TWO PHONOLOGICAL SEGMENTS, ONE MOTOR EVENT: EVIDENCE FOR SPEECH-MOTOR DISPARITY FROM ENGLISH FLAP PRODUCTION

Donald Derrick[1], Bryan Gick[1, 2], and Ian Stavness[3]
[1]Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada, V6T 1Z4
[2]Haskins Laboratories, 300 George St., 9th Fl., New Haven, CT 16511, USA
[3]Dept. of Electrical and Computer Engineering, University of British Columbia, 5500 - 2332 Main Mall, Vancouver, BC, Canada, V6T 1Z4

1. INTRODUCTION

A pervasive assumption in speech motor behaviour is that units of speech production correspond to preplanned motor routines. In a possible counterexample, Browman and Goldstein [1] showed that, kinematically, a single glottal opening event can span two adjacent segments in s + stop clusters. However, because the segments are adjacent, it has not been clear whether this is a case of preplanned motor efficiency or simply a local concatenation of two separate motor events.

Here we present an argument that in the English word `Saturday', the two flaps may be generated as one articulatory gesture at the onset of the first (upward) flap, while the rhotic and the second (downward) flap occur automatically as a result of gravity and elasticity.

We know that in English there are four subphonemic categorical kinematic variations of flaps, identifiable via B/M ultrasound [2,3]. The first is an alveolar tap (AT). The tongue (tip) moves from below the alveolar ridge, makes contact and moves back. The second is a down-flap (DF). The tongue moves from above the alveolar ridge, makes contact, and continues downwards. The third is an up-flap (UF). The tongue moves from below the alveolar ridge, makes contact, and continues. The fourth is a postalveolar tap (PAT). The tongue moves from above the alveolar ridge, makes contact, and moves back.

There is also a strong relationship between the type of flap selected and the tongue tip position before and after the flap. In the case of a non-rhotic vowel preceding and a tongue tip-up rhotic vowel following the flap, there is a higher likelihood of UF production. In the case of a tip-up rhotic vowel, there is a higher likelihood of a DF [3].

Previous work [4] found an unexplained preference for a UF-DF movement pattern in sequences of flap allophones. Also, DFs are unique among the four variants in that they take advantage of gravity and elasticity: Gravity because, while a speaker is upright, the tongue tip moves from a high position to a low position, and elasticity because the tongue moves from a retroflex to a tongue rest position vowel (see [5]).

The human nervous system does not completely compensate for the effects of gravitational load on jaw motion [6]. Jaw motion during speech differs based on whether a speaker is prone or supine, and tongue motion does not entirely compensate in place of jaw motion.

Similarly, [7] has provided experimental 2D simulation based evidence that much of the forward looping pattern of velar stop production in VCV sequences is based on the anatomical structure of the tongue such that planning may be based on target sequence as much as trajectory motion.

If our hypothesis about flaps in 'Saturday' is correct, we expect more tip-up rhotics in 'Saturday' than in the no-flap control 'peppermint'. We don't expect such differences with 'herded her', a phrase where we expect DF-UF sequences, vs. 'herbifer'. We also expect shorter duration between flaps, and less variability in the flap sequences for 'Saturday' than for the phrase 'herded her'. A production experiment tests these predictions.

To identify whether gravity and myoelasticity alone can, in principle, complete a DF closure, we will also use biomechanical simulation. Biomechanical simulation is well suited for this study as it characterizes how forces within the system interact in order to generate observed movements. ArtiSynth is a biomechanics simulation toolkit, targeted toward modeling and simulation of the human vocal tract [8]. Recently, a model of coupled jaw-tongue-hyoid dynamics has been developed within the ArtiSynth framework [9] that includes bone structures, a deformable tongue model, muscle forces, dynamic coupling (tongue muscle forces act on the jaw and vice versa), and contact (tongue-jaw and tongue-palate).

2. METHODS

Eighteen native speakers of NAE participated. Participants were seated in an adjustable chair with headrest. Microphone and B/M ultrasound recordings were taken of them producing 12 repetitions of 38 sentences, 17 control sentences, and 19 sentences with single and double flap sequences. We focused on four words within the study, 'Saturday', 'herded her', 'peppermint', and 'herbifer'. We measured flap, and rhotic type and duration between flaps.

The reference tongue model [10] uses the Finite Element Method (FEM) to represent the non-linear, large deformation tissue properties of the tongue. Muscle fibers are embedded to represent the muscle structure of the tongue. The reference jaw model [11] is composed of rigid body components for the skeletal structures (cranium, mandible, hyoid bone) connected with point-to-point Hill-type muscles and has been used to analyze forces during unilateral chewing. Simulations reported for the coupled jaw-tongue-hyoid model have shown plausible speech and chewing motions.

3. RESULTS

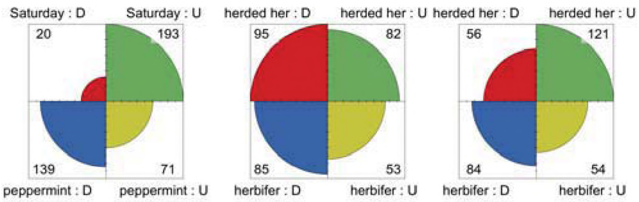

Fig. 1. Rhotic variants by subject by flap phrase (top) vs. control (bottom): U = tip-up, D = tip-down. Left = 'Saturday' vs 'peppermint', mid = initial rhotic 'herded her' vs. 'herbifer', right = final rhotic 'herded her' vs. 'herbifer'.

The difference in rhotic orientation is significant for 'Saturday' vs. 'peppermint' (AIC = 405, c = 2.94, z = 10.6, *p < 0.001), but not for the initial or final rhotic in 'herded her' vs. 'herbifer'.

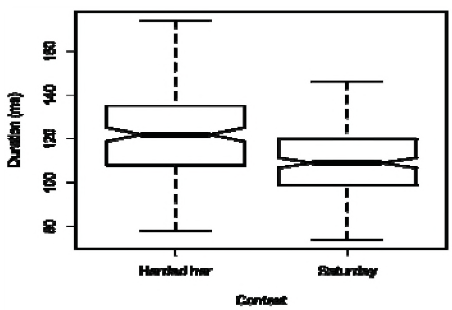

Fig. 2. Duration of 'Saturday' vs. 'herded her'.

The duration between flaps is shorter for 'Saturday' vs. 'herded her' [glmer (REML) t = -8.62 (AIC = -2091)].

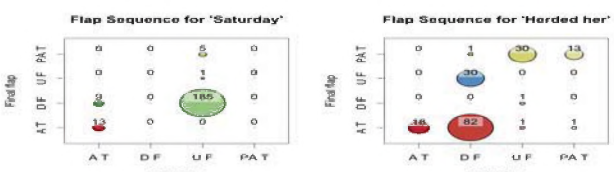

Fig. 3. Flap sequences in 'Saturday' vs. 'herded her'.

Flap sequence variability is much lower in 'Saturday' (mostly UF-DF) than 'herded her' (some DF-AT, and four other sequences prominent), as shown in by Fig. 3.

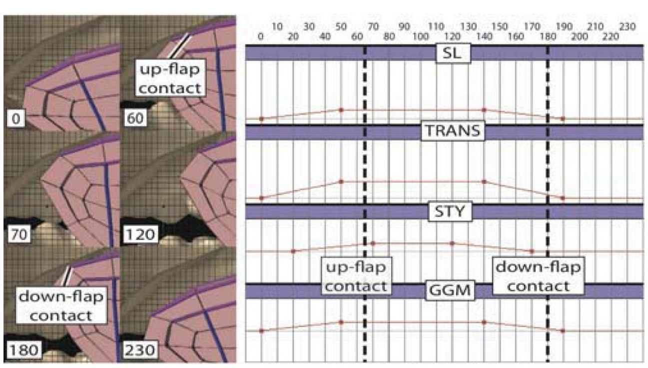

Fig. 4. Frame-by-Frame illustration of UF-DF sequence (left) in relation to muscle activation (right).

In the simulation, contracting the superior-longitudinal (SL), transverse (TRANS) and mid-genioglossus (GGM) lifts the tongue to the alveolar ridge, and the styloglossus (SG) pulls the tongue tip farther up into a retroflex. Upon relaxation, the tongue tip lowers, contacting the alveolar ridge en route. Duration between flaps is 115 ms, similar to Fig. 2.

4. DISCUSSION

The results of the experiment support all three hypotheses, indicating that an UF-DF sequence in the word 'Saturday' is preferred, and produced more quickly than flap sequences in 'herded her'. The simulation supports the hypothesis that the same cluster of muscle contractions can produce UFs and retroflex rhotics, and relaxing the muscles will produce DFs. The results demonstrate disparity between phonology and motor behaviour in that one motor event can encompass the production of two segments spanning a syllable boundary.

REFERENCES

[1] Browman, C. P. and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3:219–252.

[2] Derrick, D. and Gick. B. (2008). Quantitative analysis of subphonemic flap/tap variation in NAE. *Proceedings of the Acoustics Week in Canada*, 36-3:162-163.

[3] Derrick, D. and Gick, B. (2010). There are no fixed motor programs in speech: Evidence from categorical kinematic variants of English flaps and taps, *Under review*.

[4] Gick, B. (2006). Errors and strategy shifts in speech production indicate multiple levels of representation. *Journal of the Acoustical Society of America*, 119 (5):3302.

[5] Gick, B., Wilson, I., Koch, K., and Cook, C. (2004). Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica*, 61:220–233.

[6] Shiller, D. M., Ostry, D. J., and Gribble, P. L. (1999). Effects of Gravitational Load on Jaw Movements in Speech. *Journal of Neuroscience*, 19(20):9073–9080.

[7] Perrier, P., Payan, Y., Zandipour, M., and Perkell, J. (2003). Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114(3):1582–1599.

[8] Fels, S., Stavness, I., Hannam, A., Lloyd, J. E., Anderson, P., Batty, C., Chen, H., Combe, C., Pang, T., Mandal, T., Teixeira, B., Greena, S., Bridson, R., Lowe, A., Almeida, F., Fleetham, J., and Abugharbieh, R. (2009). Advanced tools for biomechanical modeling of the oral, pharyngeal, and laryngeal complex. In *International Symposium on Biomechanics Healthcare and Information Science*, page online.

[9] Stavness, I., Lloyd, J., Payan, Y., and Fels, S. (2010). Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. *Under review.*

[10] Buchaillard, S., Perrier, P., and Payan, Y. (2009). A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *Journal of the Acoustical Society of America*, 126(4):2033–2051.

[11] Hannam, A., Stavness, I., Lloyd, J. E., and Fels, S. (2008). A dynamic model of jaw and hyoid biomechanics during chewing. *Journal of Biomechanics*, 41(5):1069–1076.

ACKNOWLEDGMENTS

This research was funded by (NSERC) to the second author, and by National Institutes of Health (NIH) Grant DC-02717 to Haskins Laboratories. The authors thank Alan Hannam for providing the jaw model and his expertise, Gipsa-Lab, Grenoble for providing CT image data (P. Badin) and the tongue model geometry (S. Buchaillard, J.M. Gérard, P. Perrier, Y. Payan), Sid Fels and the ArtiSynth team.