Variation in the spatial position of articulators influences the relative timing between consonants and vowels: evidence from CV timing in Mandarin Chinese
Jason A. Shaw & Wei-rong Chen
Yale University

Synopsis: Patterns of relative timing between consonants and vowels appear to be conditioned in part by abstract phonological structure, such as syllables, but also modulated by the particular gestures being coordinated (e.g., Hermes, Mücke, & Auris, 2017). The most rigorous attempts to formalize phonologically relevant temporal patterns have come within the Articulatory Phonology (AP) framework, which draws a distinction between the inter-gestural level of representation and the inter-articulator level (Browman & Goldstein, 1989; Saltzman & Munhall, 1989). A key implication of this two-level (feedforward) architecture is that temporal organization at the inter-gesture level dictates relative timing and is blind to the spatial position of articulators. Some recent results suggest possible links between the spatial position of articulators and relative timing (Brunner, Geng, Sotiropoulou, & Gafos, 2014; Pastätter & Pouplier, 2017). We provide a direct test of this link, investing whether variation in the spatial position of the tongue influences consonant-vowel coordination.

Data: We recorded Electromagnetic Articulography (EMA) data from six speakers (three male) of Mandarin Chinese 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. We report C-V (onset) lag, defined as the difference in time between two articulatory landmarks, the onset of TD retraction toward the vowel target and the onset of LA narrowing for the labial consonant. Movement onsets were determined by a 20% threshold of peak velocity in the movement towards target. A C-V lag value of 0 indicates that the movements started at the same time, negative values indicate that the TD moved first, positive values indicate that the LA moved first. C-V lag was analyzed using linear mixed effects models with a random intercept for subject; significance of fixed factors was determined by nested model comparison.

Results and discussion: Both LA and TD movements typically began well before the onset of voicing, with TD movement following slightly after LA, a pattern consistent with past work on Mandarin (Gao, 2009). Labial-initial CV monosyllables produced in isolation allowed for considerable freedom in the spatial position of the tongue dorsum (TD) at the onset of movement. Fig 1 shows the distribution of TD positions (z-scored within speaker); Fig 2 shows the distribution of C-V lag values (in ms). Both variables show a roughly normal distribution and no influence of vowel quality. Fig 3 shows a strong correlation between C-V lag and TD backness, a pattern consistent across vowels. When the TD is in a more front position (farther from the vowel target) movement tends to begin earlier (less lag) than when the TD happens to be in a more back position. It is as if movement towards target is delayed when the articulator is already near to the target or initiated earlier in time when the articulator is far away. We speculate that spatially-conditioned variation in vowel movement onset (relative to the consonant) subserves coordination of production goals occurring later in time. Notably, since movement onsets occur before phonation, our results may support the claim that positional information is available to feedforward timing control through somatosensory feedback, as argued by Tremblay, Shiller, & Ostry (2003). With respect to phonological patterns, spatially-conditioned timing relations may provide insight into how syllabic organization relates both to temporal regularity and to the sequencing of spatial targets within a syllable, i.e., phonotactics.
Fig 1. distribution of TD position at the start of vowel movement

Fig 2. distribution of C-V lag values in ms

Fig 3. Scatter plot of C-V lag and the longitudinal (back-front) position of the tongue dorsum. The lag between C and V is shorter when the TD is more advanced (front).

References