COUPLED OSCILLATOR PLANNING MODEL OF SPEECH TIMING AND SYLLABLE STRUCTURE

Louis Goldstein, Hosung Nam, Elliot Saltzman & Ioana Chitoran

ABSTRACT

A fundamental problem in understanding speech production is how the temporal coherence of the speech units associated with a given lexical unit is maintained despite changes due to speaking rate, prosodic embedding, and transient perturbations. To address this, a dynamical model of temporal planning of speech has been developed [21, 13, 26, 27). In this model, each speech unit (constriction gesture) is associated with a planning oscillator, or clock, and the oscillators within the ensemble associated with a particular lexical item are coupled to one another in a pattern represented as a coupling graph.

Given this model, it is possible to account for syllable structure in terms of intrinsic modes of coupling and the topology of the coupling graph. Onset consonant gestures are hypothesized to be coupled in-phase to the tautosyllabic vowel (regardless of how many there are in an onset), while coda consonant gestures are coupled in an anti-phase pattern. This topology can account simultaneously for regularities in relative timing and variability, and examples of this will be discussed.

Despite successes obtained with the model, it is clear that there are examples in which the same syllable structure can exhibit different patterns of timing, depending, e.g., on the place of articulation of the consonants in the cluster [10], manner [7,17] or language. In this paper, we will also illustrate how these different patterns of timing can be modeled using coupling graphs with differing topologies and/or with different quantitative specification of coupling strength associated with the graph's edges.

.

Keywords: articulatory phonology, coupling graph, syllable structure

1. INTRODUCTION

Like other combinatorial systems, phonology can be can be analyzed into a set of primitive atoms and some glue that holds them together in combinations. From the perspective of articulatory phonology [3], the atoms are the distinct vocal tract actions (gestures) used by talkers and listeners to distinguish words from one another in a communication system. Words are composed of an ensemble of gestures, organized in time in a particular fashion. The temporal organization is itself informational. An example of this can be seen in Figure 1, which shows gestural scores for the words "mad" and "ban." The scores show the temporal intervals during which gestures of the various vocal tract constriction devices are active and control their corresponding articulator sets to produce the constriction task goals. The time at which the velic lowering (wide) gesture occurs is the only difference between the scores. Thus, the speech production system must include some kind of glue that insures the temporal stability of the informational pattern, for example in the case of Fig. 1, insuring that the velum lowering gesture does not stray too far during the production of "mad," so that "ban" is perceived instead.

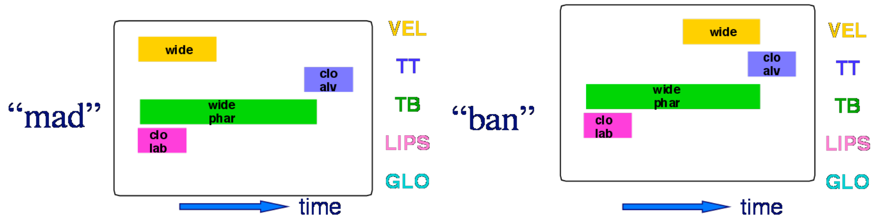

Figure 1: Gestural Scores for "mad" and "ban"

The hypothesis that underlies our recent modeling efforts [23, 26, 27] is that the required glue is to be found in dynamical coupling in a system of planning oscillators, or clocks. Coupled oscillators, when entrained, can exhibit the property of phase-locking. They oscillate with a stable relative phase and if perturbed, they return to their stable relative phase. If the gestures are associated with such clocks, and if their activations are triggered at particular phases of their clocks, stable relative phasing would be insured.

Given the decision to employ coupled oscillators in speech production planning, there are at least two alternative system architectures that could be constructed. In one, each of the gesture oscillators is coupled to an external master clock. This will allow the temporal stability of the gestural pattern to be maintained, and will allow some flexibility: as the master clock changes its frequency, this will induce temporal changes in the patterns of gestural triggering. This type of system is similar in several ways to the C/D model of Fujimura [12], in which each gesture is timed to some phase of a syllable pulse. The alternative, which we have pursued in our model, is to couple the gesture's clocks to one another in a pair-wise network, described as a coupling graph. (In more recent work [27] hierarchical coupling of gesture clocks to foot and phrase level oscillators is also implemented, but not every gesture is coupled to a higher-level oscillator).

There are at least two reasons for preferring the pair-wise coupling network to the master clock:

- The stability of relative phase of neighboring gestures can vary as a function of several factors. For example, [7] the relative phase of consonant gestures in a syllable onset is less variable than the relative phase of the same gestures in a syllable coda. If all gestures were coupled to an external clock, such differential stability would not be expected.
- Coupled oscillators can entrain in intrinsically accessible modes, and applying those modes to pair-wise inter-gestural coupling provides the basis for an embodied theory of syllable structure. As discussed in section 2, this theory can account for macroscopic universals of syllable structure, and at the same time, microscopic properties of gestural relative timing and its variability.

2. 2. COUPLING MODES AND SYLLABLE STRUCTURE

In the model of speech production we have been developing [27, 27], the planning of gesture relative timing (the gestural score) takes place through an oscillatory process. Each gesture of the utterance is associated with a planning oscillator, or clock, and gestures are coupled pairwise to one another in an utterance specific manner. As planning for an utterance is initiated, all of the gesture clocks begin oscillating at random phases with respect to one another. Over time (oscillator cycles) coupling forces cause local changes in phase of individual planning oscillators, and the system eventually settles into a stable pattern of oscillator relative phases. Once the system stabilizes, activation of a gesture (left edge of its gestural score box) is triggered at phase 0 of its component oscillator.

The pattern of inter-oscillator coupling associated with a particular lexical unit is represented using a coupling graph. The nodes of the graph represent the oscillators associated with individual gestures, the edges of the graph connect pairs of coupled oscillators. The edges (or links) specify the target relative phase of the associated pair of oscillators. In the task-dynamic model of relative phase planning, the target relative phase acts as the minimum of a potential function [24], that induces forces on the component oscillators so that relative phase of the pair is attracted to its target value.

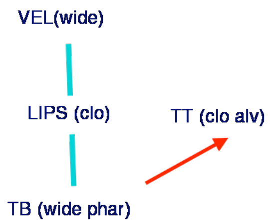

Figure 2: Coupling graph for "mad"

The coupling graph for the word "mad" is shown in Figure 2. The four gestures (also seen in the gestural score in Figure 1) composing the word are the nodes. Edges come in two types: lines and arrows. The lines represent edges whose target specification is in-phase (zero degrees relative phase), the arrows represent edges whose target specification is anti-phase. For anti-phase targets, an additional bit of information is required to specify the order in which gestures are triggered (which oscillator's phase 0 triggers its oscillator first). The arrow points from the earlier gesture to the later one.

2.1. Coupling Modes and syllable structure

Research on coordination of multiple rhythms has shown that there are two intrinsically accessible, stable modes in which subjects can, for example, coordinate the movements of multiple limbs—in-phase and anti-phase [15, 29]. Intrinsic accessibility means that the subjects can perform the task in one of these modes without training or learning. Of these two modes, in-phase is more accessible and stable. We have hypothesized that these same stable modes are also relevant to the coordination of multiple planning oscillators. Further, we hypothesize that for a system such as speech, which can be successfully acquired spontaneously (without explicit training) by most members of the population, intrinsically accessible modes are exploited as much as possible.

If a consonant (C) gesture and a vowel (V) gesture are to be coordinated in an intrinsic mode, there are just two possibilities: in-phase and anti-phase. We have hypothesized [13] that selecting in-phase mode produces the coordination underlying CV structures (C is a syllable onset), while selecting anti-phase produces VC structures (C is a syllable coda). Evidence for the in-phase coordination of onset C and V can be found in [19], which shows that onset C and V gestures are triggered in rough temporal synchrony. The example of "mad" in Figure 2 shows the onset gestures (lip closure and velum opening) coupled in-phase with the V gesture (wide pharyngeal constriction of the tongue body), and the coda gesture (tongue tip closure) coupled anti-phase with the V.

The coupling hypothesis constitutes a grounded (or embodied) theory of syllable structure. It explains the fact that there are syllables in language that have onsets or codas from the necessity of coordinating the timing of multiple gestures in one of two accessible modes. It can also explain macroscopic universal patterns associated with syllable structure, for example, why syllables with onsets (CV) are universal while those with codas are not [13]. This follows from the fact that the in-phase mode is more accessible and more stable. Similarly, it can account for the fact that onsets and Vs combine relatively freely, while combinations of V and coda Cs can be more restricted [13] and for the fact that onsets emerge earlier in phonological development [21].

2.2 Complex Onsets

A challenge to the modal theory of syllable structure can be found in syllables with consonant cluster onsets (e.g. "spade"). The oral constriction gestures associated with the /s/ and /p/ must be at least partially sequential with one another (otherwise we could not recover them both perceptually). Therefore, they cannot both be in-phase with the V gesture. If in-phase coordination defines syllable onsets, how can they both be part of the onset?

A possible solution to this [5, 21, 27] is that multiple, potentially competing couplings can be specified in the coupling graph. So for example in "spade," both onset Cs can be specified with in-phase links to the V, while they are also specified with an anti-phase link to each other, as shown in the graph in Figure 3. The coupled oscillator planning model can implement this competitive graph, and oscillators will settle at final phases that represent a stable compromise between the competing coupling forces.

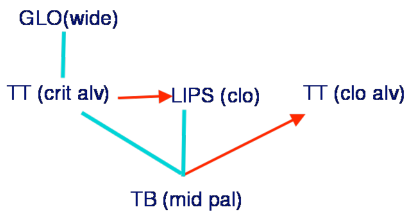

Figure 3: Coupling graph for "spade"

Evidence of this kind of competitive structure can be found in the relative timing of gestures in a complex onset [5]. In the word "spade," the /s/ is shifted earlier (leftward) with respect to the vowel, compared to its timing in words in which /s/ is the only onset C. Similarly, the timing of the /p/ in "spade" is shifted later (rightward), compared to its timing in words in which /p/ is the only onset C. These leftward and rightward shifts have also been dubbed the C-center effect [2,6], and they have been shown to emerge from competition in the planning model. Additional evidence for the coupling model has been found in the variability associated with C clusters. Model simulations [26, 27] have shown that loops in the coupling graph (e.g., the multiply-linked onset structure in "spade" in Fig. 3), help stabilize the final oscillator relative phases, and make them less sensitive to noise. This graph property can then explain the relative temporal stability of onset clusters [7], compared to coda clusters and heterosyllabic clusters, whose relative timing does not show the C-center effects exhibited by onsets [16].

While the coupling model of syllable structure has shown promising results, there is as yet only a very small amount of data demonstrating both leftward and rightward shifts in onset clusters. Moreover, the symmetry (or lack of it) between the leftward and rightward shifts has not been systematically evaluated. In the next section, we will report on an experiment designed to investigate the robustness and symmetry of these shifts. Some evidence for systematically asymmetrical shifts will be presented, and mechanisms for accounting for the asymmetries within the coupling model will be considered. In the final section, data on onset stop clusters in Georgian will be presented. These have been shown [10] to exhibit different lags as a function of the order of place in the cluster. We will consider how this kind of asymmetry can be modeled in the coupling model.

3. ENGLISH REVISITED

3.1 Method

X-ray microbeam data [20] data were collected from 6 subjects producing the utterances shown in Table 1. The utterances were designed to allow comparison of the C-V timing in words beginning with single Cs (/p/, /s/, /l/) with words beginning with CC clusters (/sp/, /pl/). Two of the six subjects were asked to produce the utterances in two different accent pattern conditions (A: accent on pa, B: accent on _eets). The other 4 subjects produced only accent pattern B. Gold markers were positioned in the mid-sagittal plane at the following locations: upper lip, lower lip, lower teeth, tongue tip (actually about 1 behind the tongue tip), and tongue dorsum (as far posterior as a subject was comfortable with) and 2 on the body of the tongue, approximately equally spaced between the tongue tip and dorsum markers. Each subject produced between 5 and 10 repetitions of each utterance.

Table 1: English Utterances

I read *pa **s**eats* again.
I read *pa **p**eets* again.
I read *pa **l**eets* again.
I read *pa **sp**eets* again.
I read *pa **pl**eats* again.

The kinematic data were used in conjunction with a trace of the subject's palate outline to estimate constriction time functions for the relevant consonants and the vowel /i/: LA (Lip Aperture)—the distance between the upper and lower lips, TTCD (Tongue Tip Constriction Degree)—the distance from the tongue tip marker to the palate, and TBCD (Tongue Body Constriction Degree)—the distance from the more posterior tongue body marker and the palate. Constriction gestures were located in the appropriate time functions (LA for /p/, TTCD for /s/ and /l/, TBCD for /i/). For each gesture, three points in time were detected using a velocity threshold—onset of constriction formation, achievement of constriction target, and release.

Mean relative timing of consonant gestures to the vowel /i/ for each subject were calculated using the lag between achievement of constriction of the C and /i/. For the subjects who produced two accent patterns, means were calculated separately for the two patterns. By comparing a given subject's mean lags for single Cs with CC clusters, we calculated the subject's leftward and rightward shifts associated with the clusters /sp/ and /pl/. For /sp/, leftward shift was calculated as the difference between /s/-/i/ timing in "seats" with /s/-/i/ timing in "speets." Rightward shift was calculated as the difference between /p/-/i/ timing in "peets" with /p/-/i/ timing in "speets." Similarly for /pl/, leftward shift was calculated as the difference between /p/-/i/ timing in "peets" with /p/-/i/ timing in "pleats." Rightward shift was calculated as the difference between /l/-/i/ timing in "leets" with /l-/i/ timing in "pleats."

3.2 Results

Figure 4 shows the leftward (light-colored bar) and rightward (dark bar) shifts for /sp/ and Figure 5 shows /pl/. The top pair of bars shows the mean shifts over all subjects, and the eight pairs below show the results for the individual subjects (separated by accent for two subjects—s1A, s1B, s2A, s2B).

In general, clusters exhibit both leftward and rightward shifts, supporting the multiply-linked onset (MLO) hypothesis. For /sp/, all subjects/accents show both leftward and rightward shifts, except for subject s1, accent pattern B, which fails to show a rightward shifts. The mean leftward shift is 47 ms, somewhat greater than the mean rightward shift is 25 ms. However, this left-right asymmetry is not consistent across speakers/accents. Five cases show a greater leftward than rightward shift, but three show the reverse. For /pl/, leftward and rightward shifts are found for all subjects except s2 (both accent conditions). The left-right asymmetry of the means is greater than for /sp/ (mean leftward = 68 ms., mean rightward =13 ms), and it is consistent across subjects. All 8 subjects/accents show a greater leftward shift than rightward shift.

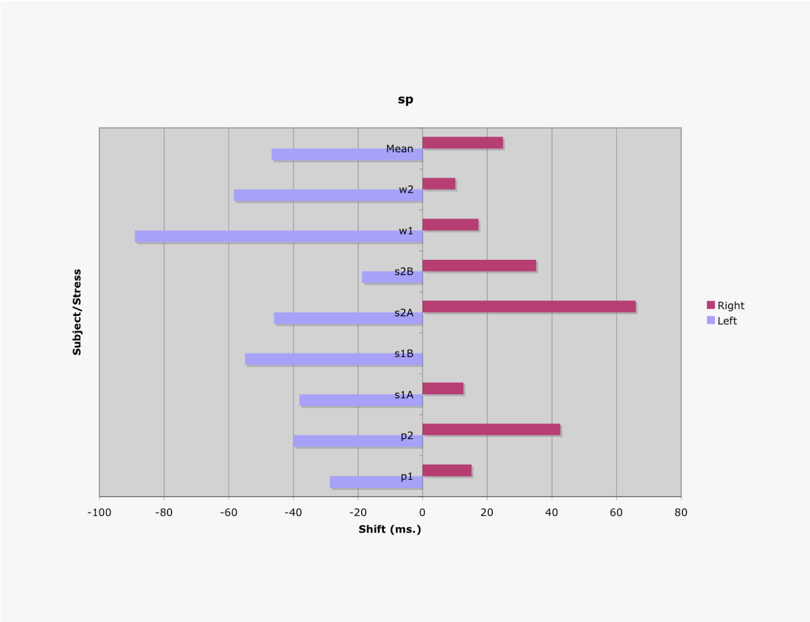

Figure 4: *Leftward and rightward shifts in /sp/*

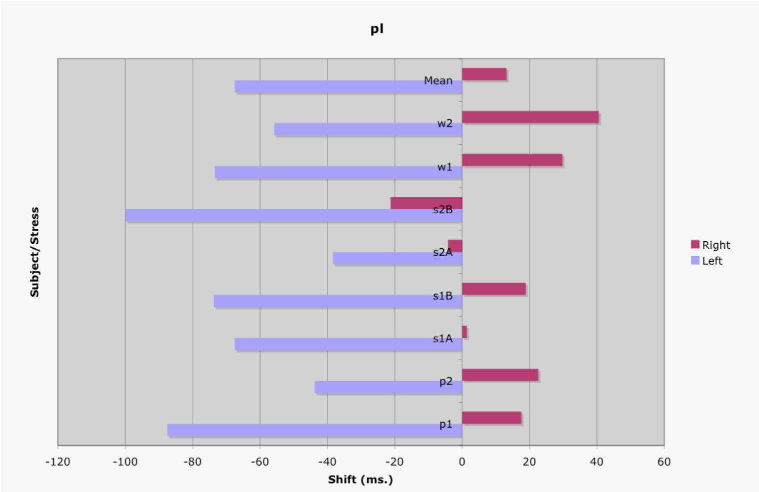

Figure 5: Leftward and rightward shifts in /pl/

How the difference in asymmetry between /sp/ and /pl/ clusters be accounted for? Since the leftward bias in the case of /pl/ appears to be quite systematic, it would be desirable to capture it in the topology of the coupling graph, ie., a different pattern of links (edges) between nodes for /sp/ and /pl/. One possible source of such a difference is that the /l/ is, in fact, composed of two gestures: a tongue tip constriction and a body constriction at the uvula [4,28]. It is possible that both of these gestures are coupled with the vowel. This is shown in Fig. 6, along with graph for /sp/. The extra link in the /pl/ graph with should result in a tighter coupling of the /l/ with respect to the vowel. This tighter coupling should also cause a reduction in rightward shift, as is confirmed in the quantitative modeling in the next section.

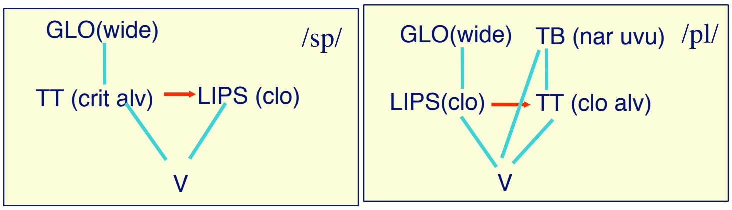

Figure 6: Hypothesized coupling graphs for /sp/ and /pl/

3.3. Quantitative modeling

While these results are in a general sense consistent with the coupling graph model that includes a competitive, multiply-linked graph of onset clusters, the details of these results are not consistent with simplest parameterization of the model that assumes that all coupling strengths are equal. That parameterization should predict an equal left and right shift. How can we model the weak left-right asymmetry exhibited by the /sp/ cluster, the much stronger leftward bias of /pl/, and an individual differences in both asymmetries?

Beginning with the weak leftward bias of /sp/, we reasoned that it would be possible to induce a leftward bias by reducing the coupling strength of C1 with V to a value less than 1, while keeping the C2-V and C1-C2 coupling strengths equal to 1. This should resolve the conflicting C-V and C-

C coupling in favor of keeping C2 relatively synchronous with V at the expense of C1-V coupling. To determine if the obtained results could be predicted by adjusting the coupling strengths, we varied the coupling strengths of all links in the graph from 0.1 to 1.0 in increments of 0.1. The resulting coupling graphs were input to TADA, the coupled oscillator simulation model [22], and the resulting gestural activations were calculated. Setting all coupling strengths to 1 resulted, as expected in symmetry, with leftward and rightward shifts equal to 30 ms. The data for the mean values of /sp/ were most closely modeled when C2-V and C1-C2 strengths remained equal to 1, and C1-V strength was reduced to 0.7 (yielding shifts left =40 ms, right=30 ms) or 0.6 (left=50 ms, right=20 ms). (Since the temporal resolution of the model is in terms of 10 ms frames, it is not possible to model the results exactly). Individual subject results can be approximated by reducing either C1-V or C2-V strengths.

For the /pl/ clusters, we manipulated coupling strength in the models with and without the TB link to the V. Coarser increments (0.135) were employed because there were more links to manipulate. Results for simulations without the additional link showed that reducing the coupling strength for C1-V to a value similar to that used to model /sp/ (0.74) while keeping all other strengths equal to 1 gave the same results as for /sp/ (shifts: left=40 ms, right=30 ms). However, when adding the additional in-phase link between the TD gesture for /l/ and the V, the shifts change to values that are close to those observed for /pl/ in our experiment: (left=60 ms, right=10 ms).

In summary, the English data support a view that patterns of coupling that seem to be regular aspects of the phonological knowledge of a language (in that they generalize across individual talkers) can be well represented qualitatively (discretely) in terms of the set of edges that define the coupling graph. Individual differences in coordination patterns can then be modeled by quantitative variation in the coupling strength values of the graph's edges. In this sense (though not in the formal sense of the mathematics of the models), coupling strength functions like articulator weights in the constriction formation task-dynamics [25], which can show individual variation in, for example, how much lip vs. jaw raising an individual talkers employs to produced a lip closure. These articulatory weights, like the coupling strengths, are assumed to not carry phonological information, which can be viewed as shared structure across community members.

4. GEORGIAN

Georgian is well known for allowing clusters of 2 or 3 stop consonants in onset position [30]. Recently [14] we found evidence that complex onset clusters in Georgian exhibit the competitive, multiply-linked structure hypothesized by the coupled oscillator model to be associated with syllable onsets. When examined using EMMA, the timing of the cluster-final C gesture with respect to the vowel shifted rightward when additional Cs were added to the onset in the following words: /riala/, /k'riala/, and /t^{s}k'riala/. This result contrasted with results using comparable words from Tashlhiyt Berber, a language in which words can also begin with complex sequences of Cs, but which are not syllabified as a complex onset, but rather as additional syllables with consonantal nuclei [11]. No evidence for the rightward shift was found in Berber. Therefore, it was argued that rightward shift could be used as a diagnostic for syllabification of consonant sequences as part of the onset.

Stop clusters in Georgian have been shown to exhibit different phonological [8, 30] and phonetic behaviors when they are sequenced such that more anterior constrictions precede more posterior constrictions ('front-to-back'), as opposed to the reverse ('back-to-front'). Most relevantly, kinematic analysis has revealed that the front-to-back sequences are produced with a shorter lag between the gesture onsets [9] than back-to-front sequences. While the earlier studies on this effect [10] hypothesized that the basis for this asymmetry could be found in the affordance of perceptual recoverability, more recent work both on Georgian [9] and other languages that also appear to show this effect (e.g., French [18]) casts some doubt on this account. But whatever the underlying cause of the evolution of this pattern, the planning of sequences in Georgian must produce different relative timing of controls for two gestures, depending on their sequential order. How can this be accomplished in a model in which coordination is controlled by means of only two intrinsically stable modes?

Since the lag differences between two place orders appears to generalize across speakers and to be phonologically relevant in Georgian, it would be desirable to derive them from topological differences in the coupling graph. One possibility involves participation in the coupling graph of the release gestures associated stop closures. Active release gestures have been shown to be necessary to account for the kinematic and perceptual properties of stops (and other consonants) [1], and to participate in some cases in planning by constituting nodes in the coupling graph [21]. However, for simplicity of exposition, the release gestures have been left out of the coupling graph figures in this paper so far (though they have been included in the actual simulations that these figures are based upon). Figure 7 (a) shows a possible coupling graph for an initial /bg/ (front-to-back) cluster in Georgian, based on the MLO principles considered so far, but with the release gestures displayed. The LIPS and TB closures are both coupled in-phase with the V, and anti-phase with each other. The release gestures are coupled only to their corresponding closures, so their presence does don't affect the relative timing of the other gestures in the graph.

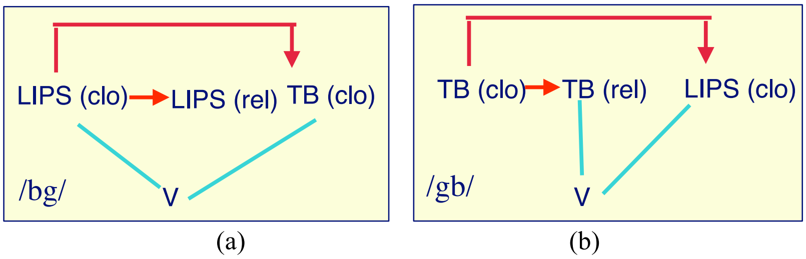

Figure 7: Hypothesized Georgian Coupling Graphs

This topology would predict a rightward shift of C2, as has been found for Georgian. If the same graph (with the opposite ordering of the two stops) were used for a /gb/ cluster, no lag differences between /bg/ and /gb/ would be predicted by such graphs. Figure 7(b) shows an alternative hypothesis for /gb/. The only difference from Figure 7 (a) is that it is now the release of C1, not its closure, that is coupled in-phase with the V. This change should, however, produce an increase in C1-C2 lag (since it is a later point in C1 that is now in-phase with the V). In fact, since C1 (clo) is now anti-phase with both C1 (rel) and C2 (clo), and both of these are in-phase with the V, the coupling specifications are all consistent—there is no competition. The final phasing of the oscillators that is expected to result from this graph should have C1 (rel) in-phase with both C2 (clo) and V, and C1 (clo) anti-phase with all of these.

Simulating these graph in 7(b) in TADA confirmed this phasing pattern. In the resulting gestural score, the lag between C1 (clo) and C2 (clo) was 90 ms. C1 (rel), C2 (clo) and V were all synchronous. This 90 ms lag can be compared to the C1 (clo) - C2 (clo) lag that results from the graph in Figure 7(a), which was 60 ms. This difference approximates closely the difference in lag reported in back-to-front (99 ms) vs. front-to-back (67 ms) in [14], averaged across the two speakers. Thus, if front-to-back stop sequences are controlled by graphs such as that in Figure 7(a), and back-to-front sequences by graphs such as in Figure 7(b), then the contrasting values in these sequences can be well accounted for, while retaining the theoretically desirable property that both onset Cs are coupled to V.

One prediction of the coupling graph in Figure 7(b) is that there is no rightward shift. In the output from this graph, C2 is synchronous with V, just as it would be if it were the only C in the onset. Here we test this prediction by examining evidence for rightward shift in Georgian stop sequences with front-to-back vs. back-to-front order.

Method. The words in Table 2 were recorded by two speakers of Georgian, the same two who participated in [14]. Articulatory kinematics were measured using EMMA, with markers placed similarly to those described for the English study in the preceding section, except that only one marker was placed on the tongue body, instead of the two employed in English. Words were recorded in a frame (**Sit'q'va __ gamoit^{h}k^{h}mis ordʒer**), and between 6 and 12 repetitions of each word type were acquired. A Lip Aperture (LA) time function was computed as the distance from the upper lip marker to the lower lip marker, and Tongue Tip (TTCD) and Tongue Body Constriction Degree (TBCD) time functions were computed as distances from the relevant marker from the palate outline. Gesture onsets, achievement of constriction target, and release were measured using the following time functions: /b/ (LA), t^{s}' (TTCD), k' (TDy), Vowels (TBCD). The time from the constriction achievement of the C gestures to the syllable's V gesture was calculated.

Table 2: Georgian words for rightward shift

Back-to-front	**Front-to-back**
bil-i 'bill'	k'ar-eb-I 'door' (pl.)
k'bil-i 'tooth'	t^{s}'k'ar-eb-i (nons.)

Results. For the front-to-back cluster, a 19 ms rightward shift was obtained (time from /k'/ to vowel was 133 ms in /k'/ and 114 in /t^{s}'k'/). The direction of shift was consistent across the two subjects. This indicates that a graph like that in Figure 7 (a) is appropriate (with an initial TT gesture substituting for the initial LIPS gesture). However, for the back-to-front example, the mean shift was only 7 ms (146 ms in /k'/ and 139 ms in /k'b/) and was not consistent in direction across subjects. This result would be predicted by a coupling graph like that in Figure 7(b) (with the addition of a glottal closure gesture).

Results from Georgian are consistent with the hypothesis that front-to-back and back-to-front clusters are represented using the different coupling graphs shown in Figure 7. These graphs satisfy the theoretical goal of using topological differences to capture robust timing effects that generalize across subjects in a language, while also satisfying the theoretical goal of capturing Georgian speakers' intuitions that even the back-to-front clusters are well-formed syllable onsets. The syllable onset intuition can be grounded in coupling graphs in which all onset Cs exhibit some in-phase coupling with the V. For the back-to-front order, it is the release, not the closure, of the initial C that bears that relation to the V. While the mean temporal pattern predicted by that graph might not differ from a graph in which C_1 is totally uncoupled from the V, different patterns of variability would be predicted. The variability predictions of such graphs could, in principle, be tested in a language in which some CCV sequences are analyzed as complex onsets and others are analyzed as an "extra-syllabic" C followed by a CV syllable.

Finally, the coupling graph differences between 7(a) and 7(b) can also be used to model other examples where different types of onset clusters have been shown to exhibit systematically different patterns of lag. The lag differences between the German initial /gl/ and /gn/ clusters as reported in [17] can be modeled in this way. Such a model also predicts that /gl/ should exhibit the rightward shift, but /gn/ should not. It is not yet known whether that prediction is born out.

5. ACKNOWLEDGMENTS

Thanks to Aaron Jacobs for help with the analysis. The work was supported by NIH grants DC008780 and DC03172 and NSF grant 0703048.

6. REFERENCES

[1] Browman, Catherine P. 1994 Lip aperture and consonant releases. *Papers in laboratory phonology III: Phonological structure and phonetic Form*. P. Keating (ed.), 331-353. Cambridge: Cambridge University Press.

[2] Browman, C., & Goldstein, L. 1988. Some notes on syllable structure in Articulatory Phonology. *Phonetica, 45*, 140-155

[3] Browman, C. P., & Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica, 49*, 155-180.

[4] Browman, C.P. & Goldstein, L. 1995. Gestural syllable position effects in American English. In F. Bell-Berti and L. Raphael (eds.). *Studies in Speech Production: A Festschrift for Katherine Safford Harris.* (pp. 19-34).Woodbury NY: American Institute of Physics.

[5] Browman, C., & Goldstein, L. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée, 5*, 25-34.

[6] Byrd, D. 1995. C-centers revisited. *Phonetica, 52*, 285-306.

[7] Byrd, D., 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics 24*, 209-244.

[8] Chitoran, I. 1998. Georgian harmonic cluster: phonetic cues to phonological representation. *Phonology, 15,* 121-144.

[9] Chitoran, I. & Goldstein, L. 2006. Testing the phonological status of perceptual recoverability: Articulatory evidence from Georgian. Presented at Laboratory Phonology 10, Paris.

[10] Chitoran, I., Goldstein, L. & Byrd, D., 2002. Gestural overlap and recoverability: Articulatory evidence from Georgian. In Laboratory Phonology VII. (C. Gussenhoven and N. Warner, editors), pp. 419-448. Berlin: Walter de Gruyter.

[11] Dell, F., Elmedlaoui, M., 1985. Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber, *Journal of African Languages and Linguistics, 7*, 105-130

[12] Fujimura, O. 2000. The C/D Model and Prosodic Control of Articulatory Behavior. *Phonetica, 57*, 128-138.

[13] Goldstein, L., Byrd, D., & Saltzman, E. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib, (Ed.). *Action to Language via the Mirror Neuron System.* New York: Cambridge University Press, Pp. 215-249.

[14] Goldstein, L., Chitoran, I, & Selkirk, E., 2007. Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. *Proc. XVIth Int. Cong. Phonetic Sciences*, Saarbrücken, 241-244.

[15] Haken, H., Kelso, J. A. S., & Bunz, H. 1985. A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347-356.

[16] Honorof, D., & Browman, C. 1995. The center or edge: How are consonant clusters organized with respect to the vowel? In K. Elenius and P. Branderud (Eds.), Proceedings of the XIIIth International Congress of Phonetic Sciences, vol. 3. Stockholm: KTH and Stockholm University. Pp. 552-555.

[17] Hoole, P., Bombien, L., Kühnert, B. & Mooshammer, C. (2008) Intrinsic and Prosodic Effects on Articulatory Coordination in Initial Consonant Clusters. Paper presented at Phonetic Congress of China. Beijing.

[18] Kühnert, B; Hoole, P.; Mooshammer, C., 2006. Gestural overlap and C-center in selected French consonant clusters. *Proc. 7th Int. Seminar on Speech Production,* Ubatuba, 327-334.

[19] Löfqvist, A., Gracco, V. 1999. Interarticulator programming in VCV sequences: lip and tongue movements. *Journal of the Acoustic Society of America 105*, 1854-1876.

[20] Nadler, R. D., Abbs, J. H., & Fujimura, O. 1987. Speech movement research using the new x-ray microbeam system. Speech Motor Control Laboratories Preprints, 181-184.

[21] Nam, H. (in press). A competitive, coupled oscillator model of moraic structure: Split-gesture dynamics focusing on positional asymmetry. In J. Cole & J. Hualde (Eds.). Papers in Laboratory Phonology IX.

[22] Nam, H., Goldstein, L. & Proctor, M. (2007). TADA (TAsk Dynamics Application) Available for downloading `http://www.haskins.yale.edu/tada_download/`.

[23] Nam, H., & Saltzman, E. 2003. A competitive, coupled oscillator of syllable structure. *Proc. XVth International Congress of Phonetic Sciences, Barcelona*, 3-9 Aug 2003.

[24] Saltzman, E., & Byrd, D. 2000. Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science, 19*, 499-526.

[25] Saltzman, E. L., & Munhall, K. G. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*, 333-382.

[26] Saltzman, E., Nam, H., & Goldstein, L. submitted. Intergestural timing in speech production: The role of graph structure. *Human Movement Science.*
[27] Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. (in press). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa, S. Madureira, & C. Reis, (Eds.).(2008). *Proceedings of the Speech Prosody 2008 Conference,* Campinas, Brazil.
[28] Sproat, R. & Fujimura, O. 1993. Allophonic variation in American English /l/ and its implications for phonetic interpretation. *Journal of Phonetics, 21*, 391-311.
[29] Turvey, M. T. 1990. Coordination. *American Psychologist, 4*5, 938-953.
[30] Vogt, H. 1971. *Grammaire de la langue géorgienne.* Oslo: Universitetsforlaget.