Gestural representations of tone in Mandarin: Evidence from timing alternations

Muye Zhang, Christopher Geissler, Jason Shaw
Department of Linguistics, Yale University

INTRODUCTION: The primary argument that lexical tones are articulatory gestures, as in the phonological primitives of Articulatory Phonology (Gafos & Goldstein, 2012), has come from the way that tones interact with other gestures. Specifically, lexical tone has been argued to condition patterns of relative timing between consonants and vowels based on evidence from kinematic data. In syllables with complex onsets (and no lexical tone), the onset of movement of a vowel tends to occur during the middle of the preceding consonant cluster, the so-called "c-center" effect (Browman & Goldstein, 1988; Hermes, Mücke, & Grice, 2013). Similarly, in languages with lexical tone, the vowel has been observed to begin movement around the midpoint between the onset consonant and the tone, a pattern reported in Mandarin Chinese (Gao, 2008), Thai (Karlin & Tilsen, 2015), and Lhasa Tibetan (Hu, 2016). Considered relative to a toneless CV syllable baseline, for which onset consonant and vowel begin movement at roughly the same time, it is striking that adding a consonant to the syllable to yield CCV and adding a lexical tone to yield, e.g., CV, result in the same pattern of relative timing—it is as if the tone is functioning as an onset consonant. Since Gao (2008), this has served as a working hypothesis within Articulatory Phonology and one that presents a viable alternative to the analysis of tones as autosegments. However, the gestural approach to tone makes a key prediction, which has yet to be tested. Within the same language, CV syllables with and without tone are predicted to *differ* in timing. Specifically, the lag between consonant and vowel gestures should increase in a syllable with tone relative to a closely matched syllable without lexical tone. In this paper, we present what is, to our knowledge, the first empirical evidence for tone-conditioned *timing alternations* of the type uniquely predicted by the Articulatory Phonology approach to tone.

EXPERIMENT: We chose Mandarin Chinese to test for timing alternations. In Mandarin, toneless syllables, often referred to as having "neutral" tone, are possible in a few ways: (1) certain grammatical morphemes are toneless, including the sentence-final question particle; (2) certain productive compounding paradigms condition tonelessness on embedded non-head members (Chen & Xu, 2006). We included both of these in our design as separate conditions (which we refer to as "Absent" for the grammatical morpheme and "Reduced" for neutral tones in compounds), along with matched syllables also embedded in compounds but fully specified for tone, the "Full" condition. An example stimulus set is provided in Table 1. The target syllable, shown in bold, was always /mu/ (or /mə/ for the question particle), produced in the "Full" condition with either falling (Tone 4) or low (Tone 3) tones and preceded by a carrier sentence and a context. The context was displayed on the screen before each item but not read aloud.

Table 1. Example stimulus set

Condition	**Context**	**Carrier**	**Target**
Full	*zhè yī lèi tùzi xǐhuān bǎ shù pí qǐ kāi, zhǎo chóngzi chī.*	*wǒmen gěi tā qǐmíng jiào*	*qǐ **mù** tù*
	This type of rabbit likes prying off tree bark to look for bugs to eat. We call it a bark-prying rabbit.		
Reduced	*zhè yī lèi tùzi hěn xǐhuān ānjìng de kàn qítā de tùzi.*	*wǒmen gěi tā qǐmíng jiào*	*qǐ **mu** tù*
	This type of rabbit likes quietly watching other rabbits. We call it an admiring rabbit.		
Absent	*zài yīgè jǐubā, xīn lái de fúwùyuán hěn bèn. lǎobǎn wèn:*	*píngzi gài dōu bù huì*	*qǐ **mə** tā*
	At a bar, there's a new employee who's incompetent. The manager asks: "he can't even open a bottle?"		

We recorded one native speaker of Beijing/Northern Mandarin producing 14 repetitions of each item (Carrier + Target) in four sets using Electromagnetic Articulography (EMA). Sensors on the upper lip and lower lip were used to compute change in lip aperture (LA) over time for the target /m/. A sensor on the tongue dorsum (5 cm behind the tip) was used to track tongue retraction towards the target for /u/. The relative timing of the onset of /m/ and the onset of /u/, i.e., C-V lag, is the key dependent variable. This was calculated by subtracting the timestamp of the onset of

movement of the vowel from the onset of movement of the consonant. To account for variations in speech rate, we divided raw C-V lags by total syllable duration. The key prediction is that C-V lag will be longer for the "Full" condition than for "Reduced" or "Absent" conditions. Except where otherwise noted, statistical significance is reported on the basis of nested comparison of linear mixed-effects models with item as a random effect.

RESULTS: Figure 1 shows only a marginal effect of condition on C-V lag (p=.0998), suggesting that mere presence or absence of tone is not a clear predictor. Figure 2 shows a significant interaction between condition and tone (p=.046). There is positive evidence for a timing alternation, but only for Tone 3. Pairwise t-tests corrected for multiple comparisons using Holm's method show that for Tone 3, the "Full" condition elicited a significantly greater C-V lag than the reduced-tone and neutral-tone conditions (p's<.02). Differences between "Reduced" and "Absent" conditions were not significant for either tone.

Figure 1: Lags by condition

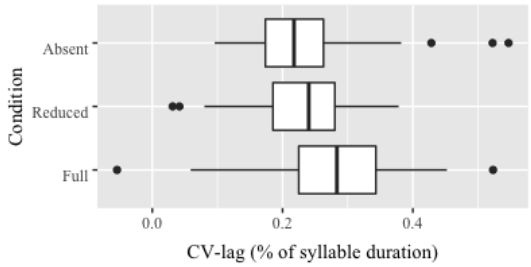
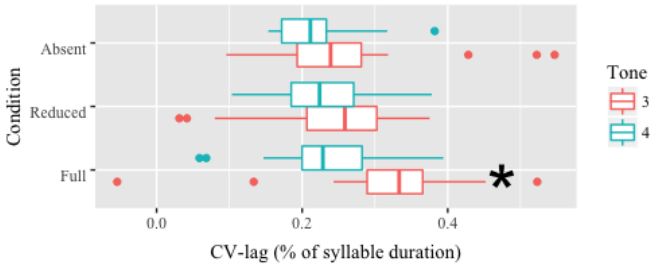


Figure 2: Lags by condition and tone



DISCUSSION: The C-V lag results provide the first piece of evidence for tone-conditioned timing alternations, a key prediction of the current working hypothesis for tonal gestures in Articulatory Phonology. The presence of tone conditions a change in C-V lag, lengthening the temporal interval between the onset of consonantal and vocalic gestures. This effect, however, was driven entirely by Tone 3. Tone 4 targets exhibited no significant difference in C-V lag when compared to reduced- or absent-tone Tone 4 targets. Although there has been some debate about the gestural analysis of Tone 3 in Mandarin (Hsieh, 2011; Yi & Tilsen, 2015), the distinction between Tone 3 and Tone 4 was not expected. One possible explanation is that Tone 4 gestures lack the anti-phase coupling relation with the onset consonant. Another is that the inventory of temporal relations is richer than just in-phase and anti-phase (c.f., the gesture-internal landmarks of Gafos, 2002; Goldstein, Nam, Saltzman, & Chitoran, 2009). Regarding the status of reduced- vs. absent-tone conditions, it seems that the morphosemantic compounding manipulations of lexical tone were effective: the C-V lag for the reduced-tone was statistically indistinguishable from the absent-tone for all targets. Our results here are based on just one speaker, but we plan to develop this paradigm further to probe the nature of tone-conditioned timing alternations and the consequences for the status of lexical tones as gestures.

References

Browman, C. P., & Goldstein, L. (1988). Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica*, *45*, 140-155.
Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory*, *20*, 269-337.
Gafos, A., & Goldstein, L. (2012). Articulatory representation and organization. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 220-231).
Gao, M. (2008). *Mandarin tones: An articulatory phonology account.* Yale University.
Goldstein, L. H., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled oscillator planning model of speech timing and syllable structure. In G. Fant, H. Fujisaki, & J. Shen (Eds.), *Frontiers in Phonetics and Speech Science: Festschrift for Wu Zongji* (pp. 239-249). Beijing: Commercial Press.
Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: The case of "impure s". *Phonology*, *30*(1), 1-25.
Hsieh, F.-Y. (2011). A gestural account of Mandarin tone 3 variation. *Proc. of the 17th international congress of phonetic sciences*, 890-893.
Hu, F. (2016). Tones are not abstract autosegmentals. *Speech Prosody 2016*, 302-306.
Karlin, R., & Tilsen, S. (2015). The articulatory tone-bearing unit: Gestural coordination of lexical tone in Thai. *Proc. of Meetings on Acoustics*, *22*(1), 1-9.
Yi, H., & Tilsen, S. (2015). Gestural timing in Mandarin tone sandhi. *Proc. of Meetings on Acoustics*, *22*(1), 1-11.