A Competitive, Coupled Oscillator Model of Syllable Structure*

Hosung Nam[†], Elliot Saltzman[‡]
† Yale University, USA
‡ Boston University, Haskins Laboratory, USA
E-mail: hosung.nam@yale.edu, esaltz@bu.edu

ABSTRACT

Syllable-initial consonant sequences have been shown to differ from syllable-final ones in two ways: they show the c-center effect [3] and they exhibit greater stability in timing [5]. Here we show that both of these differences emerge automatically from a model in which phonetic units (gestures) are modeled as oscillators, and coordination is achieved by coupling. Significantly, the necessary coupling relations are abstract and involve competitive, mutually incompatible specifications.

1. INTRODUCTION

Recently, nonlinear dynamic models (e.g., [9]) have provided a possible basis for understanding the temporal regularities exhibited by action systems. For example, in discrete, nonrepetitive utterances, articulatory actions (or gestures) can be viewed as oscillating systems with critical damping [2]. Further, intergestural coordination of multiple articulatory actions can be interpreted in terms of interaction of multiple oscillatory systems. If this interpretation is appropriate, coupled oscillator models may provide significant insights regarding the qualitative (and quantitative) behavior of the gestural organization of speed.

Saltzman & Byrd [7] developed a model in which the relative timing of a pair of gestural actions can be controlled by coupling the dynamical systems (oscillators) hypothesized to underlie the two actions. Depending on the form of the coupling and its parameterization, different relative phases (or ranges of phases) result.

In their model, an initial relative phase, $\psi_{T=0}$, changes over time toward a target relative phase, ψ_0, due to the application of a coupling force derived from a potential energy function $V(\psi)$ to each of two uncoupled oscillator systems (equations) until the relative phase (ψ) arrives at the target defined at the bottom of the potential energy curve, V. As given in (1), V is a function of relative phase, ψ ($\psi = \phi_2 - \phi_1$, ϕ_1, ϕ_2 are phases of two oscillators at a given time). Consequently, if the oscillators are parameterized identically, wherever the initial relative phase starts, the final relative phase reaches the minimum energy position, where phase locking is completed, on the potential curve of relative phase, $V(\psi)$.

$$V(\psi) = -a\cos(\psi - \psi_0) \qquad (1)$$

Here we extend their model so as to allow multiple, potentially competing coupling relations and show that by hypothesizing differences in coupling structure (i.e., the graph of which gestures are coupled to which others) between onset and coda consonants, empirical differences in *both* the mean relative timing of these clusters and their variability can be predicted.

2. C-CENTER AND COMPETING GESTURAL COORDINATION

Browman & Goldstein [3] argued that the type of gestural coordination of consonants observed in syllable onsets, called the c-center effect [1, 4], could be accounted for by local specifications of pairwise relative phases between gestures if multiple competing relative phases are specified. For onset clusters, they hypothesized that each consonant gesture was phased identically to the vowel gesture (C-V phasing) (e.g. in (2), the lines connecting C_1 and C_2 to V represent identical phase specifications). If these were the only phase specifications, the consonant gestures would be produced synchronously. However, they hypothesized that the consonant gestures are also directly phased to each other (C-C phasing) as in (2), so as to separate them in time. Thus, the C-V coupling and C-C coupling are in competition.

$$\# \; C_1 - C_2 - V \qquad (2)$$

The result of this competition is the c-center effect, in which adding additional consonants to an onset changes the resultant phase of all consonant gestures with respect to the vowel in a way that preserves the overall timing of the *center* of the consonant sequence with respect to the vowel. In coda, Browman & Goldstein hypothesized that only the first consonant is phased to the vowel (V-C phasing), and the consonants are phased to each other as in (3). Here there is no competition and no c-center effect.

$$V - C_1 - C_2 \; \# \qquad (3)$$

Here, we will test this idea by extending the Saltzman & Byrd model to allow multiple gestures and competitive couplings. It is predicted that the hypothesized intergestural

---

* Work supported by NIH grant DC-0663 to Haskins Laboratories. Louis Goldstein should also be considered an author of this paper.

structure of the onset cluster will result in a self-organized c-centered form that arises as the result of multiple competitive coupling forces, (simulation Experiment I).

Experiment 1: C-center

The original coupled oscillator model of Saltzman & Byrd [7] was extended using Matlab & Simulink. By extending their two-oscillator model to a three-oscillator model, we examined whether the c-center effect emerges from either of the coordination graphs (2, 3). Any two oscillators with a linking line in (2, 3) are in a coupling relation. For an onset simulation (2), desired target relative phases are set to 50°, 50° and 30° for C_1-V, C_2-V and C_1-C_2 respectively. Coupling strength was controlled by tuning the slope (*a*) of potential energy function in (1). Here in C_1C_2V, all *a* of C_1-V, C_2-V and C_1-C_2 couplings are set to 1 so they are all evenly competing. Initial phases of the three oscillators are set to 0°, -140° and 50° respectively. By contrast, for coda (3), only V-C_1 and C_1-C_2, (but not V-C_2) are specified. Target values are set at 50° and 30° respectively. Each component oscillator was governed by the following equation:

$$\ddot{x}_i = -\alpha_i \dot{x}_i - \beta_i x_i^{\,2} \dot{x}_i - \gamma_i \dot{x}_i^{\,3} - \omega_{0i}^{\,2} x_i ,$$

where the parameter values were $\omega_{0i} = \beta_i = \gamma_i = 1,\ \ \alpha_i = -1$ (*i* denotes oscillator identity).

For all the simulations, we set the time span to 200 sec. Each relative phase reached equilibrium state by approximately 10 seconds (The stabilizing time can be shortened by tuning up *a* in (1)). We measured mean of relative phases from 20 sec. to 200 sec.: C_1-V = 59.94°, C_2-V = 39.96°, C_1-C_2 = 19.98° for (2) whereas V-C_1 = 49.96°, C_1-C_2 = 29.94°, V-C_2 = 79.90° for (3) (See Figure 1).

The competitive couplings in onset result in preserving the overall timing of the *center* of the consonant sequence with respect to the vowel. That is the c-center effect. A schematic diagram of gestural coordination for c-center effect in onset position is given in Figure 1. The top two panels represent the target relative phases. The left one is for C-V phase coupling (50°) and the right one for C-C phase coupling (30°). These competitive temporal relations dynamically result in c-centered structure as in left bottom: the mean relative phase of the consonant cluster maintains the target relative phase (50°) phased to the vowel despite the failure of each consonant to sustain the target relative phase (50°) with respect to the vowel due to the competition. But in coda, there is no c-center effect. Each of two targeted phase relations is achieved and the resultant relative phase of V-C_2 is the simple sum of V-C_1 and C_1-C_2. Thus, the resultant relative phase between the vowel and the mean phase of the consonant sequence is 65°, which is different from the target relative phase (50°).

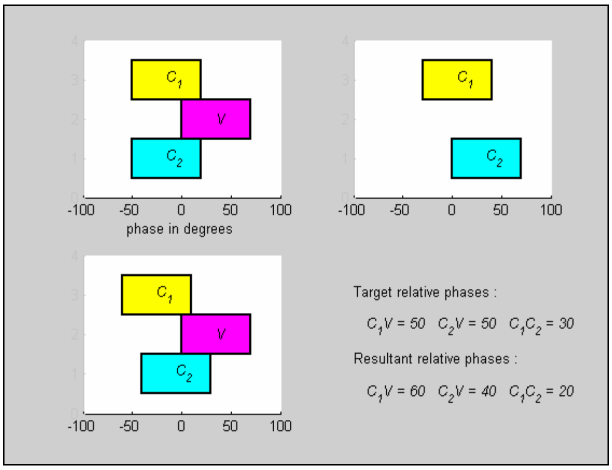

Figure 1 Schematic diagram of gestural timing for c-center effect in onset

3. STABILITY

We also hypothesized that such competitive coupling structures would account for asymmetries of relative phasing stability between onsets and codas [5]. According to Browman & Goldstein [3],

"… This may help explain the greater stability in overlap that has observed for onset consonants than for coda consonants. There is no reliable evidence that coda consonants show the c-center effect (evidence is negative in some studies or variable in others). Thus, final consonants may not be attracted into simultaneity by a V-C relation (parallel to the C-V relation)… "

In fact, reduced variability of C_1-C_2 relative phase in onset position also emerged from same coupling structure used for the c-center simulation: three inter-oscillatory couplings (C_1-C_2, C_2-V and C_1-V) in #C_1C_2V vs. two (V-C_1 and C_1-C_2) in VC_1C_2#.

In the simulation experiments described below, we performed 200 simulations trials of each utterance type, adding a random amount of noise to the coupling potential function on each trial. The result showed that relative phase of C_1-C_2 in onset exhibits less variability than in coda (Experiment 2). This idea was extended to 4-oscillator structures with word boundaries (Experiment 3): V_1#$C_1C_2V_2$, V_1C_1#C_2V_2, $V_1C_1C_2$#V_2, and to Korean data with and without a morpheme boundary provided by Cho [6] (Experiment 4).

Experiment 2: Variability in onset and codas

We added a noise term to the potential energy function to induce variability across simulation trials. One way of adding noise (*N*) to the target relative phase, ψ_0 would be:

$$V(\psi) = -a \cos(\psi - \psi_0 + N) \qquad (4)$$

However, this would only have the effect of changing the target relative phase (ψ_0) by N from trial to trial. This variable target will, of course be achieved. It thus was more interesting to simulate the situation in which we assume that a speaker is employing an invariant target across trials (and this target could be a categorical primitive), but that noise arises from some other property of the system. One such source of noise could be slight changes in the frequencies of the individual oscillators. Such changes might represent small changes in effective speaking rate from trial to trial. If the frequencies of the two oscillators change by different amounts, then they will be correspondingly de-tuned. The effect of detuning on the potential function for a set of coupled oscillators can be modeled by adding a linear function to the potential energy function as in (5) (e.g., [8]). The value of *b* represents the amount of detuning.

$$V(\psi) = -a\cos(\psi - \psi_0) + b(\psi - \psi_0) \qquad (5)$$

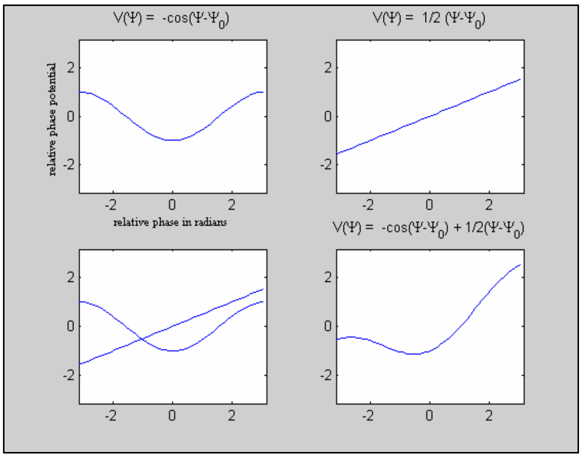

Figure 2 Adding noise by adding de-tuning (a = 1, ψ_0 = 0°, b = 1/2)

The added linear term $b(\psi - \psi_0)$ makes the potential parabola (*V*) tilted, which shifts the lowest energy point slightly. For example, by adding de-tuning (a non-zero value of b) as in the bottom right panel of Figure 2, a new "target" relative phase (< 0°) is created that diverges from the old target relative phase (= 0°).

For the simulations, syllable-initial coupling structure is given in (2) and syllable-final coupling structure in (3). The target relative phases of C_1-C_2, C_2-V and C_1-V in C_1C_2V are all set to 30°. By contrast, for VC_1C_2, V and C_2 are not coupled by setting *a* for V-C_2 to 0. For each of 200 repetitions of a given utterance, we set the value of *b* (the de-tuning factor) to a random variable, whose mean was zero, and its standard deviation σ. The value of σ (the amount of detuning noise) was manipulated in a series of five noise conditions, increasing from .05 to .85 in .20 increments. We set simulation time span to 20 sec.

We measured standard deviation of final relative phase between C_1 and C_2 across the 200 repetitions for each utterance type within each noise condition. Comparison of variability between onset and coda consonant clusters is given in Table 1 and Figure 3. Onset consonant clusters showed lower standard deviation over repetitions, which means greater stability, across the range of random noise levels (standard deviation of *b*: .05 ~ .85). A dotted line of unit slope is included in the figure to make differences in the shape of the functions more visible.

σ	**Onsets** # $\mathbf{C_1C_2}$V (2)	**Codas** V$\mathbf{C_1C_2}$ # (3)
.05	0.0394	0.0525
.25	0.2144	0.2391
.45	0.4020	0.5047
.65	0.7426	0.8888
.85	0.9374	1.0862

Table 1 Phase variability (radians) of C-C in CCV (left) and VCC (right)

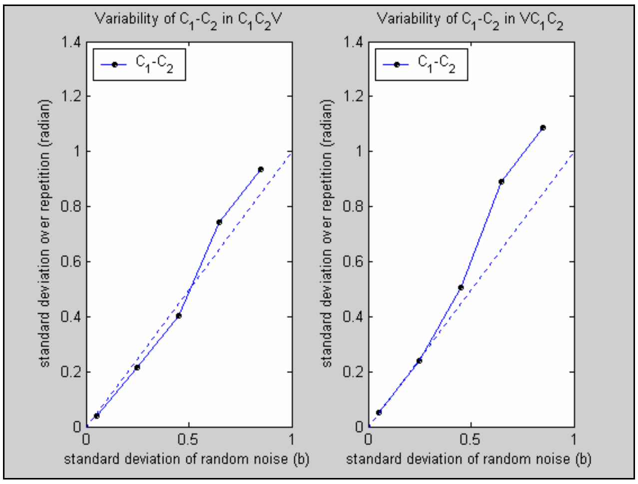

Figure 3 Phase variability of C-C in CCV (left) and VCC (right)

Experiment 3: Word boundaries

We extended our simulation to gestures defined across word (or syllable) boundaries using four oscillators. (6), (7) and (8) are possible coupling structures for utterances that span a boundary. The same principles and settings were employed here as in Experiment 2, except consonants are not directly coupled and vowels are directly linked or coupled across a word boundary.

V # C—C—V (6)

V—C # C—V (7)

V—C—C # V (8)

Comparison of variability between onset, coda and cross-boundary consonant clusters is given in Table 2. Cross-boundary clusters showed the greatest variability and the onset cluster the smallest variability for each noise level (standard deviation of *b*: .05 ~ .85), which are supported by empirical data found in Byrd [5].

σ	**Onsets** $\mathbf{V_1}$#$C_1C_2\mathbf{V_2}$ **(5)**	**X-boundary** $\mathbf{V_1}C_1$#$C_2\mathbf{V_2}$ **(6)**	**Codas** $\mathbf{V_1}C_1C_2$#$\mathbf{V_2}$ **(7)**
.05	0.0422	0.0870	0.0502
.25	0.1942	0.4330	0.2433
.45	0.4460	0.8693	0.5227
.65	0.6768	1.2608	0.7970
.85	1.0047	1.4719	1.0801

Table 2 Phase variability (radians) of C-C in (6), (7) and (8)

Experiment 4: Morphological structure

The coupling model can also account for the differences in timing variability associated with Korean morphological structure reported by Cho [6]. Monomorphemic /pi/ showed less variability than heteromorphemic /pi/. This was modeled by the presence or absence of a C-V coupling link as in (9, 10). Consonant sequences in lexicalized compounds were shown to have less variability than in non-lexicalized compounds. This was modeled by the presence or absence of a C-C link between the consonants (11, 12).

/na**<u>pi</u>**/: (Monomorphemic /pi/) $V_1 — C — V_2$ (9)

/na**<u>p-i</u>**/: (Heteromorphemic /p-i/) $V_2 — C \; \# \; V_2$ (10)

/ha**<u>kp</u>**i/: (Lexicalized Compound) $V_1 - C_1 - C_2 - V_2$ (11)

/ha**<u>k-p</u>**i/: (Non-Lex. Compound) $V_1 - C_1 \; \# \; C_2 - V_2$ (12)

C-V in (9) and C-C in (11) show less variability than in (10) and (12), respectively, as in Table 3.

σ	**Mono (9)** $V_1\mathbf{CV_2}$	**Hetero (10)** $V_1\mathbf{C}$#$\mathbf{V_2}$	**Lexical (11)** $V_1\mathbf{C_1C_2}V_2$	**Non-Lex. (12)** $V_1\mathbf{C_1}$#$\mathbf{C_2}V_2$
.05	0.0477	0.0870	0.0525	0.0710
.25	0.2553	0.4330	0.2391	0.3599
.45	0.5473	0.8693	0.5047	0.7212
.65	0.8727	1.2608	0.8888	1.2319
.85	1.0253	1.4719	1.0862	1.4670

Table 3 Phase variability (radians) of C-V and C-C with(out) a morpheme boundary

4. Implications and Prospects

This paper has provided a possible account of the asymmetrical phonetic realizations between syllable initial and final positions. The simulation experiments showed that different coupling graphs embedded within nonlinear dynamics can predict the asymmetries. Both the greater temporal stability of syllable-initial consonant sequences and the c-center effect emerge from competitive couplings between C-V and C-C that characterize onsets. In addition, the oscillator model can correctly predict the effect of boundaries on phonetic variation (temporal variability), by hypothesizing that gestures are not coupled directly across certain types of boundaries. The implication of this study is that variant phonetic realizations in different contexts can be predicted by embedding traditional linguistic boundaries in dynamical terms by means of gestural coupling graphs. We expect future studies to explore various levels of linguistic boundaries in a variety of languages, extending the current model.

REFERENCES

[1] C. Browman and L. Goldstein, “Some notes on syllable structure in articulatory phonology,” *Phonetica,* vol. 45, pp.140-155, 1988.

[2] C. Browman and L. Goldstein, “Tiers in articulatory phonology, with some implications for casual speech,” *Papers in laboratory phonology I: Between the grammar and physics of speech*, pp.341-376, Cambridge University Press, 1990.

[3] C. Browman and L. Goldstein, “Competing constraints on intergestural coordination and self-organization of phonological structures,” *Les Cahiers de l'ICP, Bulletin de la Communication Parlée*, vol. 5, pp.25-34, 2000.

[4] D. Byrd, “C-centers revisited,” *Phonetica,* vol. 52, pp.285-306, 1995.

[5] D. Byrd, “Influences on articulatory timing in consonant sequences,” *Journal of Phonetics,* vol. 24, pp.209-244, 1996.

[6] T. Cho, “Effects of Morpheme Boundaries on Intergestural Timing: Evidence from Korean,” *Phonetica*, vol. 58, pp.129-162, 2001.

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

[8] D. Sternad, E. Amazeen and M. Turvey, “Diffusive, Synaptic, and Synergetic coupling: An Evaluation Through In-Phase and Antiphase Rhythmic Movements,” *Journal of Motor Behavior*, vol. 28, no. 3, pp.255-269, 1996.

[9] M. Turvey, “Coordination,” *American Psychologist*, vol. 45, pp.938-953, 1990.