

The influence of prosodic structure on tone variation and change in Chengdu Mandarin

Introduction As a tonal language, Mandarin Chinese relies on pitch contours rather than stress to distinguish lexical meaning. However, in connected speech, syllables still vary in relative degrees of prosodic prominence, which may systematically alter the phonetic realization of lexical tones. Stronger accounts even argue that prosodic structure directly triggers tone sandhi (Chen, 1993; Guo, 2020; Jiang, 2011). For disyllabic words in particular, it has been proposed that syllables with greater prosodic prominence tend to preserve the citation forms, whereas less prominent syllables are more likely to receive sandhi variation (Duanmu, 2014; Yue, 1987). Specifically, in a left-dominant structure, the first syllable of a disyllabic word is expected to retain its citation form and carry the tonal traces onto the second syllable, where tone sandhi occurs. In a right-dominant structure, the initial syllable of the word alters its tonal realization to better transit to the following tone. Words without prominence-based preference might have a local modification on the phonetic realization of tones in both syllables.

Mandarin Chinese is rich in its regional varieties, each exhibiting a unique tonal system, but how the prosodic structure modulates tonal realization remains unclear. Chengdu Mandarin, for example, has a left-dominant structure, where the prosodic stress mostly occurs in the initial syllable of the disyllabic word and usually with a longer duration (Liao, 2005; Qin, 2015). According to previous work, there are four lexical tones in Chengdu Mandarin and their citation forms are assumed as follows: Tone 1 [35], Tone 2 [31], Tone 3 [53], Tone 4 [312] (He, 2015; Zhao, 2023). It has been shown that tone sandhi in Chengdu Mandarin generally occurs at the right edge of the prosodic words; the only exception is Tone 3 which changes its phonetic realization in the initial syllable (Lin, 2014). Whether this exception overrides the presumed influence of prosodic structure on tone variation and whether it persists in present-day speech require further investigation using empirical data. In addition, not much research has been done to examine how sandhi patterns change over time. The present study investigates preliminarily whether tone sandhi in Chengdu Mandarin consistently occurs at the left edge of the disyllabic words and whether the sandhi patterns exhibit diachronic change by comparing production patterns between older and younger speakers using an apparent-time paradigm.

Method The recordings of twenty disyllabic words were collected from five native Chengdu Mandarin speakers. The speakers were divided into two age groups: old ($N = 2$; aged 66 and 71) and young ($N = 3$; aged between 22–28). Each word represented a distinct combination of lexical tones over two syllables—four possible tone categories for the first syllable and five possible categories for the second syllable. All the syllables in the disyllabic words allowed for a full set of tone realizations in Chengdu Mandarin. The recordings were force aligned at the word, syllable, and phone levels using a combination of the Montreal Forced Aligner (MFA, McAuliffe, et al., 2017) and Praat (Boersma & Weenink, 2020). Word- and syllable-level alignments were generated using a custom Praat script; phone-level alignments were performed by MFA (v3.2.1) with the pretrained Mandarin acoustic model. Manual corrections were applied to misaligned boundaries.

F0 contours were used to represent the phonetic realization of the lexical tone (Jongman et al., 2006; Tupper et al., 2020). Eleven equally spaced F0 values in hertz were extracted over the sonorant portion of each syllable in the disyllabic words, and then converted to Chao tone numerals using a formula of t-values to normalize variability between speakers on a scale of 0-5 (Shi, 2008). The f0 measurement of Chengdu monosyllabic words were taken from a spoken corpus

of Mandarin dialects, the ManDi corpus (Zhao & Chodroff, 2022), using the same method. Second-order polynomial models were implemented to predict time series data in a curve; the fixed effects included F0 point and age group to test for statistically significant differences in sandhi forms between the age groups; a random slope was set for the types of tone combination.

Results We found that tone sandhi occurred at varying positions in the disyllabic words in Chengdu Mandarin. Chengdu Tone 1 preserved its citation form [35] when it was in the initial syllable of the words (Figure 1). However, in the second syllable, Tone 1 changed from [35] to [45] for both age groups (Figure 3). Unlike Tone 1, sandhi variation of Tone 2 and Tone 3 occurred in the initial syllable of the words: Tone 2 changed from a falling tone [31] to a level tone [33] (Figure 4), and Tone 3, also falling in the citation form, changed from [53] to a high-level contour of [44/45] (Figure 7). Tone 4 was realized as a dipping tone in both syllables, but schematically differed in the position of the turning point across syllables (Figure 10). In the first syllable of the word, Tone 4 reached the lowest pitch level around 40% into the sonorant portion, resembling the shape of its citation form, while in the second syllable, the turning point of Tone 4 was delayed around 70% in the vocalic part especially for the older speakers (Figure 10 & 12).

For age differences, the sandhi forms of Tone 1, Tone 2 and Tone 3 did not differ significantly between the older and younger speakers (Tone 1 \sim age: $\beta = -0.06$, $t = -1.13$; Tone 2 \sim age: $\beta = -0.08$, $t = -1.45$; Tone 3 \sim age: $\beta = -0.04$, $t = -0.91$). For the non-sandhi forms, the younger speakers had a curvier rise for Tone 1 in the first syllable than the older speakers (Tone 1 \sim age: $\beta = 2.60$, $t = 5.10$) (Figure 2); they also had a steeper fall for Tone 2 and Tone 3 in the second syllable (Tone 2 \sim age: $\beta = -2.51$, $t = -4.48$; Tone 3 \sim age: $\beta = -0.85$, $t = -1.78$) (Figure 6 & 9).

Conclusion Though Chengdu Mandarin has a left-dominant prosodic pattern, it seems that tone sandhi does not consistently occur at the left edge of the disyllabic words. It is possible that the relation between prosodic structure and tonal variation is more complicated than a one-to-one correspondence—the direction of tone sandhi might be regulated by other factors apart from prosodic prominence and can be highly language- or variety-specific. For variation across age groups, the sandhi forms seemed relatively stable over a thirty-year span, which might suggest that sandhi patterns are systematic for a language variety regardless of the intrinsic variation over time. For the non-sandhi forms, the younger speakers seemed to have curvier contours than older speakers in general. The above findings are preliminary given the limited size of the data; more data will be obtained to further confirm our conclusions.

The influence of prosodic structure on tone variation and change in Chengdu Mandarin

Tone 1

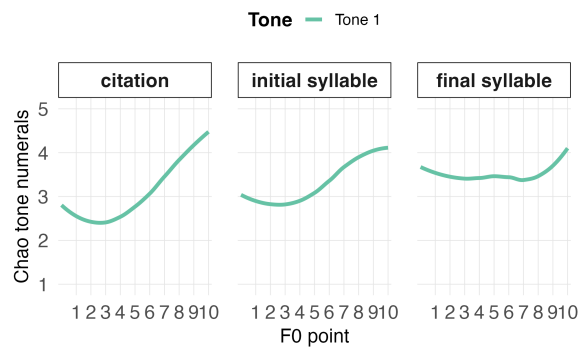


Figure 1: Tone 1 in monosyllabic and disyllabic words (as the initial and the final syllable)

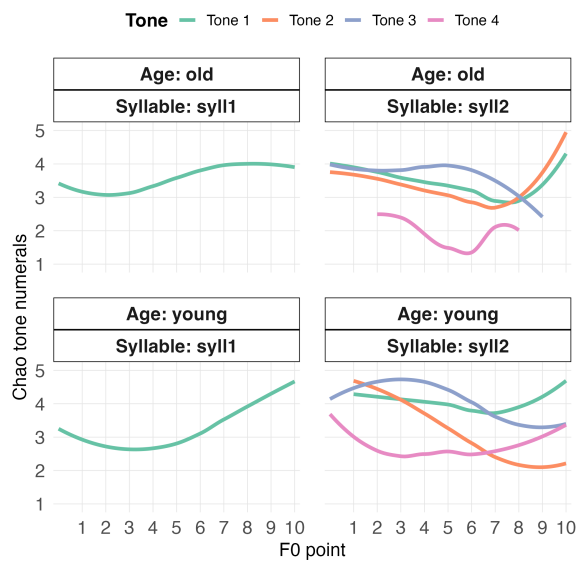


Figure 2: Tone 1 in the initial syllable of disyllabic words between the age groups

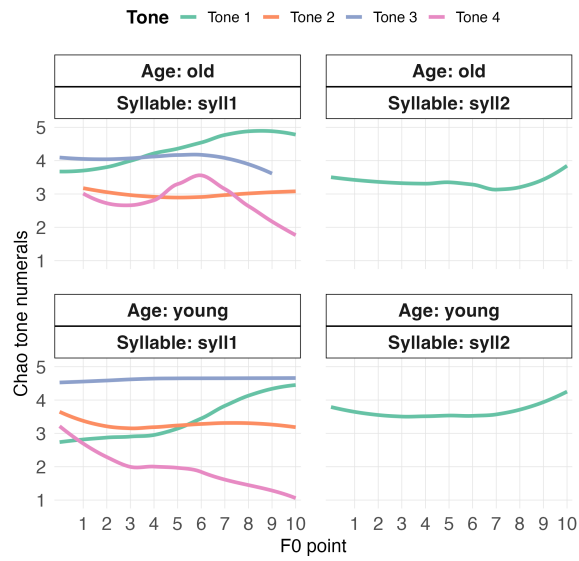


Figure 3: Tone 1 in the final syllable of disyllabic words between the age groups

Tone 2

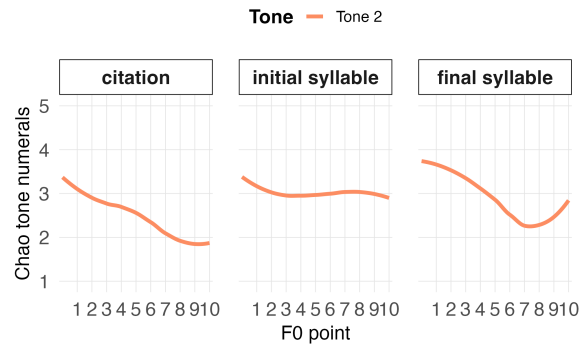


Figure 4: Tone 2 in monosyllabic and disyllabic words (as the initial and the final syllable)

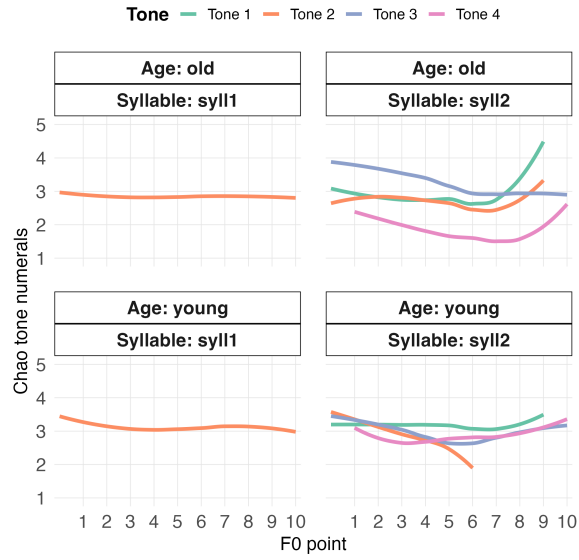


Figure 5: Tone 2 in the initial syllable of disyllabic words between the age groups

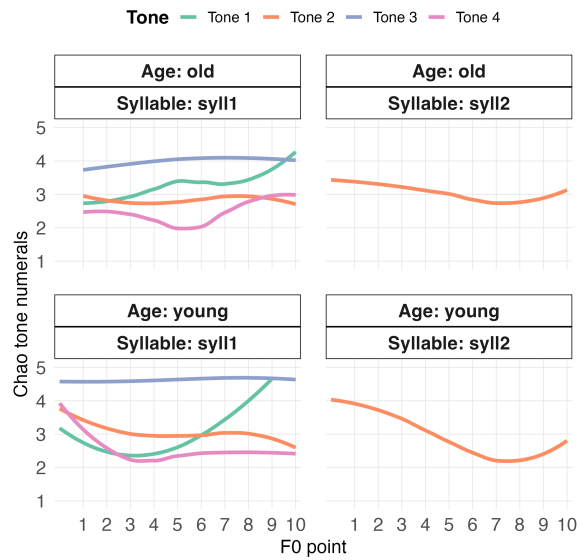


Figure 6: Tone 2 in the final syllable of disyllabic words between the age groups

Tone 3

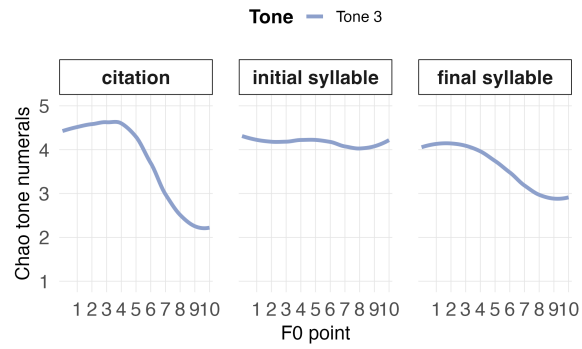


Figure 7: Tone 3 in monosyllabic and disyllabic words (as the initial and the final syllable)

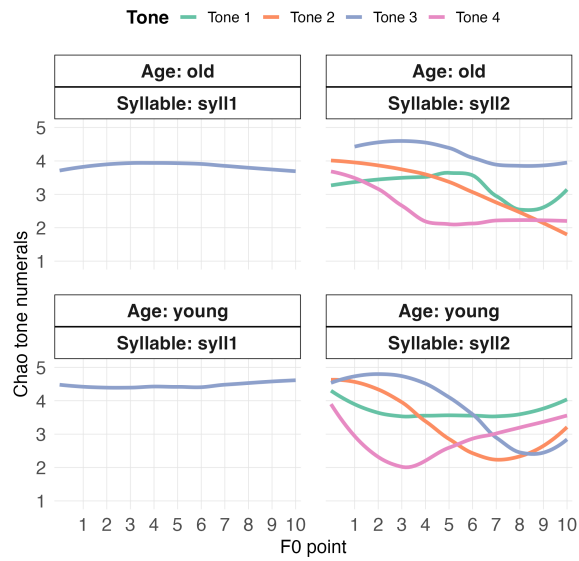


Figure 8: Tone 3 in the initial syllable of disyllabic words between the age groups

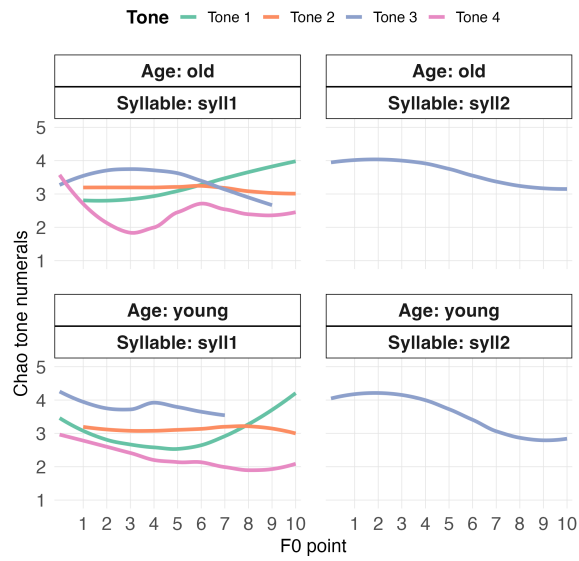


Figure 9: Tone 3 in the final syllable of disyllabic words between the age groups

Tone 4

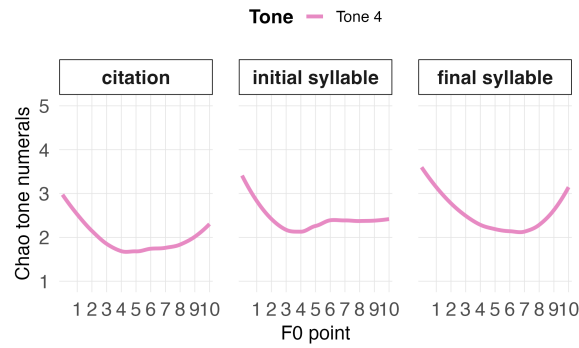


Figure 10: Tone 4 in monosyllabic and disyllabic words (as the initial and the final syllable)

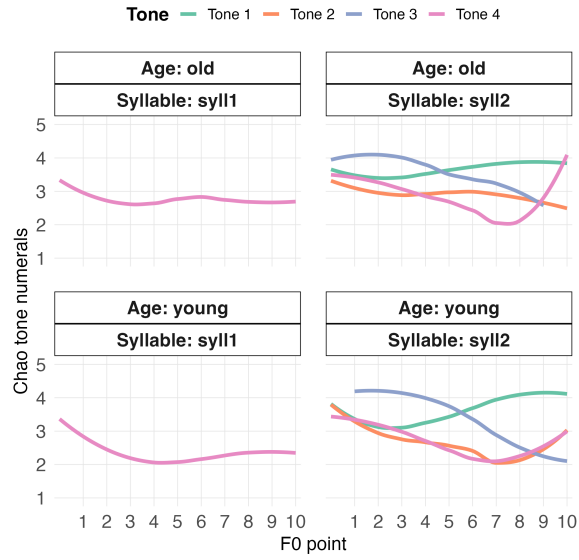


Figure 11: Tone 4 in the initial syllable of disyllabic words between the age groups

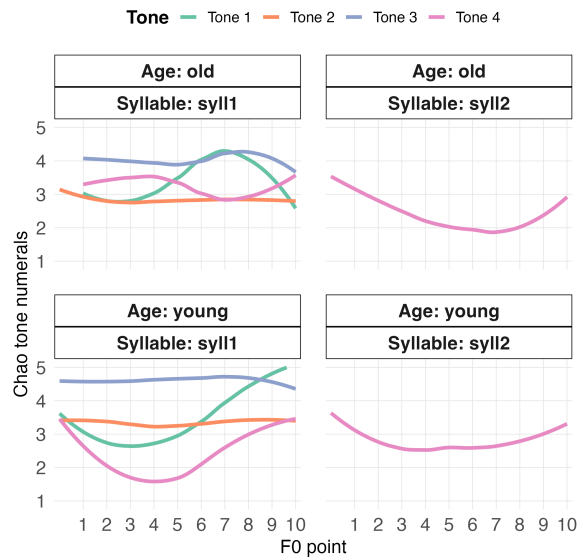


Figure 12: Tone 4 in the initial syllable of disyllabic words between the age groups