# Common coordination patterns in prosodic and segmental domains

**Jason A. Shaw**
*Yale University*
Jason.Shaw@yale.edu

In Articulatory Phonology, phonological representations take the form of temporally coordinated action units, known as gestures [1, 2]. Frequently observed coordination relations include synchronous (or in-phase) timing, whereby gestures start at the same time, and sequential (or anti-phase) timing, whereby the onset of one gesture is coordinated to the offset of another gesture [3]. In contrast to these coordination relations observed in the segmental domain, the pitch accents of intonation phonology have been argued to have target-based timing [4, 5]. On this view, pitch peaks (or troughs) of accentual tones, rather than the start or end of pitch control, are aligned to segments. Accordingly, pitch contours are derived from interpolation between L(ow) and H(igh) tone specifications, which are aligned to segments according to language-specific rules. It appears from this past work that segmental timing differs from prosodic timing in that segmental gestures (consonants and vowels) are coordinated with reference to gesture onsets and offsets while pitch accents (tones) are coordinated with respect to their targets (F0 maxima/minima). There has been some work trying to reconcile these differences, focusing on pitch patterns that can be modelled as tone gestures aligned according onsets (instead of pitch targets) [6]. In this talk, I argue based on data from Mandarin Chinese that consonant and vowel gestures can also show target alignment, offering another route to convergence between timing control in prosodic and segmental domains.

The key evidence comes from analysis of Electromagnetic Articulography (EMA) data from six speakers (three male) of Mandarin Chinese. Data was collected using the NDI Wave Speech production system. Sensors were attached to the tongue tip (TT), blade (TB), dorsum (TD), lips, jaw, nasion and mastoids. Lip Aperture (LA) was computed as the difference between the upper and lower lip sensors. Target items were a set of CV monosyllables that crossed all four lexical tones with two labial consonants {/m/,/p/} and three back rounded vowels {/o/,/u/,/ou/} yielding 24 items, which were repeated 6-12 times by each speaker producing a corpus of 949 tokens for analysis. Items were randomized with fillers and displayed one at a time on a monitor in pinyin.

To investigate the alignment of the vowel gesture relative to the consonant, a series of gesture landmarks were parsed from continuous kinematic trajectories: (1) **onset** of movement, (2) achievement of **target** (3) the **release** from target and (4) the **offset** of controlled movement. The alignment of consonants and vowels was assessed by computing C-V lag measures, subtracting consonantal landmarks from vowel landmarks. Fig. 1 compares three such lag measures. On average, the achievement of target of the vowel occurs around the offset of the consonant (right).

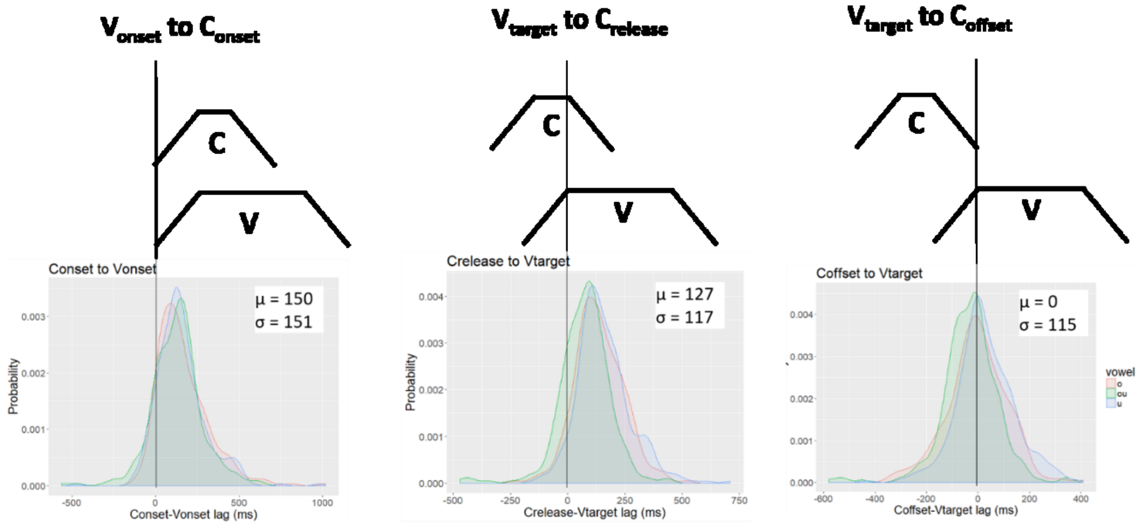


**Fig.1** temporal lag between three sets of C-V landmarks: (left) $V_{onset}$ to $C_{onset}$; (middle) $V_{target}$ to $C_{release}$; (right) $V_{target}$ to $C_{offset}$. The top row shows a schema for the lag measurement. The bottom row shows the distribution of lag values by vowel. The average lag between the $V_{target}$ and $C_{offset}$ is 0, indicating that these landmarks are roughly synchronous.

In contrast to the in-phase C-V pattern sometimes assumed for Mandarin Chinese [7], we found that the vowel gesture typically begins well after the consonant gesture; rather, it appears that the *vowel target* is more stably timed to the consonant than the *vowel onset*. Target-based timing makes an additional prediction that can be tested in the data. If the vowel target is the landmark that is temporally coordinated it would mean that the timing of the gesture onset is free to vary. Movement towards the target could start earlier or later in time depending, for example, on the distance of the tongue dorsum to its target. To investigate this possibility, we evaluated whether the lag between the consonant onset and vowel onset (leftmost schema in Fig. 1) is correlated with the spatial position of the tongue dorsum at movement onset. Fig 2 shows this relation. There is a significant negative correlation. The further the tongue dorsum is from the target, the earlier in time the vowel begins its movement. This pattern of spatially conditioned relative timing is not expected if gestures are timed according to their onsets and offsets [8]; however, the pattern is analagous to the finding in the intontation literature that the slope of F0 between pitch accents decreases with the number of syllables [9] or with speech rate [4].

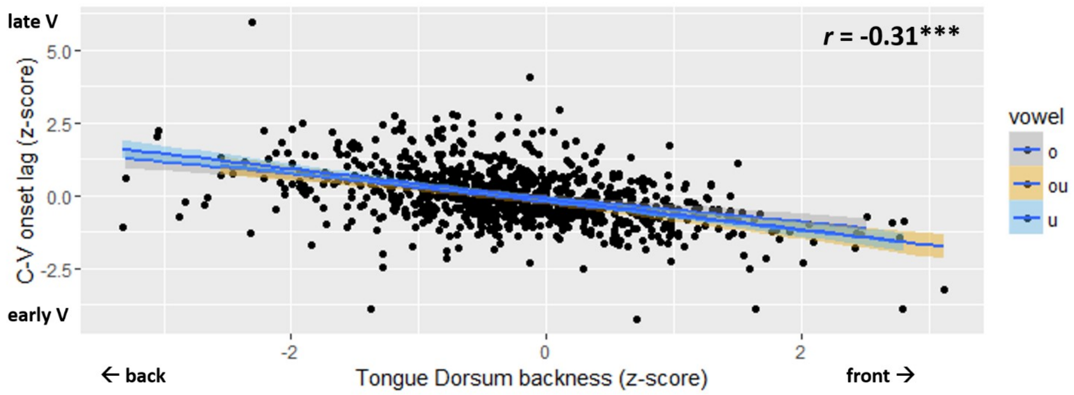


**Fig.2** scatter plot of C-V lag (y-axis) and Tongue Dorsum backness (x-axis). The legend shows the Pinyin for the vowels, which correspond to: ‘o’ /uo/, ‘u’ /u/, ‘ou’ /ou/

Overall, the data suggest that C-V coordination in Mandarin is based on the alignment of the vowel target to the offset of the consonant. This pattern is in contrast to other languages which can be modelled with gestures timed either in-phase or anti-phase [3]. Target-based vowel alignment resembles tone alignment in the Autosegmental Metrical framework, offering a point of convergence between timing control in the segmental and prosodic domains.

References

[1] A. Gafos, and L. Goldstein, "Articulatory representation and organization," *The Oxford handbook of laboratory phonology*, A. C. Cohn, C. Fougeron and M. K. Huffman, eds., pp. 220-231, 2012.

[2] C. Browman, and L. Goldstein, “Towards an Articulatory Phonology,” *Phonology Yearbook,* vol. 3, pp. 219-252, 1986.

[3] H. Nam, L. Goldstein, and E. Saltzman, "Self-organization of syllable structure: a coupled oscillator mode," *Approaches to phonological complexity*, F. Pellegrino, E. Marisco and I. Chitoran, eds., pp. 299-328, Berlin, New York: Mouton de Gruyter, 2009.

[4] D. R. Ladd, D. Faulkner, H. Faulkner, and A. Schepman, “Constant “segmental anchoring” of F 0 movements under changes in speech rate,” *The Journal of the Acoustical Society of America,* vol. 106, no. 3, pp. 1543-1554, 1999.

[5] A. Arvaniti, D. R. Ladd, and I. Mennen, “Stability of tonal alignment: the case of Greek prenuclear accents,” *Journal of phonetics,* vol. 26, no. 1, pp. 3-25, 1998.

[6] A. Katsika, J. Krivokapić, C. Mooshammer, M. Tiede, and L. Goldstein, “The coordination of boundary tones and its interaction with prominence,” *Journal of Phonetics,* vol. 44, pp. 62-82, 2014.

[7] Y. Xu, and M. Wang, “Organizing syllables into groups—Evidence from F< sub> 0</sub> and duration patterns in Mandarin,” *Journal of phonetics,* vol. 37, no. 4, pp. 502-520, 2009.

[8] J. A. Shaw, and W.-r. Chen, “Spatially-conditioned speech timing: evidence and implications,” *Frontiers in Psychology: language science*, pp. 1-17, submitted.

[9] M. Beckman, and J. Pierrehumbert, “Intonational structure in English and Japanese,” *Phonology,* vol. 3, pp. 255-309, 1986.