FROM QUANTAL BIOMECHANICS TO WHOLE EVENTS: TOWARD A MULTIDIMENSIONAL MODEL FOR EMERGENT LANGUAGE

Bryan Gick[1, 2]

[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T1Z4 gick@mail.ubc.ca
[2]Haskins Laboratories, 300 George St., New Haven, CT, USA 06511

1. INTRODUCTION

Nature is replete with nonlinearities. In speech, the simple act of closing the lips corresponds with a flood of mechanical and sensory nonlinearities: visual, kinematic, somatosensory, aerodynamic, acoustic-auditory, and so on. The term "quantal" has been applied to a subset of nonlinear effects in speech that help in achieving some auditory-perceptual goal (Stevens 1972, 1989). Stevens & Keyser (2006) define quantal effects as discontinuities "between the displacement of an articulatory parameter and the acoustical attribute that results from this articulatory movement," arguing that these effects provide the basis for phonological categories. Nonlinear effects fitting this definition of quantal thus share the three properties of: (a) originating in articulation, (b) having demonstrable auditory-perceptual implications, and (c) being advantageous in the communicative process.

According to this view, some of the nonlinearities described above for lip closure qualify as "quantal" while others – specifically those with no direct articulatory-auditory link – do not. However, many studies in the last decade have identified effects that are important in speech but that are independent of auditory-perceptual goals. For example, Tremblay et al. (2002) show that jaw movements associated with speech-related constrictions maintain their own proprioceptive trajectories. Likewise, Ménard & al. (2009) found that sighted speakers use lip movements differently from blind speakers for speech. Similar work has shown independent effects in aerodynamic (Gick & al. 2012), tactile (Gick & al. 2008) and aerotactile (Gick & Derrick 2009, Francis & al. this volume) dimensions.

In this model, speakers do not aim to achieve "movement goals/targets" as defined in physical, auditory-perceptual or somatosensory space, but rather "whole events" in a fully multidimensional space. The present paper demonstrates how a biomechanical nonlinearity (as described in Gick & al. 2011; see below) can be translated into a multidimensional density map, creating a topology of relative densities. Within this space, those nonlinearities that facilitate the production of ecologically successful events function as attractors to behavior. These "facilitiative nonlinearities" become the common currency of human ecological space, and coupled with an iterative learning simulator, are the basis for an emergent model of phonology or other patterned behavior.

2. VOCAL TRACT SPHINCTERS AS QUANTAL BIOMECHANICAL DEVICES

Previous models of speech articulation have been based predominantly on spatial targets in a 2-dimensional (midsagittal) vocal tract space. The present paper describes an approach to speech articulation based on a view of vocal tract constrictors as physiological sphincter mechanisms. In this model, a particular constrictor (sphincter) can produce only one kind of constriction, and it does so using the inherent quantal biomechanical properties of sphincters. Sphincters in this model differ from articulators in previous models in that they: (a) generate constrictions rather than targeted motions; (b) constrict from multiple directions in 3 spatial dimensions; (c) are functionally idiosyncratic to fixed locations in the vocal tract; (d) are highly constrained in degrees of freedom; (e) are fully predictive of location, degree, articulator and shape; (f) provide no inherent hierarchical structure; (g) exhibit robust mechanical nonlinearities (quantal effects) that generate discrete output from continuous, noisy input; (h) enable entirely feed-forward control; and (i) function as attractors to behavior in a multidimensional behavior space.

Figure 1 shows an example of three biomechanical nonlinear regions (highlighted) corresponding to quantal-like regions in the actions of the three lip sphincters used to produce labial approximants, fricatives and stops (adapted from Gick & al. 2011, which gives a more detailed description of this graph). These regions represent regions of biomechanical efficiency.

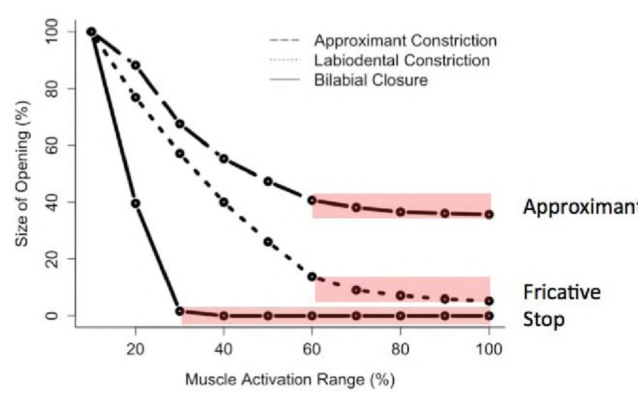

Figure. 1. Facilitative nonlinearities (highlighted) in lip opening size as a function of muscle activation.

3. GETTING FROM QUANTAL REGIONS TO MULTIDIMENSIONAL SPACE

Figure 2 shows how a potentially advantageous nonlinearity (shown in Figure 1) can be reduced in dimensionality in order to be incorporated into a multidimensional functional space. This process reduces the 2-dimensional size-by-activation space of Figure 1 to a single "biomechanical advantage" dimension, with three attractor regions based on the slope of the Figure 1 curves.

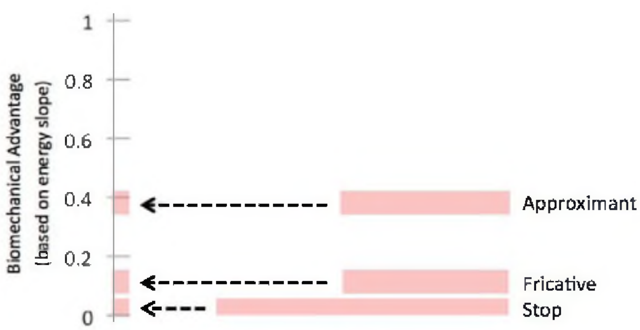

Figure. 2. Dimensionality reduction of lip data.

Figure 3 shows the lip data in the biomechanical advantage dimension now plotted in a multidimensional space, along with hypothetical "auditory advantage" and "visual advantage" dimensions. In this density- or heat-map plot, high-advantage areas are plotted as dense or hot regions.

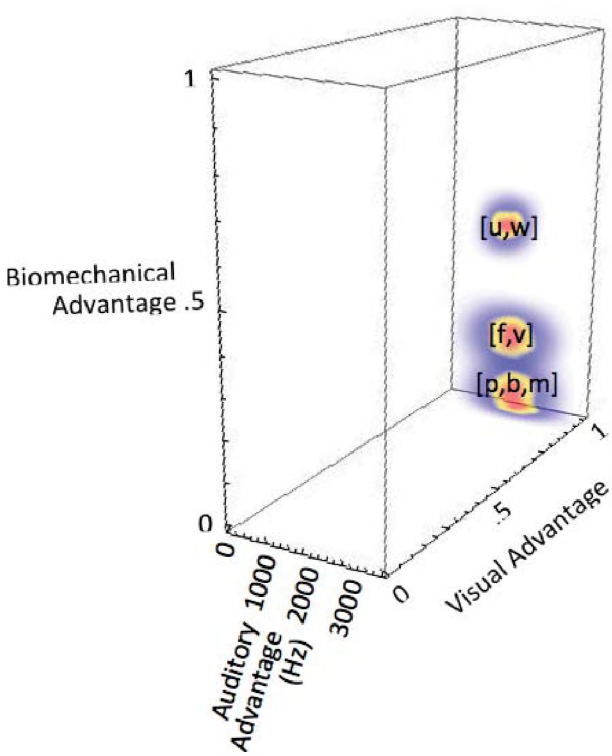

Figure. 3. Labial sound categories plotted in 4-D space.

4. DISCUSSION

The present paper is part of a research program whose aim is to identify a much wider range of potentially advantageous nonlinearities, and explore their role in an emergent model of phonetics and phonology. This work models the operational space in which spoken communication takes place as a fully multidimensional (physical, social, ecological) "whole event" space. Within this space, behavior is attracted to nonlinearities that afford users communicative advantage, allowing previously unrelated dimensions of human behavior to interact in a single functional space. The common currency of this space is communicative advantage, and it is only by offering demonstrative communicative advantage that a behavior may be adopted as exerting an effect. This approach interfaces with previous models that use nonlinear dynamical systems to simulate emergent language (e.g., Smith & Thelen 2003, Gafos 2006, etc.), models that have sought nonlinear relationships as important keys to understanding speech and language (Perkell 2010), and other empirically based, multimodal approaches to emergence of language (Mielke 2011).

REFERENCES

Gafos, A.I. and Benus, S. (2006). Dynamics of Phonological Cognition. *Cognitive Science 30*, 1-39.

Gick, B., Chiu, C., Flynn, C., Stavness, I., Francis, N. and Fels, S. (2012). Producing whole speech events: Differential facial stiffness across the labial stops. *Proc. of Acoust. 2012 HK*. Ppr 2pSC1, 6pp.

Gick, B. and Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*. *462*. 502-504.

Gick, B., Jóhannsdóttir, K., Gibraiel, D. and Muehlbauer, J. (2008). Tactile enhancement of auditory and visual speech perception in untrained perceivers. *JASA* 123(4), EL72-76.

Gick, B., I. Stavness, C. Chiu and S. S. Fels. (2011). Categorical variation in lip posture is determined by quantal biomechanical-articulatory relations. Proceedings of Acoustics Week in Canada 2011. *Canadian Acoustics* 39(3).

Mielke, J. (2011). A phonetically based metric of sound similarity. *Lingua*. doi:10.1016/j. lingua.2011.04.006.

Ménard, Lucie, Sophie Dupont, Shari R. Baum, & Jérôme Aubin. (2009). Production and perception of French vowels by congenitally blind and sighted adults. *JASA*, 126 (3) 1406-1414.

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

Smith, L.B. and E. Thelen. Development as a dynamic system. TiCS 7, 343-348.

Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. David & P. Denes (Eds.), *Human communication: A unified view*. NY: McGraw-Hill. 51-66.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics* , 17, 3–4.

Stevens, K. N. and Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, 38, 10-19.

ACKNOWLEDGEMENTS

The author acknowledges the many collaborators who have contributed to this work, and funding from NSERC.