Quantal biomechanical effects in speech postures of the lips

Bryan Gick [1, 2], Chenhao Chiu [3], Erik Widing [4], Francois Roewer-Despres [4],
Connor Mayer [5], Sidney Fels [6], Ian Stavness [4]

[1] Department of Linguistics, University of British Columbia
[2] Haskins Laboratories, New Haven, CT
[3] Graduate Institute of Linguistics, National Taiwan University
[4] Department of Computer Science, University of Saskatchewan
[5] Department of Linguistics, University of California, Los Angeles
[6] Department of Electrical and Computer Engineering, University of British Columbia

gick@mail.ubc.ca, chenhaochiu@ntu.edu.tw, ehw785@mail.usask.ca, francois.roewerdespres@usask.ca,
connormayer@ucla.edu, ssfels@ece.ubc.ca, ian.stavness@usask.ca

ABSTRACT

A robust typological finding is that languages produce different degrees of labial constriction using distinct muscle groupings and concomitantly distinct lip postures. Past research has suggested that these lip postures exhibit quantal biomechanical properties that allow movement goals to be realized despite variable muscle activation. We perform two sets of computer simulations showing that these postures are biomechanically robust, first to variation in the activation levels of participating muscles, and second to interference from surrounding muscles. These results provide quantitative support for the hypothesis that quantal biomechanical properties are an important factor in selecting the groupings of muscles used for speech movements.

Keywords: biomechanical simulation, quantal effects, phonetic typology, motor control

1. INTRODUCTION

Languages display a robust tendency to use different lip shapes for different degrees of labial constriction. Of the 451 languages in the UCLA Phonological Segment Inventory Database (UPSID) [1, 2], 446 (99%) have a bilabial stop (e.g. /b/, /p/, /m/ or variants), while only one language is reported as having a labiodental stop (0.2%). Conversely, for labial fricatives, 82 languages (18%) are reported to have bilabial fricatives (e.g., /ɸ/, /β/), while 199 (44%) have at least one labiodental fricative (e.g., /f/, /v/). For approximants the story is different yet again: only 6 languages (1.33%) have labiodental approximants, while 336 (75%) have rounded approximants (e.g., /w/, /ʍ/).

Although it is not without exceptions, particularly in the case of fricatives, there is a clear generalization across languages that labial stops are typically produced using a flat constriction of the margins of both lips (Fig. 1a), labial fricatives using contact between the lower lip and upper teeth (Fig. 1b), and labial approximants using lip rounding and protrusion (Fig. 1c).

Figure 1: Typical lip postures for speech events.

(a) (b) (c)

In principle, this does not have to be the case. A language could just as well produce different degrees of constriction by varying the activation of a single labial movement. For example, a language could generate the voiceless bilabial stop, fricative, and approximant as something like [p], [ɸ], and [ɸ̞].

Although this is conceptually straightforward, it is typologically unattested. That is, given the knowledge that a labial sound is produced with a certain degree of constriction, we may predict other properties of the constriction, and vice versa. This suggests that types of constriction at the lips should not be specified by combinations of independent parameters like the degree of constriction plus the articulators involved, but rather by more coherent, dedicated structures consisting of a set of muscles that activate in fixed proportion to one another, and are organized to generate a specific output (e.g. [3, 4]).

An outstanding question is why these structures are employed time and time again across languages. In general, we may expect that structures built for a task will tend to be robust to noisy, everyday conditions. Such structures should allow a large margin of error and optimize for feed-forward function (i.e., operating without correction based on immediate sensory feedback). Some speech mechanisms have been described as having properties of this kind, and are often associated with the term “quantal.”

The term “quantal” has been applied to a subset of non-linear effects in speech – traditionally those that facilitate some auditory-perceptual goal [5, 6, 7]. 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 little response in some other (task) dimension. Although this literature focuses on the auditory-perceptual domain, these effects have been predicted to obtain across a variety of sensory domains so long as such effects enhance the feed-forward aspect of speech control [8].

Few biomechanical quantal effects have been discussed in the literature (e.g., [9, 10, 11]). Such effects have, however, been shown in simulation studies of vocal tract structures, such as the soft palate [12] and the larynx [13]. Similar effects have been predicted for labial movements, specifically labiodental fricatives [8], and have been found for labial movements to limited degrees in simulation studies: e.g., studies have shown that the rounded, protruded lip posture for approximants is robust to varying muscle activation when muscle stiffness is simulated [14], and that both lip closure for stops and protrusion for approximants are robust to variation in muscle activation [15].

The goal of the present paper is to test for quantal effects in the three canonical lip postures described above. We did so by simulating lip constrictions using a three-dimensional finite-element face model implemented using the biomechanical modelling platform Artisynth (e.g. [16]; www.artisynth.org). Artisynth is capable of simulating tissue biomechanics and the actions of spatially fixed groupings of muscles in the vocal tract. In addition to passive tissue mechanics, the face model accounts for active muscle stress and intrinsic stiffness, volume preservation, and gravity. The facial muscles in the model that are relevant to the simulations in this paper are shown in Fig. 2. This paper extends the work in [15], adding labiodentals to the simulations in Section 2, and presenting new simulation results in Section 3.

Under the hypothesis that speech movements are the outputs of discrete, functionally independent structures, and that these structures are selected for use in speech in part because they take advantage of intrinsic quantal 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 (see [11]). This is tested in Section 2. In addition, we expect the outcomes will be robust to interference caused by surrounding muscle activations. This is tested in Section 3. If borne out, these results will provide a biomechanical basis for the robust typological generalizations described above.

2. ROBUSTNESS TO VARYING ACTIVATION

To test for robustness to varying muscle activation, we defined groupings of muscles based on known muscle involvements in the three most widely attested speech lip postures, corresponding to a stop constriction (as in [p]), a fricative constriction (as in [f]), and an approximant constriction (as in [w]). Lip shapes were achieved by activating muscles up to a maximum stress as indicated in Table 1. Simulated postures for rest position, bilabial closure, labiodental constriction, and approximant constriction are shown in Fig. 3.

Figure 2: Muscles in the face model: superior and inferior peripheral orbicularis oris (OOPs/OOPi), superior and inferior marginal orbicularis oris (OOMs/OOMi), mentalis (MENT), risorius (RIS), levator labii superioris alaeque nasi (LLSAN), and levator labii superioris (LLS).

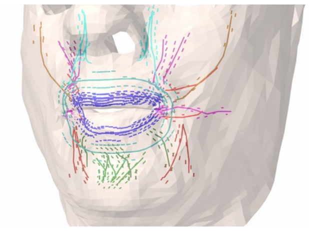
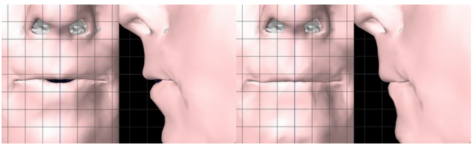


Table 1: Maximum muscle stress (kPa) used for the three lip constrictions.

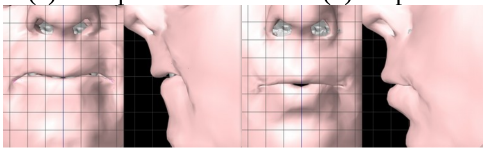
	OOPs	OOPi	OOMs	OOMi	MENT	RIS	LLSAN	LLS
Bilabial closure	–	–	30	30	20	20	–	–
Labio-dental	–	–	–	26	26	26	36	50
Approx.	40	40	–	–	–	–	–	–

The plot in Fig. 4 shows nonlinearities occurring, as predicted, for all three labial postures. The maximum stress values in Table 1 (equivalent to 100% of the x-axis in Fig. 4) correspond to an average of about half of the absolute maximum voluntary muscle contraction. Quantal regions (relatively “flatter” parts of the plots) are enclosed in gray boxes in Fig. 4, indicating the region of the graph beyond which 95% of total distance from maximum opening to maximum constriction has been covered. Opening size for the labiodental was calculated as the area between the lower lip and the teeth, rather than the area between the lips, giving it a larger initial opening (and prompting plotting as percentages on the y-axis). In all cases, the area was calculated by counting pixels in the opening from coronal images of the face and converting these to mm^2. Increased activation beyond the 100% threshold yielded further compression for the stop and fricative, and slightly more protrusion for the approximant, but did not affect gross posture.

Figure 3: Postures for labial constrictions.

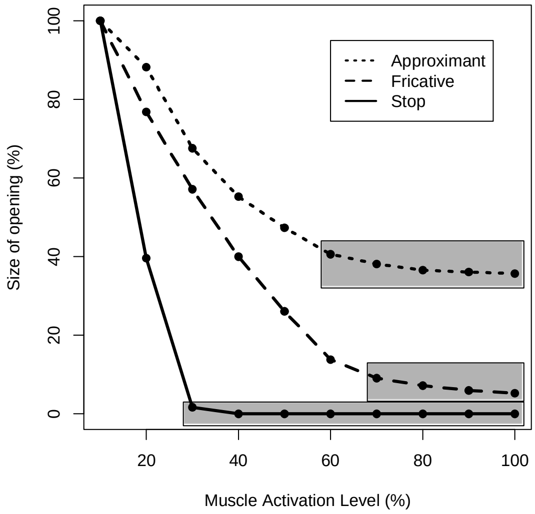


(a) Rest position (b) Stop



(c) Fricative (d) Approximant

Figure 4: Size of lip opening as a function of muscle activation for the three canonical postures. Quantal regions (95%) are indicated in gray.



The degree of quantality of a function (here from activation level to size of lip opening) can be quantified using the *quantality score* (Q-score) proposed in [13]. Higher Q-scores indicate that the function shows a region of significant change at lower input values, but quickly stabilizes into a region of low change at higher input values.

The Q-scores for the stop, fricative, and approximants from the above simulations are 1.681, 0.670, and 0.728 respectively. Using the ranges provided in [13], 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 desired equivalent postures using fixed sets of muscles.

3. ROBUSTNESS TO SURROUNDING MUSCLES

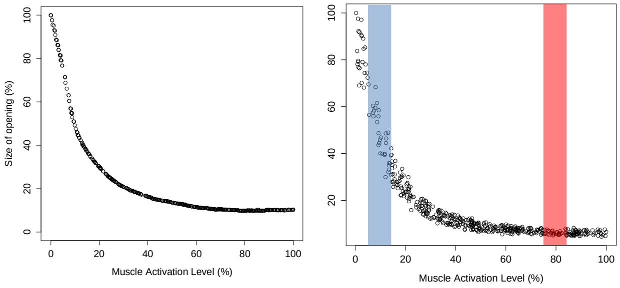
The previous simulations showed that the canonical postures used to produce the three degrees of labial constriction are robust to variation in the activation levels of the relevant muscle groups. A second question is whether these postures are also robust to interference from activation of surrounding muscles.

We performed two types of simulations focusing on the robustness of the rounding gesture used in approximants (generated by activating the OOP) to interference from surrounding muscle activations. The first type tested whether lip closure produced by OOP activation remained stable in the presence of surrounding muscle noise, while the second tested the effect of different degrees of OOP activation on this stability.

Rather than counting pixels, an automatic measurement scheme was devised to handle the large number of simulations. This consisted of calculating the minimum opening size along a series of cutting planes placed between the lips, and resulted in smaller minimum opening sizes than those reported in the previous section. To run the simulations, we used the BatchSim tool, which permits automatic probabilistic sampling of any model parameter: in this case, muscle activations.

In the first simulation type, we sampled the activation space of the OOP uniformly from 0% to 100% activation both without activation of any other muscles (500 total samples; 481 successful), and with other muscles (the muscles in Table 1 plus depressor anguli oris, buccinator, depressor labii inferioris, levator anguli oris, and zygomaticus) excited randomly from a uniform distribution between 0% and 10% activation (500 total samples; 475 successful). Unsuccessful samples were those that resulted in invalid model states.

Figure 5: The results of the probabilistic sampling simulations on the full range of OOP activation.



The results of these simulations are shown in Fig. 5. The left side shows activation of only the OOP. Lip opening decreases with increased OOP activation, but never to the point of complete closure, as in the previous simulations. The right side shows that even in the presence of noise from surrounding muscle activations, the lip closure area remains fairly stable, particularly at high OOP activation levels, although the overall constriction size decreases across the board. This greater overall constriction is not surprising, since the majority of the perturbed muscles serve to close the lips.

The second set of simulations sampled OOP activation from a normal distribution with a mean of 10% and a standard deviation of 10% (low activation; 500 total samples; 474 successful), and from a normal distribution with a mean of 80% and a standard deviation of 10% (high activation; 500 total samples; 481 successful). Both sets of samples were in the presence of noise from other muscles again sampled uniformly from 0% to 10% activation. These results are shown in Fig. 6. The correlation between OOP activation and lip opening area in the low activation case was -0.88, and the standard deviation of lip area was 17.37%. In the high activation case, the correlation was -0.19, and the standard deviation of lip area was 0.96%. Welch’s two sample *t*-test shows that the opening sizes are significantly different between the two conditions [t(475.84) = 52.083, p < 2.2e-16], and Levene’s test shows that the variance is significantly different as well [F(1, 953) = 689.07, p < 2.2e-16]. Comparison of the two correlations using Fisher’s transformation shows that the correlations are significantly different [Z = -18.865, p < 2.2e-16]. The OOP muscle activation accounts for about 79% of lip area variation at low activation, but only 3.5% of lip area variation at high excitation.

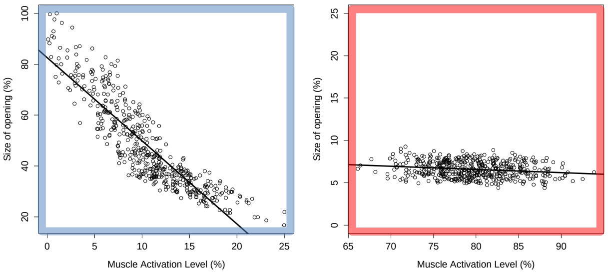
These results indicate that the activation of OOP for lip rounding gestures maintains relatively consistent degrees of lip opening, even in the presence of noise from the activation of surrounding muscles, and that higher OOP activation results in reduced interference from surrounding muscles. Note that the area of higher activation tested in these simulations corresponds roughly to the quantal region shown in Fig. 4. This indicates that the same quantal region of activation space is robust to both intrinsic and extrinsic activation noise.

4. DISCUSSION

The results of the simulations presented above lend support to the hypothesis that the typologically most frequent lip postures used in speech correspond with quantal properties of those postures. It is important to note that quantality is not a property of all sets of muscle activations, as shown by [12].

The simulations in Section 2 show that each set of muscles associated with the three canonical postures generates large biomechanical quantal regions where the target constriction is met, indicating robustness to intrinsic activation noise. The simulations in Section 3 show that the quantal region for lip rounding is also robust to extrinsic noise from surrounding muscle activations.

Figure 6: The results of the probabilistic sampling simulations on selected ranges of OOP activation.



The hypothetical unattested labial inventory [p], [ɸ], and [ɸ̞] is less desirable because, as shown in Fig. 4, the regions in which frication and approximation are achievable using this muscle configuration are biomechanically unstable, providing uncertain, highly variable outcomes.

There is a general correspondence between the degree of quantality of each constriction type and its typological prevalence. The exception is the labiodental fricative, which has a similar Q-score to the approximant, but is less common cross-linguistically. It may be the case that the mechanism used for bilabial fricative constriction is not the same as that for bilabial stop closure (e.g., lip compression, as in [18]), and that this mechanism serves as a competing, though perhaps less effective, quantal alternative to labiodental closure. These simulations also do not include potentially relevant factors such as aerodynamics that may help explain tendencies such as voiced labiodental fricatives alternating with labiodental approximants.

In addition to suggesting that biomechanics contributes to the typological distribution of canonical lip postures, this work also bears on theories of speech motor control. The consistent use of these specialized postures for different speech sounds suggests that the primitive units of speech motor organization should be modular muscle groupings that activate in fixed proportion to one another to perform a particular task – in this case, to achieve a particular degree of labial constriction (see [17, 3]). The present paper, we hope, demonstrates the value of such structures, in the lips or otherwise, for providing explanatory power for linguistic phenomena like the typological patterns described here, as well as narrowing the scope of potential solutions to general problems in speech motor control and articulatory speech synthesis. We anticipate similar value in understanding phenomena in phonetics, phonology and sound change.

5. REFERENCES

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

[2] Maddieson, I. & Precoda, K. 1990. Updating UPSID. *UCLA Working Papers in Phonetics*. Department of Linguistics, UCLA, 104–111.

[3] Gick, B., Stavness, I. 2013. Modularizing Speech. *Frontiers in Psychology* 4, 977.

[4] Gick, B., Schellenberg, M., Stavness, I., Taylor, R. 2018. Articulatory Phonetics. In: Katz, W.F., Assmann, P. (eds), *The Routledge Handbook of Phonetics.* Ch 5. New York: Taylor & Francis.

[5] Stevens, K.N. 1972. The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data. In: David, E.E. Jr., Denes, P.B. (eds), *Human Communication: A Unified View*. New York: McGraw-Hill, 51-66.

[6] Stevens, K. 1989. On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.

[7] Stevens, K., Keyser, S.J. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics*, 38, 10-19.

[8] Fujimura, O. 1989. Comments on “On the Quantal Nature of Speech,” by K.N. Stevens. *Journal of Phonetics*, 17, 87-90.

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

[10] Perkell, J.S., Matthies, M.L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., Stockmann, E., Guenther, F.H. 2004. The distinctness of speakers’ /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research*, 47(6), 1259-1269.

[11] Perkell, J.S. 2012. Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, 25(5), 382-407.

[12] Gick, B., Anderson, P., Chen, H., Chiu, C., Kwon, H.B., Stavness, I., Tsou, L., Fels, S. 2014. Speech function of the oropharyngeal isthmus: A modeling study. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2(4), 217-222.

[13] Moisik, S., Gick, B. 2017. The quantal larynx: The stable regions of biomechanics and implications for speech production. *Journal of Speech, Language and Hearing Research*, 60, 540-560.

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

[15] Gick, B., Stavness, I., Chiu, C., Fels, S.S. 2011. Categorical variation in lip posture is determined by quantal biomechanical-articulatory relations. *Canadian Acoustics*, 39(3), 178-179.

[16] Stavness, I., Lloyd, J.E., Fels. S.S. 2012. Automatic Prediction of Tongue Muscle Activations Using a Finite Element Model. *Journal of Biomechanics*, 45(16), 2841-2848.

[17] Safavynia, S.A., Ting, L.H. 2013. Sensorimotor feedback based on task-relevant error robustly predicts temporal recruitment and multidirectional tuning of muscle synergies. *Journal of Neurophysiology*, 109, 31-45.

[18] Okada, H. 1991. Japanese. *Journal of the International Phonetic Association*, 21(2), 94-96.