Title
Quantal biomechanical effects in speech postures of the lips

Journal
Journal of Neurophysiology, 124(3)

ISSN
0022-3077

Authors
Gick, Bryan
Mayer, Connor
Chiu, Chenhao
et al.

Publication Date
2020-09-01

DOI
10.1152/jn.00676.2019

Quantal biomechanical effects in speech postures of the lips

Bryan Gick[1], Connor Mayer[2], Chenhao Chiu[3], Erik Widing[4], Francois Roewer-Despres[4], Sidney Fels[5], & Ian Stavness[4]

[1] Department of Linguistics, University of British Columbia

[2] Department of Linguistics, University of California, Los Angeles

[3] Graduate Institute of Linguistics, National Taiwan University

[4] Department of Computer Science, University of Saskatchewan

[5] Department of Electrical and Computer Engineering, University of British Columbia

Corresponding author: Connor Mayer. connormayer@ucla.edu. Department of Linguistics. 3125 Campbell Hall, UCLA. Los Angeles, CA. 90095-1543.

Abstract

The unique biomechanical and functional constraints on human speech make it a promising area for research investigating modular control of movement. The present paper illustrates how a modular control approach to speech can provide insights relevant to understanding both motor control and observed variation across languages. We specifically explore the robust typological finding that languages produce different degrees of labial constriction using distinct muscle groupings and concomitantly distinct lip postures. Research has suggested that these lip postures exploit biomechanical regions of non-linearity between neural activation and movement, also known as *quantal regions*, to allow movement goals to be realized despite variable activation signals. We present two sets of computer simulations showing that these labial postures can be generated under the assumption of modular control, and that the corresponding modules are biomechanically robust: first to variation in the activation levels of participating muscles, and second to interference from surrounding muscles. These results provide support for the hypothesis that biomechanical robustness is an important factor in selecting the muscle groupings used for speech movements, and provide insight into the neurological control of speech movements and how biomechanical and functional constraints govern the emergence of speech motor modules. We anticipate that future experimental work guided by biomechanical simulation results will provide new insights into the neural organization of speech movements.

New and noteworthy: This paper provides additional evidence that speech motor control is organized in a modular fashion, and that biomechanics constrain the kinds of motor modules that may emerge. It also suggests that speech can be a fruitful domain for the study of modularity, and that a better understanding of speech motor modules will be useful for speech research.

Finally, it suggests that biomechanical modeling can serve as a useful complement to experimental work when studying modularity.

Keywords: biomechanical simulation, quantal effects, speech, phonetics, motor control

Introduction

Decades of literature have substantiated the view that biological movement control is organized in a modular fashion: that is, the central nervous system does not exert control at the level of individual muscles, but rather has access to a limited repertoire of coordinated patterns of muscle activity that may be combined freely to achieve different functional goals (e.g., Bernstein 1967, Tresch et al. 2002, d'Avella et al. 2003, d'Avella and Bizzi 2005, Bizzi et al. 2008, Berniker et al. 2009, de Rugy et al. 2012, Overduin et al. 2012, Ting et al. 2015).

Bernstein (1967) proposed that this modularization simplifies motor control by reducing the degrees of freedom in the control space. Subsequent research has questioned this claim (e.g., Radhakrishnan et al. 2008, Overduin et al. 2012), suggesting instead that modules might facilitate effective control by simplifying the generation of frequently executed movements and allowing them to be produced with lower variability (Diedrichsen and Classen 2012), as well as facilitating the learning of new movement patterns (d'Avella 2016).

Evidence for modular control has come in two forms. First, there is evidence that motor behavior can be adequately *described* using a combination of a small set of coordinated patterns of muscle activity. Many of the results from kinematic (e.g., Santello et al. 1998) and electromyographic (EMG) studies (e.g., d'Avella et al. 2003, d'Avella and Bizzi 2005) fall into this category. These observations, while suggestive, do not necessarily imply that movements are *controlled* in a modular fashion: as pointed out by Diedrichen and Classen (2012), the descriptive adequacy of modules may emerge from constraints imposed by specific tasks (Diedrichsen et al. 2010) and the properties of the musculoskeletal system (Kutch and Valero-Cuevas 2012). The second type of evidence, that movements are controlled in a modular fashion, is thus crucially important for supporting claims of modular organization. It comes primarily from studies that have applied electrical neuronal stimulation to non-human vertebrates (e.g., Bizzi et al. 1991, Overduin et al. 2012), though some indirect support has come from records of kinematic activity following TMS stimulation (e.g., Gentner and Classen 2006) and from clinical studies (e.g., Cheung et al. 2009, Clark et al. 2010).

The literature on modular control generally agrees that motor modules are *learned* or *emergent*. Previous work has suggested that modules emerge from associations between neural control strategies on the one hand and biomechanical and task constraints on the other (e.g., Ting and McKay 2007, de Rugy et al. 2012, Ting and Chiel 2017). Ting and Chiel (2017) write that "biomechanical considerations are critical for understanding the neural control of movement," and that "the effects of neural signals on motor outputs is highly non-linear and context dependent, relying on the specific biomechanical constraints and affordances of the motor periphery" (401).

An open question here concerns what factors might cause a particular structure to emerge over another one: that is, given competing candidate modules that are sufficient to meet the requirements of a particular task given the constraints imposed by the biomechanics, what might prioritize the adoption of certain ones over others? Loeb et al. (2000) use computer simulations to identify a small set of muscle synergies in the frog that reliably produce leg stabilization in predictable locations under activation noise. These synergies are found to be similar to synergies identified from experimental work, suggesting that *biomechanical robustness* may be an important criterion in the development of motor modules.

In the present paper, we propose on the basis of neuromechanical simulations that biomechanical robustness may also govern the emergence of speech motor modules, which must similarly produce predictable outcomes in the presence of noise. Specifically, we suggest that the neural controller may adopt modules that exploit the non-linear, or *quantal*, biomechanical relationships between neural activation and movement in order to produce consistent outcomes in the face of variable activation and interference from surrounding muscles.

Modular organization in speech motor control

Speech is a promising and little-studied area of modular motor organization. Very little research on motor modularization has dealt with speech directly, despite proposals in the literature that speech should be further studied from this perspective (Fowler et al. 1980, Gick and Stavness 2013, Ting et al. 2015, Gick et al. 2019a, Gick et al. 2019b).

Ting et al. (2015) note that while there is some interpersonal variation in walking style, it is relatively limited due to strict biomechanical constraints and the possibility that unfit variants might lead to a fall. Speech, on the other hand, involves countless combinations of highly specialized movement sequences, and is less biomechanically constrained than walking, as unfit variants do not produce the same disastrous outcomes. This has allowed for the emergence of a wide range of modules for handling similar tasks, even within a language (e.g., Delattre and Freeman 1968, Benguerel et al. 1977, Derrick and Gick 2011). Speech thus provides a rich context for studying variability and emergence of motor modules.

A challenge in studying speech from the perspective of modularization is that the speech articulators are biomechanically different from the more frequently studied musculoskeletal structures of limbs and digits, and many of their functional properties are not well understood. The tongue, for example, is often described as a muscular hydrostat (Kier and Smith 2005), composed primarily of muscle fiber and lacking any intrinsic rigid support that might mechanically constrain its degrees of freedom. Muscular hydrostats maintain a constant volume, so that compressions along one dimension result in expansions distributed along other dimensions, complicating the relationship between muscle activation and movement.

An additional challenge is the difficulty of collecting the neurophysiological data commonly used in such studies. Speech muscles are generally very small, hard to access, and heavily interdigitated, making them hard to measure using traditional neurophysiological methods (see, e.g., Kuehn et al. 1982, Anderson et al. 2019). Even for simple, non-speech tongue movements, recording representative neuromuscular activity from EMG has been shown to be difficult (Pittman and Bailey 2009). Furthermore, the muscles used in speech are primarily innervated by cranial nerves, reducing the value of insights from previous experiments on non-human spinal structures (e.g., Tresch et al. 2002).

Due in part to these difficulties, the literature on speech motor control has focused primarily on modeling other aspects of speech, such as control space (e.g., Turvey 1978, Kelso et al. 1986) and sensory processing (e.g., Guenther 2016). A limited number of studies have looked for indirect evidence of modular organization in speech based on imaging data (e.g., Ramanarayanan et al. 2013, Moisik et al. 2017). Other studies have applied an auditory startle paradigm (Valls-Solé et al. 1999) to speech, showing that the jaw and lip movement involved in producing the syllable ‘ba’ can become decoupled when participants are startled before producing them, which suggests that independent jaw and lip modules may contribute to the closure for ‘b’ (Chiu and Gick 2015, Chiu 2016). Psycholinguistic studies based on speech errors or reaction time measurements have also suggested that constituents of varying sizes, such as syllables, may be drawn upon in a modular fashion for speech production (e.g., Kozhevnikov and Chistovich 1965, Crompton 1981, MacNeilage and Davis 1990, Levelt and Wheeldon 1994, Levelt et al. 1999, Cholin et al. 2006, Guenther et al. 2006, Cholin et al. 2011, Guenther 2016).

The difficulty of obtaining reliable and comprehensive neurophysiological measurements has made biomechanical simulation a useful method for studying the function of muscle groupings in speech. Such neuromechanical models have been proposed to be an important tool for the study of modular control (Ting and McKay 2007) and have been applied in much previous work (e.g., Loeb et al. 2000, Berniker et al. 2009, Kutch and Valero-Cuevas 2012; see Ting et al. 2015 for additional references). Biomechanical simulations predict movements that result from prescribed forces, e.g., muscle forces, by solving equations that represent the mechanics of the musculoskeletal system. These methods have been applied to various complex structures involved in speech, such as the tongue (Stavness et al. 2011, Gick et al. 2017, Mayer et al. 2018), the soft palate (Gick et al. 2014, Anderson et al. 2019), the larynx (Moisik and Gick 2017), and the lips (Nazari et al. 2010, 2011, Stavness et al. 2013, Payan et al. 2019).

The lips in speech

The lips are highly complex, despite their superficially simple appearance. Though speech researchers have alternately treated them as a single functional sphincter (e.g., The UCLA Phonetics Laboratory 2002), or as two or more independent structures, like the ‘upper lip’ and ‘lower lip’ (e.g., Kelso et al. 1986, Guenther 2016), neither of these views appears to be completely correct. The lips are controlled by a network of muscles that runs from the top of the forehead to the sternum. Many of these muscles exhibit a high degree of morphological complexity. The orbicularis oris (OO) muscle, for example, which encircles the mouth and contributes to many of the speech movements we will discuss below, consists of four distinct loops of muscle which are layered at three different depths (Lightoller 1925, Kraus et al. 1966, Larrabee and Makielski 1993, Rogers et al. 2008, Stavness et al. 2013, Standring 2015). Because of this complexity, the function of the lip muscles in speech is not well understood.

It is common for different speech sounds to be produced using different sized openings at roughly the same anatomical location. In English, for example, we may produce sounds at the lips with complete closure, known as *stops* (such as the first sounds in ‘pail’, ‘bail’, and ‘mail’), sounds with narrow constrictions that produce turbulent noise, known as *fricatives* (e.g., the first sounds in ‘fail’ and ‘veil’), and sounds with constrictions that are not sufficiently narrow to produce turbulence, known as *approximants* (e.g., the first sound in ‘wail’). Although we may consider all these constrictions to be made ‘using the lips’, they use quite different parts of the lips, and are generated by categorically different mechanisms. The English labial stops are *bilabial*, with oral airflow being completely blocked by bringing both lips together in a wide, flat constriction; labial fricatives are *labiodental,* with a narrow constriction between the lower lip and upper teeth while the upper lip is relaxed or pulled upward; and labial approximants are *rounded,* with constriction produced by rounding and protruding the lips. Each of these mechanisms is known to be associated with different muscle groupings, suggesting separate motor modules underpinning their control.

The use of these particular mechanisms for these different degrees of labial constriction is strikingly robust across languages of the world. Of the 451 languages in the UCLA Phonological Segment Inventory Database (Maddieson 1984, Maddieson and Precoda 1990), 446 (99%) contain bilabial stops. Only one language (0.2%) is reported as having a labiodental stop, where complete closure is made between the lower lip and upper teeth, and none are reported with a stop produced exclusively by rounding. Conversely, for labial fricatives, 199 (44%) languages have at least one labiodental fricative, while only 82 (18%) are reported to have bilabial fricatives, where similar turbulence is generated by forming a narrow constriction between both lips (though Blasi et al. (2019) suggest that the distribution of labiodental fricatives across languages may be governed in part by differences in human physiology). For approximants, 336 languages (75%) have rounded approximants while only 6 (1.33%) have labiodental approximants, and none are described as having a flat bilabial approximant.

Though not without exceptions, the nearly universal generalization across languages is clear: complete labial closures are typically produced using a flat constriction of the margins of both lips, narrow constrictions are produced using contact between the lower lip and upper teeth, and more open constrictions are produced using a rounded and protruded lip shape. This observation explains why it is possible for speech researchers to use the same phonetic writing systems to describe the speech sounds of so many of the world's different languages: a remarkably small set of characters (e.g., /p b m f v w/) can characterize almost all of the lip sounds in the world. It is by no means obvious why this should be the case. In principle, different degrees of constriction could just as well be produced by varying the activation of a single labial movement (the bilabial movement for the first sound in 'pile', for example). Although this is conceptually straightforward, it is unattested across human languages.

Quantality and robustness

Why should these particular structures be so commonly employed across languages? In general, we may expect that any structure constituted to achieve a task will be robust to noisy, everyday conditions (e.g., Loeb 2012). Such a structure should allow a large margin of error and optimize for feed-forward function (i.e., for operation without correction based on immediate sensory feedback; Perkell 2012, Guenther 2016). Some speech mechanisms have been described as having such properties, and have been associated with the term ‘quantal’. This term has been applied to a subset of non-linear effects in speech – traditionally those corresponding to some auditory-perceptual goal (e.g., Stevens 1972, Stevens 1989, Stevens and Keyser 2010). These nonlinearities correspond directly with error range, such that a quantal region in some function is a region in which large variation (error) in one dimension effects comparatively little response in some other (task) dimension (Fig. 1). Although this literature focuses on the auditory-perceptual domain, these effects have long been predicted to obtain across a variety of domains, including biomechanics (Fujimura 1989).

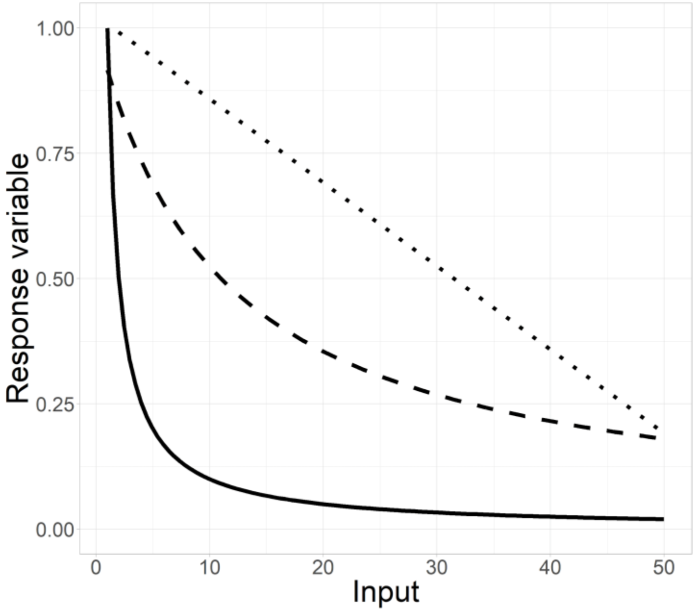

Figure 1. A schematic representation of functions that are strongly quantal (solid line), moderately quantal (dashed line), and not quantal (dotted line).

Relatively few biomechanical quantal effects have been discussed in the speech literature (e.g., Fujimura and Kakita 1979, Perkell et al. 2004, Perkell 2012). Such effects have, however, been shown in simulation studies of vocal tract structures such as the soft palate (Gick et al. 2014) and the larynx (Moisik and Gick 2017); both of these previous studies found that not all possible sets of muscle activations exhibit this quantality. Similar effects have been predicted for lip movements, specifically labiodental fricatives (Fujimura 1989), and have been observed to a limited degree in previous lip simulations: Nazari et al. (2011) show that the rounded, protruded lip posture for approximants is robust to variations in muscle stiffness.

In this paper, we investigate the quantal properties of labial movements using an anatomically and biomechanically realistic three-dimensional finite-element face model, implemented using the Artisynth modeling platform (e.g., Stavness et al. 2012; www.artisynth.org; Fig. 2). Artisynth simulates the biomechanics of soft- and hard-tissue structures, and can predict movements and shape changes of the vocal tract due to the actions of fixed groupings of muscles. In addition to passive tissue mechanics, the face model accounts for active muscle stress and intrinsic stiffness, volume preservation, and gravity.

The present paper has two primary goals. The first goal is to test whether the three canonical lip postures described above emerge under an assumption of modular control: that is, where the basic objects of control are groups of muscles that activate in fixed proportion to one another in response to a single control signal. The second goal is to test for quantal properties of these postures when they are implemented in this fashion.

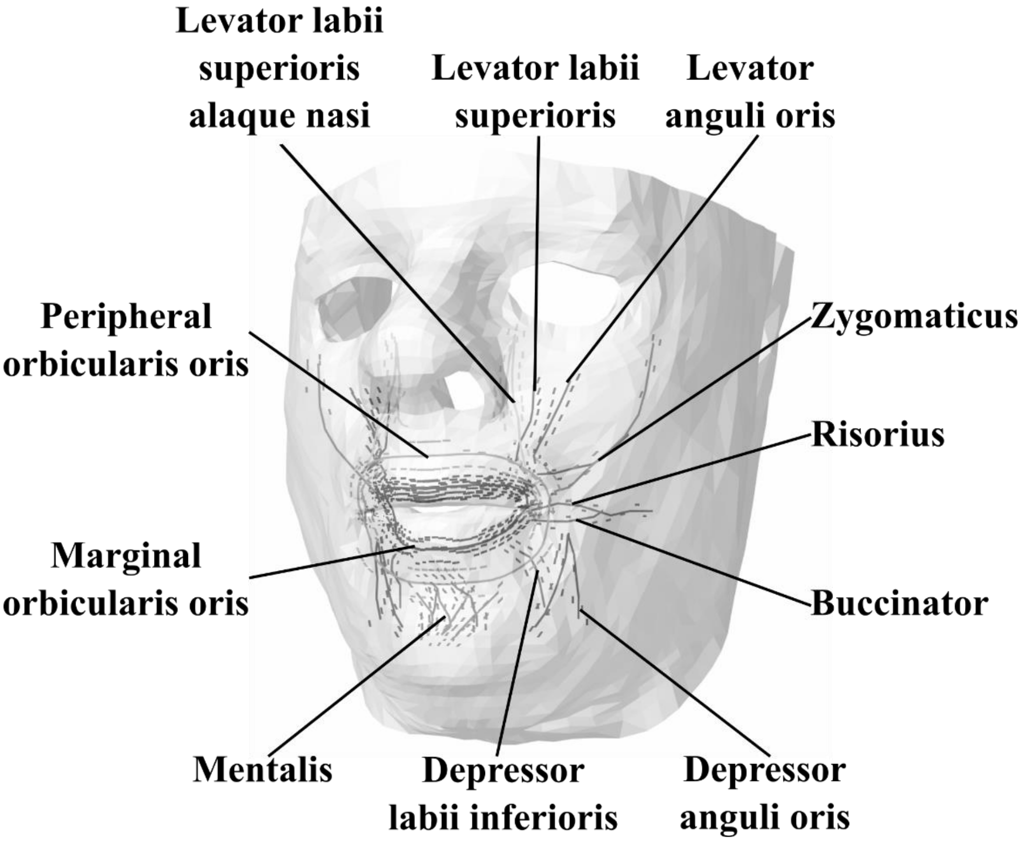

Figure 2. The Artisynth face model with muscle labels. Solid contours denote the geometric centerline of each muscle as defined in Nazari et al. (2010). Dashed contours represent the finite element associated with the muscle, with the direction of the dash showing the principle direction of muscle material within the element (i.e., the direction in which contractile stress is generated when the muscle is activated).

Under the hypothesis that speech movements are the outputs of functionally independent modules which are selected for use in speech in part because they take advantage of intrinsic quantal (non-linear) biomechanical properties that help to produce reliable outcomes, we predict that varying a single parameter in this model – the activation at fixed ratios of an appropriate set of muscles – will allow lips to achieve consistent speech postures over a wide range of activation levels, without reliance on feedback-based control (Simulation 1). In addition, we expect the outcomes will be robust to interference caused by surrounding muscle activations (Simulation 2). If borne out, these results will suggest a biomechanical contribution to the typological generalizations described above, and, more generally, that biomechanical robustness to intrinsic and extrinsic activation noise may be a contributing factor to the selection of motor modules.

Methods

To carry out the simulations presented in the present paper, we use a biomechanical model of the skull, jaw, lips, and face. This model was previously described in Stavness et al. (2013), and is based on existing models that have been employed in biomechanical and neurophysiological research. The model is implemented using the Artisynth biomechanical modeling toolkit. ArtiSynth is an open source, biomechanical simulation platform written in Java that combines multibody and finite-element (FE) methods to allow for accurate models of the rigid and deformable structures found in the human body (Lloyd et al. 2012). Multibody dynamics is used to model rigid structures in the body (Delp et al. 2012), while the FE method is used for movements of deformable structures (Lloyd et al. 2012, Maas et al. 2012). Artisynth models consist of *dynamic components* that have mass and maintain position and velocity, and *force effectors* and *constraints*, which govern interactions between dynamic components. Dynamic components may be particles, rigid bodies, or FEM nodes, and force effectors include various types of point-to-point muscles and both linear and non-linear finite elements. Constraints can be implemented as unilateral (e.g., contact detection and joint limits) and bilateral (e.g., joints or FEM incompressibility). Inputs to force effectors may be varied over time, and various output measurements recorded. This allows a wide range of mechanical systems to be modeled using Lagrangian representations.

The face model was created by registering and integrating existing models of the face and skull (Nazari et al. 2010) and the jaw-tongue-hyoid complex (Stavness et al. 2011) to computed tomography (CT) data from the same speaker. The face tissue model is implemented using a large deformation FE simulation framework, which is well suited to approximate the behavior of facial tissue (Chabanas et al. 2003). The outer and inner surfaces of the facial tissue were extracted from the CT data, and the volume between these surfaces was manually meshed using 6,342 eight-node hexahedral elements. These elements were arranged in three layers corresponding to the hypodermis, dermis, and epidermis. The elastic properties of the material are assumed to follow a hyperelastic law, and are implemented using a simplified fifth-order Mooney-Rivlin model. The parameterization used is based on previous models (Payan and Perrier 1997, Buchaillard et al. 2009), the latter of which calculated parameters based on indentation measures from a cadaver's tongue (Gerard et al. 2005).

The maxilla and mandible are implemented as rigid bodies based on the volumes calculated from the CT scan. A number of inner-surface nodes in the face mesh were attached to adjacent locations on the jaw and maxilla, while nodes in the cheeks and lips were left unattached. For the present simulations the jaw posture was held static. Contact between the rigid and deformable structures in the model is implemented using Artisynth's mesh-based collision detection and contact handling using dynamic constraints.

Muscle mechanics in the face are incorporated using a transverse-isotropic material (Wiess et al. 1996). This material allows the computation of muscle stress based on both the active force produced by muscle activation and the passive force produced by muscle stiffness. The passive stress in the fiber direction is assumed to increase exponentially with increasing fiber stretch. The parameters for the active contractile stress generated along the fiber direction are based on Blemker et al. (2005). The muscle model formulation incorporates stress stiffening effects, wherein the along-fiber stiffness increases with increasing contractile stress. The model uses muscle activations as input, which approximate the neural drive to muscle groups. A more detailed description of the model, particularly its mathematical properties and specific parameter values, can be found in Stavness et al. (2013). This work also includes a validation of the range of motion of bilabial postures generated by the model including lip protrusion and rounding.

An important property of this model is that it was not designed to perform a specific task, such as producing speech movements, but rather to be as accurate a model of the skull, jaw, face, and lips as is permitted by our knowledge of the morphological and mechanical properties of these structures and the state of current biomechanical modeling techniques, with a wide range of possible applications including feeding and swallowing, respiration, surgical modeling, and so on. Throughout this paper, we present our results as lip opening area as a function of the activation of presumed motor modules corresponding to various lip postures. The variations in lip opening at different activation levels and the rate at which the opening changes as a function of activation are determined solely by the biomechanical properties of the model and the selection of the muscles and their relative degrees of activation that contribute to each lip posture.

We provide additional methodological detail specific to each simulation below.

Simulation 1 Methodology: Robustness to varying activation

Simulation 1 tests whether the canonical lip postures described above can be produced under the assumption of modular organization, and whether these postures are robust to variations in the degree of activation of these modules.

We defined groupings of muscles corresponding to modules based on presumed muscle involvements in the three most widely attested speech lip postures (Lightoller 1925, Stavness et al. 2013). It is important to note that there is no ‘right’ set of muscles for a given posture, because there are likely to be a variety of solutions that will achieve the desired functional outcome (Loeb 2012; see also the results of Simulation 2 below for the approximant).

	OOPs	**OOPi**	**OOMs**	**OOMi**	**MENT**	**RIS**	**LLSAN**	**LLS**
Bilabial stop	–	–	30	30	20	20	–	–
Labiodental fricative	–	–	–	20	20	20	30	40
Approximant	40	40	–	–	–	–	–	–

Table 1: Maximum muscle stress (kPA) used for the three lip constrictions. Muscle abbreviations are superior (OOPs) and inferior (OOPi) peripheral orbicularis oris, superior (OOMs) and inferior (OOMi) marginal orbicularis oris, mentalis (MENT), and risorius (RIS), levator labii superioris alaeque nasi (LLSAN), and levator labii superioris (LLS). The bilabial closure, labiodental, and approximant constrictions correspond to the postures shown in Figs. 3b, 3c, and 3d respectively.

We assume a maximum specific muscle tension of 100 kPa following Stavness et al. (2013). This value is somewhat less than what is commonly assumed (200 kPa; Rospars and Meyer-Vernet 2016), but has been shown in previous work to result in realistic lip rounding and protrusion for 50% muscle activation (Stavness et al. 2013). The three postures were achieved by activating the involved muscles up to a maximum stress as indicated in Table 1. The stresses correspond to an average of roughly 30% of the maximum specific muscle tension, and were chosen by trial and error to generate representative postures. An appropriate ratio between muscle activations for each posture is important to achieve the correct outcome, but, as the results of these simulations will show, the outcomes are not particularly sensitive to the specific choice of maximum stress so long as it is sufficiently high. Some additional discussion of the model parameters is provided in the Discussion section.

To perform Simulation 1, we activated each muscle grouping from 0% to 100% of the maximum muscle stress in Table 1 in increments of 1%, and measured the resulting lip opening automatically using a green-screen technique, where a green plane was placed inside the model's mouth and the number of visible pixels of this color were counted from orthographic coronal images at each increment. These pixel counts were then converted to lip opening area in mm^2 based on the orthographic projection. The opening size for the stop and approximant was calculated as the area between the lips, while the area for the labiodental fricative was calculated as the area between the lower lip and the upper teeth. This resulted in a larger initial opening for the fricative than for the other sounds.

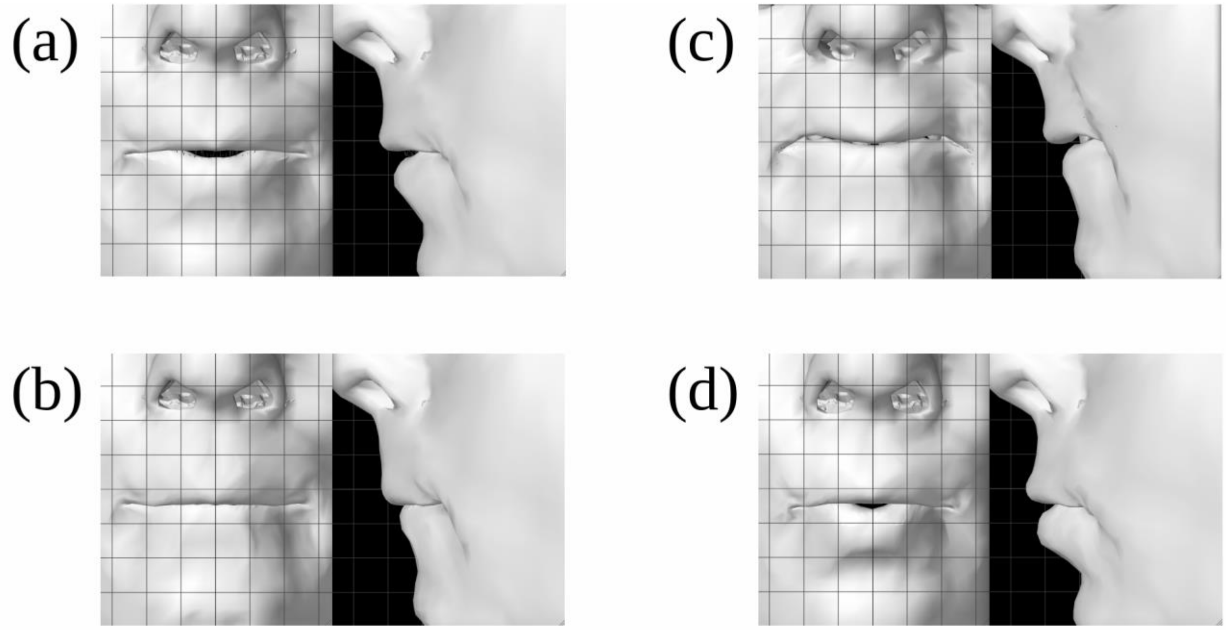

Figure 3. Model postures for (a) rest position; (b) stop; (c) fricative; and (d) approximant.

Simulation 2 Methodology: Robustness to surrounding muscles

We expect that postures observed to be robust to variations in degree of activation in Simulation 1 will also be robust to interference from activation of surrounding muscles. To test this, we performed two additional types of simulations testing whether the rounding posture used in approximants is robust to interference from surrounding muscle activations.

Lip rounding is generated by constricting the peripheral orbicularis oris (OOP). Orbicularis oris is a complex muscle whose fibers encircle the mouth and make up much of the substance of the lips. It can be broadly divided into two parts: the marginal orbicularis oris (OOM), which is closest to the lip opening, and OOP, which is farther from the lip opening (see Fig. 1). OOM and OOP can be further divided into superior (OOMs; OOPs) and inferior (OOMi; OOPi) portions, which run above and below the lip opening respectively and can function independently. Contracting OOM serves to close the lips, while contracting OOP protrudes them (see Table 1).

We chose the approximant for this simulation because it does not result in medial contact between the lips or the lips and teeth. In the stop and fricative, motion perpendicular to the contact surface is directly obstructed, and accordingly we may expect these movements to be more robust to interference from other muscles once contact is made. In the approximant, however, there is no such obstruction, and thus the effects of perturbations from surrounding muscles on lip opening size can be measured across a wide range of OOP activation. Simulation 2 included two types of simulation: the first tested whether lip closure produced by OOP activation remained stable in the presence of surrounding muscle noise, while the second tested the effect of degree of OOP activation on this stability. Both used the same green-screen technique employed in Simulation 1 to measure lip opening area.

We used Artisynth's BatchSim tool to run the simulations, permitting automatic probabilistic sampling of muscle activations. Each simulation consisted of 500 samples, but a small number of samples resulted in invalid inputs or invalid model states, where the finite elements encounter numerical problems, such as element inversion or area-measurement inconsistencies, producing invalid results. These samples are not reported. In the first simulation type, we sampled the activation space of OOP with activation of other muscles (all those listed in the previous section plus depressor anguli oris, buccinator, depressor labii inferioris, levator anguli oris, and zygomaticus) excited randomly from a uniform distribution between 0% and 5% activation in the same range as the OOP, which corresponds to a maximum stress of about 2 kPa (473 samples). Uniform sampling was done to avoid bias in the choice of inputs, and the activation range of 0% to 5% across the set of other muscles was chosen as a reasonable assumption of motor noise from small labial movements that might co-occur with speech without completely disrupting it.

The second set of simulations sampled OOP activation from a normal distribution with a mean of 10% and a standard deviation of 5% (low activation; 486 samples), and from a normal distribution with a mean of 80% and a standard deviation of 5% (high activation; 487 samples). Both sets of samples were in the presence of noise from other muscles, again sampled uniformly from 0% to 5% activation. Samples that fell outside the valid range of OOP muscle activations (0-100%) were discarded and resampled.

Results

Simulation 1 results

As predicted, the plot in Fig. 4 shows a nonlinear relationship between module activation and opening size for all three labial postures. Quantal regions (relatively ‘flatter’ parts of the plots) are enclosed in gray boxes in Fig. 4, indicating the region of the graph in which at least 95% of the total distance from maximum opening to maximum constriction has been covered.

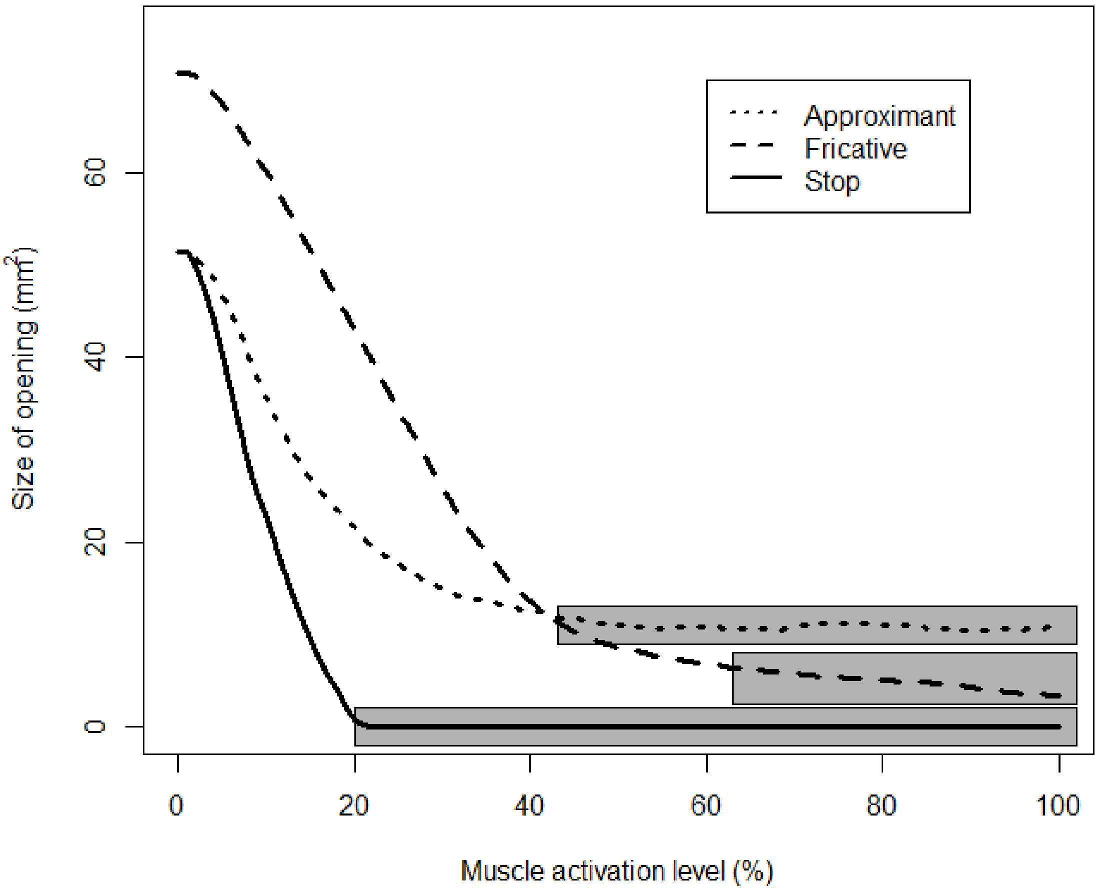

Figure 4. Lip opening area as a function of muscle activation for the three canonical postures. Gray boxes indicate quantal regions where 95% of the distance from maximum opening to maximum closure has been covered.

The degree of quantality of a function (here from activation to size of lip opening) can be quantified using the *Q-score* proposed by Moisik and Gick (2017). The Q-score provides a quantitative measure of quantality by comparing the first derivative of a response variable (here the proportion of maximum lip opening) at earlier and later points in the series (here different degrees of muscle activation). It is calculated as follows:

$$\text{Q-score}(f_{norm}(x)) = \ln\left(\frac{\sum_{i=1}^{n} w_i |f_{norm}'(x_i)|}{\sum_{i=1}^{n} w_i}\right) - \ln\left(\frac{\sum_{i=1}^{n} z_i |f_{norm}'(x_i)|}{\sum_{i=1}^{n} z_i}\right)$$

where x is a vector of length n containing input values, f_{norm} is a function that maps values of x to a normalized response variable between 0 and 1, $z_i = (i - 1) / (n - 1)$, and $w_i = 1 - z_i$. Higher Q-scores indicate that the function shows a region of significant change at lower input values, but stabilizes into a region of low change at higher input values. The Q-scores for the stop, fricative, and approximants are 2.18, 1.01, and 1.37 respectively. Using the heuristics in Moisik and Gick (2017), the stop output is *strongly quantal*, while the fricative and approximant outputs are *moderately quantal*. These results indicate that a large range of possible muscle activation levels in a feed-forward model can produce consistent equivalent postures using fixed sets of muscles.

Simulation 2 results

The results of the first simulation type, where OOP was activated uniformly between 0 and 100% activation with other muscles activated uniformly between 0 and 5% activation, are shown in Fig. 5. We see that, even in the presence of noise from surrounding muscle activations, the lip closure area remains fairly stable, particularly at higher OOP activation levels. Although the overall constriction size is lower, only about 5% of activations resulted in complete closure. This overall decrease is opening size is not surprising, since the majority of the perturbing muscles serve to close the lips.

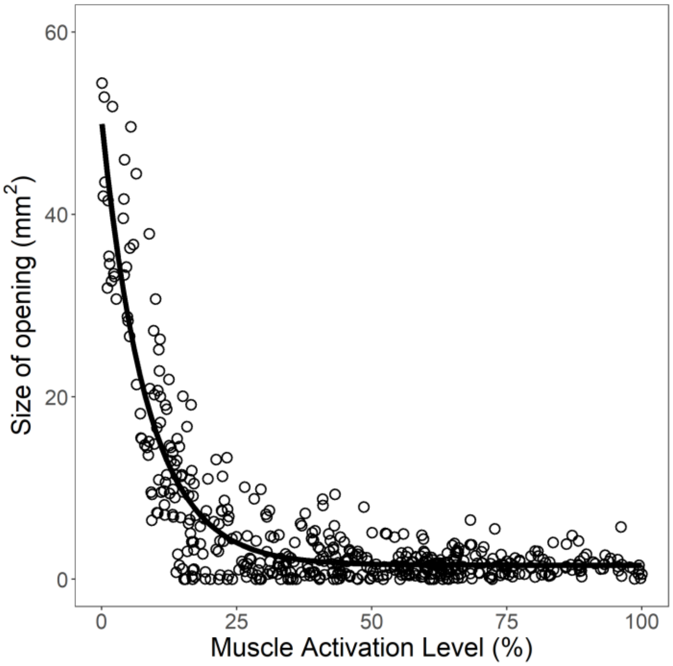

Figure 5. The results of the probabilistic sampling simulations on the full range of OOP activation with noise from surrounding muscles, overlaid with an asymptotic regression line of the form *a + b(1 - exp(-exp(c)x))*.

These results from the second simulation with OOP activation normally distributed at 10% activation and 80% activation with a standard deviation of 5% are shown in Fig. 6. The correlation between OOP activation and lip opening area in the low activation case (Fig. 6, left panel) was -0.75, and the standard deviation of lip area was 12.4 mm^2. In the high activation case (Fig. 6, right panel), the correlation was 0.04, and the standard deviation of lip area was 1.5 mm^2. Fisher's transformation shows that these correlations are significantly different [Z = -15.65, p < 2.2e-16]. OOP muscle activation accounts for about 56% of lip area variation at low activation, but only 0.2% at high activation. Welch's two-sample *t*-test shows that the opening sizes are significantly different between the two conditions [t(502.33) = 27.207, p < 2.2e-16]. Few simulations generated complete closure (6 in the low activation case, and 3 in the high activation case).

To determine whether the variance caused by surrounding muscle activation is different for the low and high activation cases, we applied Levene's test to the residuals from fitting an asymptotic regression model of the form *a + b(1 - exp(-exp(c)x))* to the low activation and a linear regression model to the high activation data with OOP activation as the predictor and opening size as the response variable (Fig. 6). Testing on residuals is necessary because OOP activation contributes more to variation in opening size at lower activation levels. Levene's test shows that the variance due to surrounding muscle activations is significantly lower in the higher activation condition [F(1, 971) = 637.72, p < 2.2e-16].

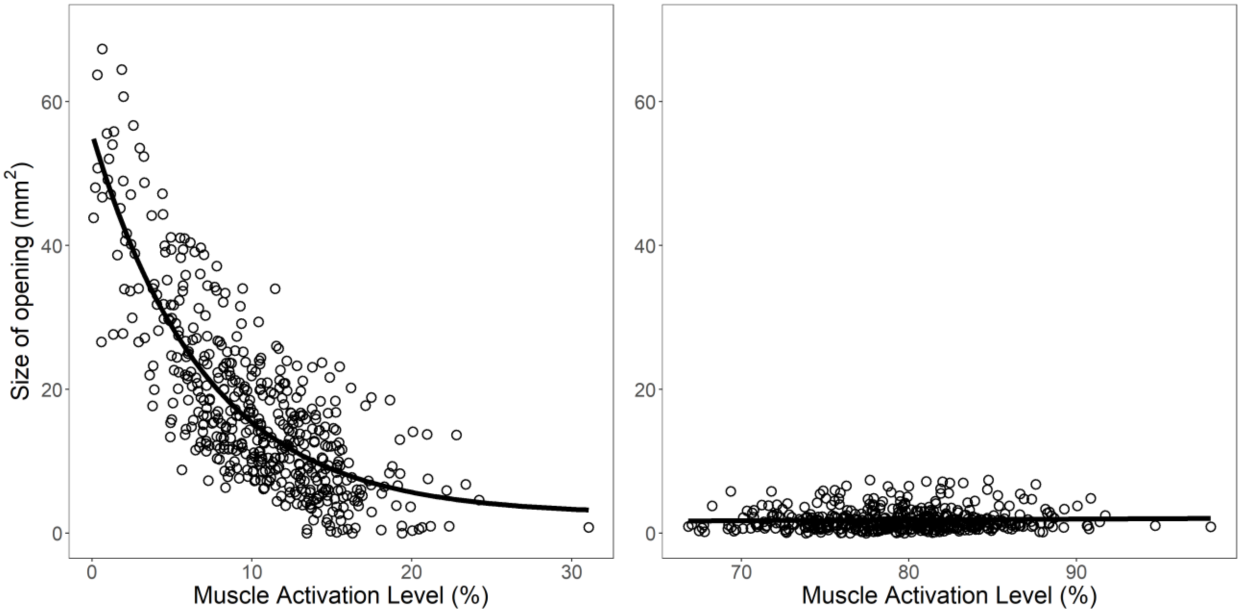

Figure 6. Results on selected ranges of OOP activation. Left: mean of 10% activation overlaid with an asymptotic regression line of the form *a + b(1 - exp(-exp(c)x))*. Right: mean of 80% activation overlaid with a linear regression line.

These results indicate that activation of OOP maintains a relatively consistent lip posture, even in the presence of modest perturbation from the activation of surrounding muscles, and that increasing OOP activation results in reduced interference from surrounding muscles, due to increased stiffness at higher activation levels. The area of higher activation tested in these simulations corresponds roughly to the quantal region shown in Fig. 4, indicating that the quantal regions of the postures produced using OOP are robust to both intrinsic perturbation (variations in OOP activation) and extrinsic perturbation (variations in activation of surrounding muscles).

Discussion

The simulations presented in this paper lend support to the hypothesis that modular control is a viable strategy for generating the three most widely attested lip postures for speech. Our simulations assume that each lip posture is generated using a structure that follows the motor theoretic definition of a module: a set of muscles that activate in fixed proportion to one another in response to a single control signal that may vary in amplitude. The fact that this assumption results in outcomes that look very much like the lip postures used in speech provides support for modularization as a viable strategy for motor control.

These results also support the hypothesis that emergent speech motor modules corresponding to these postures are characterized by quantal biomechanical properties. That is, motor learners do not converge on particular speech motor modules based solely on their ability to generate a particular outcome, but also based on their ability to do so robustly. Modules with quantal biomechanical properties leverage the non-linear relationship between neural activations and movement afforded by biomechanics to provide this robustness in the face of both intrinsic and extrinsic perturbation.

Each of the modules we have tested displays robustness to intrinsic variations in activation levels, and the labial approximant is also robust to activations from surrounding muscles. As to why these particular structures might exhibit this robustness, the bilabial case is straightforward: once contact is achieved between the upper lip and lower lip, additional activation will serve only to close the lips more tightly, resulting in a large range of activations that achieve the goal of bilabial closure. In the case of the labiodental, the retraction of the upper lip and the uneven surface provided by the upper teeth conspire to make complete closure difficult to achieve, reliably resulting in a small opening (Fujimura 1989, Blasi et al. 2019). In the case of the approximant, the morphology of OOP is such that complete closure cannot be achieved by increasing activation up to the threshold we tested. Stavness et al. (2013) suggest that the highly peripheral and moderately deep location of OOP is important to effectively produce protrusion and rounding: more superficial implementations tend to produce complete closure. In addition, Nazari et al. (2010) suggest that the passive stiffening of OOP with increasing activation is crucial for achieving the desired lip shape. We leave a more detailed examination for future research.

We may speculate on why the hypothetical labial inventory described above, wherein each type of closure is generated by activating a single module to different degrees, is unattested. If robustness is a criterion for the development of modular structure, then such a system is undesirable because the regions in which frication and approximation are achievable using this muscle configuration are biomechanically unstable, providing highly variable outcomes (see Fig. 4). Indeed, we see a general correspondence between the degree of quantality for each constriction type and its typological prevalence.

Based on the results presented here, and on previous work that has demonstrated quantal properties for movements made in the oropharyngeal isthmus (Gick et al. 2014) and the larynx

(Moisik and Gick 2017), we predict that any speech sound that appears with any cross-language regularity will exhibit this property of robustness to perturbation, and that robustness may exert selectional pressures on motor modules in general. Many other factors doubtless contribute to the emergence of motor modules, and the relationship between modularity and robustness is not perfect. In this case, this claim does not preclude the existence of sounds with weaker quantal properties, but suggests that these sounds should be less common due in part to pressures on language change over time that favor sounds with similar perceptual outcomes that are more biomechanically stable (e.g., Purnomo et al. 2019). In this paper we present evidence that several typologically common sounds exhibit quantal properties. What is needed to further support this prediction is to show that typologically uncommon sounds have weaker quantal properties.

The prediction that typological prevalence depends in part on biomechanical stability suggests that the bilabial fricative, which is observed with some regularity cross-linguistically, is produced using a different mechanism than bilabial closure. The line in Fig. 4 that corresponds to bilabial closure has an extremely small range of activation levels in which constriction sufficient for frication is possible. Small deviations in either direction will result in approximation or complete closure. One possible mechanism that could be used in bilabial fricatives is lip compression (e.g., Okada 1991), where the corners of the mouth are drawn together without simultaneous protrusion. This mechanism may resist complete closure of the lips (perhaps via increasing stiffness), and provide a larger stable region in which frication is achievable. Additional experimental and modeling work is necessary to determine whether this is the case.

The consistent use of these specialized postures for different speech sounds suggests that the primitive units of speech motor organization are best modeled as modular muscle groupings that activate in fixed proportion to one another to perform a particular task (see, e.g., Ting et al.

2015) – in this case, to achieve a particular degree of labial constriction. The second set of simulations also demonstrates spatial robustness in these groupings, i.e., that there are many different sets of muscles that can be selected to produce similar postures. This has been demonstrated on a simplified lip model (de Vries et al. 2018), and is shown here to hold for a higher-dimensional model as well.

Because of the difficulty in collecting physiological measurements from muscles used in speech, muscle groupings and their relative activation levels used in the present simulations were chosen based on their representative postural outcomes. Validating these against experimental data will be an important future step, though we expect that the results presented here should be robust across a range of muscle groupings and activation levels.

Similarly, the choice of 5% as the maximum activation level for the surrounding muscles in Simulation 2 is somewhat arbitrary. To our knowledge there are no quantitative results describing the forces generated by muscle contractions in behaviors that typically co-occur with speech. We expect that higher levels of activation that constitute volitional activation will disrupt the approximant to a greater degree. This will depend on a number of factors, such as the stress stiffening of OOP (Nazari et al. 2010) and the relative timing and magnitude of force generated by surrounding muscles: for example, co-contracting LLS and DLI would exert opposing forces on OOP, and might therefore be less likely to disrupt the approximant. Understanding how speech modules respond to particular types of extrinsic perturbation, rather than perturbation generated at random, will be an important task for future research.

A sensitivity analysis of this model will be a useful component in future work to provide a clearer understanding of how robust these findings are with respect to the model's parameterization. Due to the computational complexity of FEM models and the large number of parameters, such an analysis warrants a study of its own. A further complication with such an analysis concerns the interdependence of model parameters. For example, in the model of passive tissue mechanics we employ (Blemker et al. 2005), muscle stiffness and muscle activation are coupled, with stiffness increasing exponentially as a function of activation (cf. Nazari et al. (2010), whose model of passive tissue mechanics allows stiffness and activation to be decoupled).

Finally, this model omits several mechanisms which are known to be important for motor control, such as reflex mechanisms and sensory afferents, which provide position- and length-based contributions to muscle force. Their omission here is a simplifying assumption, and it will be important to incorporate them into future work as modeling techniques advance.

We see this type of computational modeling as complementary, rather than supplementary, to experimental work (e.g., Hannam 2011). That is, biomechanical models can help to identify plausible or implausible muscle synergies for performing various tasks, and these predictions can be validated against experimental work in neurophysiology, kinematics, etc. In turn, this experimental work can guide subsequent modeling studies that probe for new effects. Thus, the present study should be seen as a milestone in a larger research program that explores the assumptions and hypotheses of modular control. We are optimistic that further experimental work testing the extent and validity of quantal biomechanical effects in speech production will provide additional insights into the organization of speech movements.

Glossary

stop: a sound produced with complete closure of the oral cavity; **fricative**: a sound produced with a constriction in the oral cavity narrow enough to produce turbulent noise; **approximant**: a sound produced with a constriction in the oral cavity that is not sufficiently narrow to produce turbulent noise; **bilabial**: a sound produced with a constriction between the two lips; **labiodental**: a sound produced with a constriction between the lower lip and upper teeth; **rounded sound**: a sound with constriction produced by rounding and protruding the lips; **OOPs**: superior peripheral orbicularis oris; **OOPi**: inferior peripheral orbicularis oris; **OOMs**: superior marginal orbicularis oris; **OOMi**: inferior marginal orbicularis oris; **MENT**: mentalis; **RIS**: risorius; **LLSAN**: levator labii superioris alaeque nasi; **LLS**: levator labii superioris.

Grants

This work was funded by NIH Grant DC-002717, NSERC Grants RGPIN-2015-05099 and RGPIN-2019-07017, and by a SSHRC Doctoral Award to CM.

Disclosures

The authors declare that they have no competing interests.

Availability of data and materials

The code necessary for running the simulations can be found at https://github.com/artisynth. The *artisynth_core* repository contains the base Artisynth platform, while *artisynth_models* contains a variety of models, including the one used in this paper, which can be found here: https://github.com/artisynth/artisynth_models/blob/master/src/artisynth/models/face/BadinFaceDemoLipOpening.java. Instructions for installing and running Artisynth can be found here: https://www.artisynth.org/Documentation/InstallGuide. Instructions for using the BatchSim tool, which allows the combinatorial and probabilistic sampling used in these simulations can be found here: https://github.com/artisynth/artisynth_models/blob/master/doc/batchsim/batchsim.pdf. We also welcome correspondence from readers who are interested in learning more about Artisynth.

Acknowledgements

The authors acknowledge Yohan Payan for his assistance in creating the face-jaw model, Pascal Perrier and Mohammad Ali Nazari for providing the face model geometry, and Jonathan de Vries for helpful commentary on early drafts of the manuscript.

Authors' contributions

BG and CM are joint first authors. BG conceived of the study, coordinated the research, designed simulations, and drafted early versions of the manuscript. CM carried out simulations and analyzed simulation results, performed typological research, and drafted and revised later versions of the manuscript. CC, EW, and FRD designed and carried out simulations. SF coordinated the simulations, and IS designed, carried out, and coordinated simulations. The authors have read and approved the final manuscript.

References

Anderson P, Fels S, Stavness I, Pearson W, Gick B. Intravelar and extravelar portions of soft palate muscles in velic constrictions: a three-dimensional modeling study. *J Speech Lang Hear R* 64(4): 802-814, 2019.

Benguerel A-P, Hirose H, Sawashima M, Ushijima T. Velar coarticulation in French: an electromyographic study. *J Phonetics* 5: 159-167, 1977.

Bernstein N. *The co-ordination and regulation of movements*. Oxford: Pergamon Press, 1967.

Berniker M, Jarc A, Bizzi E, Tresch MC. Simplified and effective motor control based on muscle synergies to exploit musculoskeletal dynamics. *P Natl Acad Sci USA* 106: 7601–7606, 2009.

Bizzi E, Cheung VCK, d'Avella A, Saltiel P, Tresch M. Combining modules for movement. *Brain Res Rev* 57: 125-133, 2008.

Bizzi E, Mussa-Ivaldi FA, Giszter S. Computations underlying the execution of movement: A biological perspective. *Science* 253(5017): 287-291, 1991.

Blasi DE, Moran S, Moisik SR, Widmer P, Dediu D, Bickel B. Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* 363(6432): 2019.

Blemker S, Pinsky PM, Delp SL. A 3D model of muscle reveals the causes of nonuniform stress in the biceps brachii. *J Biomech* 38: 657-665, 2005.

Buchaillard S, Perrier P, Payan Y. A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *J Acoust Soc Am* 126(4): 2033-2051, 2009.

Chabanas M, Luboz V, Payan Y. Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery. *Med Image Anal* 7: 131-151, 2003.

Cheung VCK, Piron L, Agostini M, Silvoni S, Turolla A, Bizzi E. Stability of muscle synergies for voluntary actions after cortical stroke in humans. *Proc Natl Acad Sci U.S.A.* 106: 19563-19568, 2009.

Clark DJ, Ting LH, Zajac FE, Neptune RR, Kautz SA. Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke. *J Neurophysiol* 103: 844-857, 2010.

Chiu C. Startling auditory stimulus as a window into speech motor planning. Unpublished doctoral dissertation, University of British Columbia, 2016.

Chiu C, Gick B. Decoupling functional units in speech production using auditory startle. In: *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, edited by The Scottish Consortium for ICPhS 2015. Glasgow, UK: University of Glasgow. Paper Number 966, 2015.

Cholin J, Dell GS, Levelt WJM. Planning and articulation in incremental word production: Syllable-frequency effects in English. *J Exp Psychol Learn* 37(1): 109-122, 2011.

Cholin J, Levelt WJM, Schiller NO. Effects of syllable frequency in speech production. *Cognition* 99(2): 205-235, 2006.

Crompton A. Syllables and segments in speech production. *Linguistics* 19(7-8): 663-716, 1981.

d'Avella A, Bizzi E. Shared and specific muscle synergies in natural motor behaviors. *P Natl Acad Sci USA 102*: 3076–3081, 2005.

d'Avella A. Modularity for motor control and motor learning. *Adv Exp Med Biol* 957: 3-19, 2016.

d'Avella A, Saltiel P, Bizzi E. Combinations of muscle synergies in the construction of a natural motor behavior. *Nat Neurosci* 6(3): 300-308, 2003.

de Rugy A, Loeb GE, Carroll TJ. Muscle Coordination is Habitual Rather Than Optimal. *J Neurosci* 31(21): 7384-7391, 2012.

de Rugy A, Loeb GE, Carroll TJ. Are muscle synergies useful for neural control? *Front Comput Neurosci 21*: 2013.

de Vries J, Stavness I, Fels S, Gick B. Speech-like movements emerge from simulated perioral muscle activation space without neural control. *Can Acoust* 46(4): 62-63, 2018.

Delattre P, Freeman D. A dialect study of American R’s by x-ray motion picture. *Linguistics* 44: 29–68, 1968.

Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, Guendelman E, Thelen DG. OpenSim: open-source software to create and analyze dynamic simulations of movement. *IEEE T Bio-Med Eng* 54(11): 1940-1950, 2007.

Derrick D, Gick B. Individual variation in English flaps and taps: A case of categorical phonetics. *Can J Ling* 56(3): 307-319, 2011.

Diedrichsen J, Classen J. Stimulating news about modular motor control. *Neuron* 76(6): 1043-1045, 2012.

Diedrichsen J, Shadmehr R, Ivry RB. The coordination of movement: Optimal feedback control and beyond. *Trends Cogn Sci* 14(1): 31-39, 2010.

Fowler CA, Rubin PE, Remez RE, Turvey MT. Implications for speech production of a general theory of action. In: *Language Production*, edited by Butterworth G. New York: Academic Press, 1980, 373-420.

Fujimura O, Kakita K. Remarks on quantitative descriptions of the lingual articulation. In: *Frontiers of Speech Communication Research*, edited by Lindblom B, Öhman S. New York: Academic Press, 17-24, 1979.

Fujimura O. Comments on ‘On the Quantal Nature of Speech,’ by K.N. Stevens. *J Phonetics* 17: 87-90, 1989.

Gentner R, Classen J. Modular organization of finger movements by the human central nervous system. *Neuron* 52(4): 731-742, 2006.

Gerard JM, Ohayon J, Luboz V, Perrier P, Payan Y. Non-linear elastic properties of the lingual and facial tissues assessed by indentation technique, Application to the biomechanics of speech production. *Med Eng Phys* 27: 884-892, 2005.

Gick B, Stavness I. Modularizing Speech. *Front Psychol* 4: 977, 2013.

Gick B, Allen B, Roewer-Despres F, Stavness I. Speaking tongues are actively braced. *J Speech Lang Hear R* 60(3): 494-506, 2017.

Gick B, Anderson P, Chen H, Chiu C, Kwon HB, Stavness I, Tsou L, Fels S. Speech function of the oropharyngeal isthmus: A modeling study. *Comput Methods Biomech Biomed Eng: Imaging & Visualization* 2(4): 217-222, 2014.

Gick B, Schellenberg M, Stavness I, Taylor R. Articulatory phonetics. In: *The Routledge Handbook of Phonetics*, edited by Katz WF, Assmann P. Ch 5. New York: Taylor & Francis, 2019a.

Gick B, Chiu C, Widing E, Roewer-Depres F, Mayer C, Fels S, Stavness I. Quantal biomechanical effects in speech postures of the lips. In: *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, edited by Calhoun S, Escudero P, Tabain M, Warren P. Canberra, Australia: Australian Speech Science and Technology Association Inc.: 1749-1753, 2019b.

Guenther FH. *Neural control of speech.* Cambridge, MA: MIT Press, 2016.

Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang* 96(3): 280-301, 2006.

Hannam A. Current computational modelling trends in craniomandibular biomechanics and their clinical implications. *J Oral Rehab* 38: 217-234, 2011.

Hug F. Can muscle coordination be precisely studied by surface electromyography? *J Electromyogr Kinesiol* 21: 1-12, 2011.

Kozhevnikov VA, Chistovich LA. *Speech: Articulation and perception*. Oxford: Nauka.

Kraus BS, Kitamura H, Latham RA. *Atlas of Developmental Anatomy of the Face*. New York: Harper & Row.

Kuehn D, Folkins J, Cutting C. Relationships between muscle activity and velar position. *Cleft Palate J* 19(1): 25-35, 1982.

Kutch JJ, Valero-Cuevas FJ. Challenges and new approaches to proving the existence of muscle synergies of neural origin. *PLoS Comput Biol* 8: e1002434, 2012.

Larrabee WF Jr, Makielski KH. *Surgical Anatomy of the Face*. New York: Raven Press.

Levelt WJM, Wheeldon L. Do speakers have access to a mental syllabary? *Cognition* 50(1-3): 239-269, 1994.

Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behav Brain Sci* 22(1): 1-38, 1999.

Lightoller GHS. Facial muscles: The modiolus and muscles surrounding the rima oris with some remarks about the panniculus adiposus. *J Anat* 60: 1-85, 1925.

Lloyd JE, Stavness I, Fels S. Artisynth: A fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In: *Soft tissue biomechanical modeling for computer assisted surgery*, edited by Payan, Y. Berlin: Springer: 355-394, 2012.

Loeb EP, Giszter SF, Saltiel P, Bizzi E, Mussa-Ivaldi FA. Output units of motor behavior: An experimental and modeling study. *J Cogn Neurosci* 12(1): 78-97, 2000.

Loeb GE. Optimal isn't good enough. *Biol Cybern* 106: 757-765, 2012.

Maas SA, Ellis BJ, Ateshian GA, Weiss JA. FEBio: finite elements for biomechanics. *J Biomed Eng* 134(1): 011005, 2012.

MacNeilage PF, Davis B. Acquisition of speech production: Frames, then content. In: *Attention and performance 13: Motor representation and control*, edited by Jeannerod M. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc: 453-476, 1990.

Maddieson I. *Pattern of Sounds*. Cambridge, UK: Cambridge University Press, 1984.

Maddieson I, Precoda K. Updating UPSID. *UCLA Working Papers in Phonetics* 74: 104-114, 1990.

Mayer C, Roewer-Despres F, Stavness I, Gick B. Do innate stereotypies serve as a basis for swallowing and learned speech movements? *Behav Brain Sci* 40: 2017.

Moisik S, Gick B. The quantal larynx: The stable regions of biomechanics and implications for speech production. *J Speech Lang Hear R* 60: 540-560, 2017.

Mosik S, Gick B, Esling JH. The quantal larynx in action: Smooth and abrupt aspects of laryngeal motion observed in laryngoscopic videos. *Proceedings of the International Seminar on Speech Production (ISSP 2017)*. Tianjin, China: 4pp, 2017.

Nazari MA, Perrier P, Chabanas M, Payan, Y. Simulation of dynamic orofacial movements using a constitutive law varying with muscle activation. *Comput Methods Biomec* 13: 469-482, 2010.

Nazari MA, Perrier P, Chabanas M, Payan Y. Shaping by Stiffening: A Modeling Study for Lips. *Motor Control* 15(1): 141-168, 2011.

Okada H. Japanese. *J Int Phon Assoc* 21(2): 94-96, 1991.

Overduin SA, d’Avella A, Carmena JM, Bizzi E. Microstimulation activates a handful of muscle synergies. *Neuron* 76(6): 1071-1077, 2012.

Payan Y, Perrier P. Synthesis of V-V sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis. *Speech Commun* 22: 185-205, 1997.

Perkell JS. Movement goals and feedback and feedforward control mechanisms in speech production. *J Neurolinguist* 25(5): 382-407, 2012.

Perkell JS, Matthies ML, Tiede M, Lane H, Zandipour M, Marrone N, Stockmann E, Guenther FH. The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *J Speech Lang Hear R* 47(6): 1259-1269, 2004.

Pittman LJ, Bailey EF. Genioglossus and intrinsic electromyographic activities in impeded and unimpeded protrusion tasks. *J Neurophysiol* 101: 276-282, 2009.

Purnomo G, Mellesmoen G, Shamei A, Gick B. Production study of Spanish spirantization in naturalistic speech. *Can Acoust* 47(3): 90-91, 2019.

Radhakrishnan SM, Baker SN, Jackson A. Learning a novel myoelectric-controlled interface task. *J Neurophysiol* 100(4): 2397-2408, 2008.

Ramanayaranan V, Goldstein L, Narayanan SS. Spatio-temporal articulatory movement primitives during speech production: Extraction, interpretation, and validation. *J Acoust Soc Am* 134(2): 1378-1394, 2013.

Rogers CR, Mooney MP, Smith TD, Weinberg SM, Waller BM, Parr LA, Docherty BA, Bonar CJ, Reinholt LE, Deleyiannis FW, Siegel MI, Marazita ML, Burrows AM. Comparative microanatomy of the orbicularis oris muscle between chimpanzees and humans: Evolutionary divergence of lip function. *J Anat* 214(1): 36-44, 2008.

Rospars JP, Meyer-Vernet N. Force per cross-sectional area from molecules to muscles: a general property of biological motors. *R Soc Open Sci* 3(7): 160313, 2016.

Santello M, Flanders, M, Soechting JF. Postural hand synergies for tool use. *J Neurosci* 18(23): 10105-10115, 1998.

Sartori M, Farina D, Lloyd, DG. Hybrid neuromusculoskeletal modeling to best track joint moments using a balance between muscle excitations derived from electromyograms and optimization. *J Biomech* 47: 3613-3621, 2014.

Standring S. *Gray's Anatomy: The Anatomical Basis of Clinical Practice*. London: Elsevier Health Sciences, 2015.

Stavness I, Lloyd JE, Fels SS. Automatic Prediction of Tongue Muscle Activations Using a Finite Element Model. *J Biomech* 45(16): 2841-2848, 2012.

Stavness I, Lloyd J, Payan Y, Fels S. Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. *Int J Numer Meth Bio*: 367-390, 2011.

Stavness I, Nazari MA, Perrier P, Demolin D, Payan Y. A biomechanical modeling study of the effects of the orbicularis oris muscle and jaw posture on lip shape. *J Speech Lang Hear R* 56: 878-890, 2013.

Stevens KN. The Quantal Nature of Speech:Evidence from Articulatory-Acoustic Data. In: *Human Communication: A Unified View*, edited by David EE Jr, Denes PB. New York: McGraw-Hill: 51-66, 1972.

Stevens KN. On the quantal nature of speech. *J Phonetics* 17, 3-45, 1989.

Stevens K, Keyser SJ. Quantal theory, enhancement and overlap. *J Phonetics* 38: 10-19, 2010.

Ting LH, Chiel HJ. Muscle, biomechanics, and implications neural control. In: *Neurobiology of Motor Control: Fundamental Concepts and New Directions*, edited by Hooper SL, Büschges, A. New York: Wiley, 2017.

Ting LH, Chiel HJ, Trumbower RD, Allen JL, McKay JL, Hackney ME, Kesar TM. Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron* 86: 38-54, 2015.

Ting LH, McKay JL. Neuromechanics of muscle synergies for posture and movement. Curr Opin *Neurobiol* 17: 622-628, 2007.

Tresch MC, Saltiel P, d'Avella A, Bizzi E. Coordination and localization in spinal motor systems. *Brain Res Rev* 40: 66-79, 2002.

The UCLA Phonetics Laboratory. Dissection of the Speech Production Mechanism. In: *UCLA Working Papers in Phonetics 102*, edited by Epstein M, Hacopian N, Ladefoged P. Los Angeles: UCLA Department of Linguistics, 2002.

Valls-Solé J, Rothwell JC, Goulart F, Cossu G, Muñoz E. Patterned ballistic movements triggered by startle in healthy humans. *J Physiol* 516(3): 931-938, 1999.

Weiss JA, Maker BN, Govindjee S. Finite element implementation of incompressible, transversely isotropic hyperelasticity. *Comput Method Appl M* 135(1-2): 107-128, 1996.