An articulatory study of word-level prominence in two Mandarin dialects Jing Huang, Feng-fan Hsieh, Yueh-chin Chang

National Tsing Hua University

xia okuidaren 94 @ 163. com, ffhsieh @ mx.nthu.edu.tw, ycchang @ mx.nthu.edu.tw

Introduction: The existence of word-level prominence/lexical stress in Mandarin Chinese is controversial. In this study, we investigated whether word-level prominence is attested in two relatively understudied dialects of Mandarin, Southwestern Mandarin (SWM) and Taiwanese Mandarin (TWM). Unlike Beijing Mandarin, SWM and TWM do not have "unstressed" (neutral tone) syllables in lexical words, meaning that no apparent strong-weak patterns have been reported for them, at least impressionistically. Nevertheless, little is known if there is any articulatory difference between the syllables in a word. Given that stressed syllables may involve longer, larger, and faster gestures than their counterparts in unstressed positions (e.g., Katsika and Tsai 2021), we set out to assess the supralaryngeal kinematic correlates of word-level prominence/lexical stress in these two Mandarin dialects, using electromagnetic articulography (EMA).

Method: We analyzed and reported data from four SWM speakers (1 female) and two TWM speakers (1 male) in their twenties. The stimuli are personal names with two or three identical syllables (note that these are *not* reduplicated forms): {papa, pepe, tutu, titi} in SWM (all with the mid tone) and {pi, pai, tai, two/papapa, tatata, kakaka} in TWM (all with the high-level tone). Seven repetitions for each token were collected for SWM and ten repetitions for TWM. For SWM, the target words were embedded in the carrier phrase, " $p^he A$, $pu p^he B$ ", meaning "(Please) pat A! not <u>B</u>." The tokens in the unfocused position B were analyzed. For TWM, the target words were embedded in X in the carrier phrase "wŏ xiǎng X bèi Y", meaning "I think <u>X</u> memorizes Y." The target words were produced in a position with the *least* higher-level prominence (i.e., phrasal/sentential stress) and without the well-established effect of domain-initial strengthening (Keating et al. 2003) to highlight the word-level prominence.

Results: In disyllables, the initial syllables are significantly longer in duration (p<.05) in SWM, while the final syllables are significantly longer (p<.05) in TWM. See (1). Note that the raw data are presented here for ease of interpretation only.

(1) SWM-Initial σ	SWM-Final σ	TWM-Initial σ	TWM-Final σ
209 ms (SD: 56)	172 ms (SD: 55)	135 ms (SD: 22)	153 ms (SD: 25)

Regarding gesture duration, the onset gestures (i.e., Lip Aperture and Tongue Tip along the vertical dimension) are significantly longer (p < .05) in initial position than in final position in TWM. In contrast, no significant difference is found for SWM. See (2).

(2) SWM-Initial σ	SWM-Final σ	TWM-Initial σ	TWM-Final σ
55 mm (SD: 29)	51 mm (SD: 27)	78 mm (SD: 39)	42 mm (SD: 16)

Regarding the peak velocities of the onset gestures, our results reveal that there is no significant difference between the initial and final syllables in SWM and TWM (3).

(3) SWM-Initial σ	SWM-Final σ	TWM-Initial σ	TWM-Final σ
13 cm/s (SD: 7)	12 cm/s (SD: 8)	15 cm/s (SD: 7)	15 cm/s (SD: 6)

The acoustic and articulatory results of the trisyllables in TWM are given in (4). It appears that there is a "mismatch" in acoustics and articulation: the non-initial syllables are significantly longer in duration (p<.05); on the other hand, the onset gestural duration is the longest in word-initial position (p<.05). Finally, other acoustic cues such as F₀ and intensity are not significantly different across the board (not shown here).

(4) TWM	Initial σ	Medial σ	Final σ
Syllable duration	134 ms (SD: 28)	150 ms (SD: 31)	159 ms (17ms)
Gestural duration	83 mm (SD: 41)	37 mm (SD: 17)	37 mm (SD: 18)
Peak Velocity	13 cm/s (SD: 7)	17 cm/s (SD: 7)	15 cm/s (SD: 8)

Discussion: The present results reveal a novel cross-dialectal difference: both word- and phrase-level prominences are cued by phonetic duration in SWM because the initial syllable is longer, although no kinematic differences (gestural duration/peak velocity) were found. On the other hand, TWM shows mixed results; the non-final syllables are longer in acoustics, but articulatorily, the onset gestures are significantly longer in word-initial positions. This cross-dialectal difference might be attributed to the possibility that in TWM, the word-level prominence does not coincide with the phrasal/edge prominence. In other words, the longer gesture duration in word-initial positions can be regarded as an exponent of word-level prominence/lexical stress and phrasal prominence (longer duration) falls on the final syllables in TWM (see also Duanmu 2007 for a similar view). The cross-dialectal distinction is schematized in (5).



From a broader perspective, both the SWM and TWM results further suggest that the gestural duration in articulation might be orthogonal to the syllable duration in acoustic measurements.

Conclusion: In this study, we have shown that there are different sources of prosodic prominence in Mandarin: specifically, word-level prominence in TWM is more reliably detected in kinematics but not in acoustics. In sum, the present results indicate that the articulatory properties also contribute to the realization of prosodic prominence, thus casting doubt on the widely accepted claim that Mandarin has no lexical stress.

References

[1] Katsika, A. and K. Tsai. 2021. The supralaryngeal articulation of stress and accent in Greek. *JPhon* 88.

[2] Keating, P., Cho, T. H., Fougeron, C., & Hsu, C. S. 2003. Domain-initial strengthening in four languages. In Local J. et al. (Eds.), *Laboratory Phonology VI: Phonetic Interpretation*, pp.145-163. CUP.

[3] Duanmu, S. 2007. The Phonology of Standard Chinese. OUP.