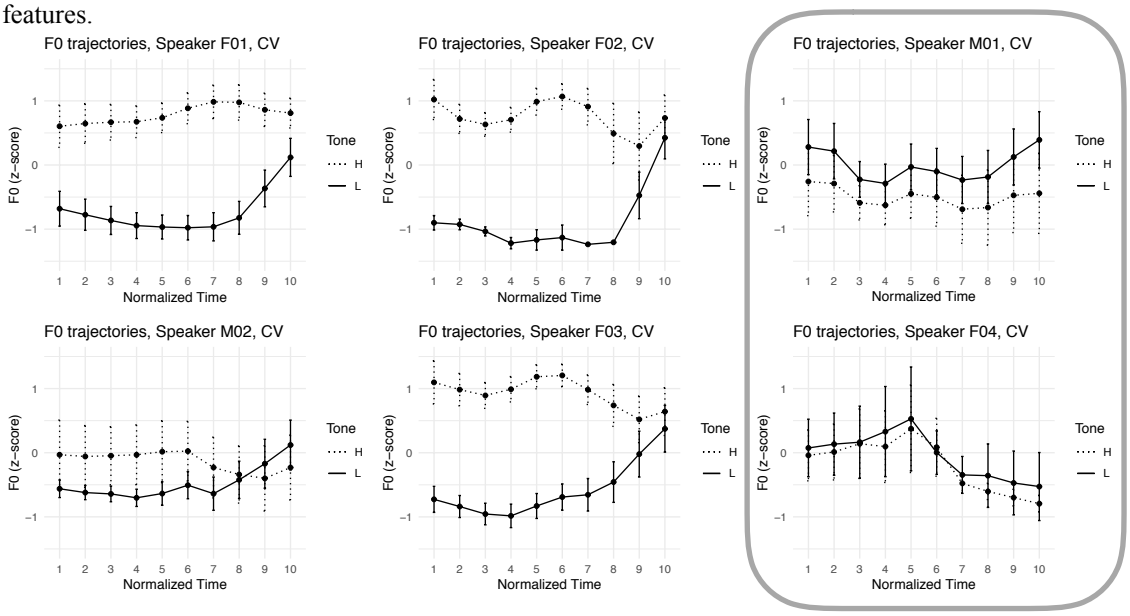
Consistent C-V timing across speakers of diaspora Tibetan with and without lexical tone contrasts

Languages differ in how articulatory gestures are coordinated in time, including the relative timing between consonants and vowels, i.e., C-V timing. Previous work has observed that in lexical tone languages, such as Mandarin (Gao 2008), Thai (Karlin and Tilsen 2014), and Lhasa Tibetan (Hu 2016), the lag between consonant and vowel gestures (C-V lag) is longer than in non-tonal languages, such as English (Lofqvist and Gracco 1999). Additionally, C-V lag is longer for tonal syllables than non-tonal syllables in Mandarin (Zhang et al. 2019). In the present study, we focus on diaspora Tibetan, which furnishes a unique example of a language community where some speakers contrast tones and others do not, allowing comparison across speakers who do and do not contrast tone in the same language. We predicted that C-V lag would be longer for speakers that maintain the lexical tone contrast than for those that do not. The prediction follows from the empirical observations cited above as well as the model of tone as an articulatory gesture (Gao 2008; Niemann et al., 2011).

Six speakers (four female) of Tibetan raised in India and Nepal and living in the United States participated in this study. Using Electromagnetic Articulography (EMA), we recorded the movement trajectories of the lips and tongue dorsum while speakers produced target words in a carrier phrase. Target items (N = 71) were one and two syllable words that varied in word-initial consonant (/m/, /p/, /p^h/) and tone (high and low/rising); first syllable vowels drew from the set of back vowels in Tibetan: /u/, /o/, /a/. Speakers produced each item 4-10 times. Gestures for lip aperture and tongue dorsum retraction were identified as starting where 20% of peak velocity toward target was reached. Time-normalized F0 trajectories were analyzed for systematic differences by lexical tone category; two of the six participants (one female, one male) did not contrast tone (Fig. 1). However, contrary to the predictions of the model of tone as an articulatory gesture, tonal and non-tonal speakers did not differ in C-V lag. We also investigated C-V lag across aspirated and unaspirated stops, finding no consistent differences (Fig. 2).

We interpret the lack of difference in C-V lag across tonal and non-tonal speakers as evidence against a tone gesture interacting with the other gestures in a word to affect C-V lag. Instead, it appears that the Tibetan speakers in our study learned a pattern of C-V timing common to their linguistic community, which includes speakers that maintain lexical tone. C-V lag also did not differ across aspirated and unaspirated stops, showing greater stability across consonant voicing contrasts than has been observed in other languages (Spanish: Gibson et al 2017; English: Sotiropoulou et al 2016). Taken together the results indicate that C-V timing can remain consistent across speakers, including those that have lost contrastive lexical tone, and across consonants that vary in laryngeal features.



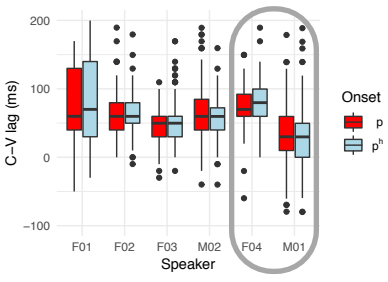


Fig. 1 (above): F0 trajectories for each speaker. The two speakers at right (circled) were found not to contrast tone. Mean time-normalized F0 is plotted, with error bars at 95% confidence intervals.

Fig. 2 (right): Tukey-style boxplot of C-V lag (ms) measurements by speaker and onset. Speakers F04 and M01, at right and circled, do not contrast tone, while other speakers do contrast tone.