389–394.

Uemura-Sumi, M., N. Mizuno, S. Nomura, N. Iwahori, Y. Takeuchi, and R. Matsushima (1981). Topographical representation of the hypoglossal nerve branches and tongue muscles in the hypoglossal nucleus of macaque monkeys. *Neurosci Lett*, **22**(1), pp. 31–5.

van den Doel, K., F. Vogt, R. E. English, and S. Fels (2006). Towards Articulatory Speech Synthesis with a Dynamic 3D Finite Element Tongue Model. In *Proceedings of ISSP '06*.

van Leeuwen, J. L. (1992). Muscle function in locomotion. In Alexander, R. M. (ed.) *Advances in Comparative and Environmental Physiology: Mechanics of Animal Locomotion*, pp. 191–250. Springer-Verlag.

Verestóy, J., D. Chetverikov, and M. Nagy (1999). *Digital PIV: a Challenge for Feature Based Tracking.* Available: `http://visual.ipan.sztaki.hu/pivweb/pivweb.html`.

Weiss, J., B. Maker, and S. Govindjee (1996). Finite element implementation of incompressible, transversely isotropic hyperelasticity. *Computer Methods in Applied Mechanics and Engineering*, **135**(1), pp. 107–128.

Wilhelms-Tricarico, R. (1995). Physiological modeling of speech production: methods for modeling soft-tissue articulators. *Journal of Acoustical Society of America*, **97**(5).

Wilhelms-Tricarico, R. (1996). A biomechanical and physiologically–based vocal tract model and its control. *Journal of Phonetics*, **24**(1), pp. 23–38.

Wilhelms-Tricarico, R. (2000a). Development on a tongue and mouth floor model for normalization and biomechanical modelling. In *Proceedings the 5th Speech Production Seminar*. Kloster Seon, Bavaria.

Wilhelms-Tricarico, R. (2000b). Development on a tongue and mouth floor model for normalization and biomechanical modelling. In *Proceedings of the 5th Speech Production Seminar, Kloster Seon, Bavaria*.

Wilhelms-Tricarico, R. (2005). A 3D dynamical biomechanical tongue model to study speech motor control. University of Alberta.

Wožniak, W. and P. A. Young (1969). Further observations on human hypoglossal nerve. *Anatomischer Anzeiger*, **125**.

Yamaguchi, G. T. (2001). *Dynamic modeling of musculoskeletal motion : a vectorized approach for biomechanical analysis in three dimensions*. Kluwer Academic Publishers.

Yang, C. and M. Stone (2002). Dynamic programming method for temporal registration of three-dimensional tongue surface motion from multiple utterances. *Speech Communication*, **38**(1-2), pp. 201–209.

Zajac, F. E. (1989). Muscle and tendon: properties, models, scaling, and applications to biomechanics and motor control. *Critical Reviews in Biomedical Engineering*, **17**(4), pp. 359–411.

Zemlin, W. R. (1988). *Speech and Hearing Science: Anatomy and Physiology, Third Edition*. Prentice Hall.