

Disentangling Focus and Lexical Tone in Taiwan Mandarin Articulation

Sherry Chien¹, Argyro Katsika¹

¹*University of California, Santa Barbara*

In Taiwan Mandarin, focus is marked by raised F0, increased intensity, and lengthened duration of the entire focused word [1]. In Standard Chinese, focus has also been shown to expand lexical tone F0 contours relative to post-focus contexts, enhancing tonal distinctiveness within the focused constituent [2]. If similar focus-driven tonal expansion characterizes Taiwan Mandarin, this raises a central question for tone languages: how pitch simultaneously encodes both lexical and intonational information, and whether these modulations draw on shared or distinct phonetic resources. In parallel, studies of stress languages have shown that focused elements exhibit longer, larger, and faster constriction gestures [3], [4], [5], and that these phonetic modulations systematically scale across focus types—from unfocused to broad, narrow, and contrastive focus [6], [7], [8]. Whether focus types are kinematically encoded in tone languages remains unknown. This study therefore investigates whether focus affects supralaryngeal kinematics in Taiwan Mandarin and whether such effects differentiate among focus types in ways comparable to those observed in stress languages.

Electromagnetic articulography (EMA) was used to investigate the kinematic effects of focus and its interaction with lexical tone in Taiwan Mandarin. Target stimuli were two monosyllabic lexical-tone minimal sets (Table 1), elicited phrase-medially under five conditions: contrastive focus (CF), narrow focus (NF), broad focus (BF), pre-focus (pre-F), and post-focus (post-F). All target words were preceded and followed by the particle 的 (*de*), a nominal modifier in Mandarin, which carries the neutral tone (T0). The particle preceding the target word is referred to as the *pre-target word*, and the particle following the target word is referred to as the *post-target word*. The onset consonant (C) gestures were labeled for target, pre-target, and post-target words. Separate linear mixed-effects models were fit to two kinematic measures—formation duration and displacement—at each word position. Fixed effects were Focus Type (CF, NF, BF, pre-F, post-F) and Lexical Tone (T1, T2, T3, T4). Random intercepts included Speaker and Segment Composition (/ma/, /mau/, /nau/) for all models, as well as the preceding segment for pre-target words, and the following segment for post-target words.

Analyses of five speakers reveal systematic kinematic modulation associated with both lexical tone and focus, with effects distributed differently across target, pre-target, and post-target words. Notably, no significant interaction between focus and lexical tone was observed, suggesting that their effects on articulatory gestures operate independently. For target words, lexical tones primarily affect the displacement of articulatory gestures: C gestures co-produced with contour tones (T2, T3, T4) show significantly greater displacement than those with high-level tone (T1) (Fig. 1), revealing larger gestures for contour tones. Focus is instead marked kinematically through temporal adjustment: narrow (NF) and contrastive focus (CF) elicited longer gestures than broad focus (BF) and unfocused contexts (pre-F and post-F) (Fig. 2). Individual types of focus were not distinguished.

Tone-dependent effects of lexical tone are also observed in pre-target words, with larger C-gesture displacement before T1 and reduced displacement before T4 (Fig. 3), as well as shorter formation duration before T3 (Fig. 4), indicating systematic pre-target modulation associated with lexical tone. In post-target words, focus shows a strong post-target influence: gestures following contrastively and narrowly focused targets show significantly larger displacement than those following broad-focus or post-focal targets (Fig. 5).

Taken together, these results indicate that lexical tone modulates articulatory gestures primarily through displacement, with effects localized to the target and extending anticipatorily to pre-target positions. Focus, by contrast, primarily shapes the temporal organization of gestures at the target, while also influencing displacement in post-target contexts.

Table 1. Target words used in the experiment.

Tone	/ma/	/mao/ or /nao/
T1	ma1/媽 ‘mother’	mao1/貓 ‘cat’
T2	ma2/麻 ‘hemp’	mao2/毛 ‘fur’
T3	ma3/馬 ‘horse’	nao3/腦 ‘brain’
T4	ma4/罵 ‘scold’	mao4/帽 ‘hat’

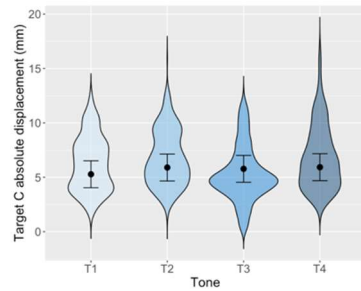


Fig. 1. Absolute displacement of target onset C gesture across lexical tones. Violin plots show raw data distribution; dots show model-estimated means (± 1 SE).

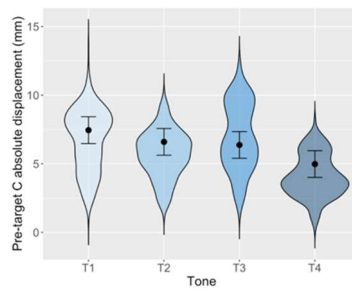


Fig. 3. Absolute displacement of pre-target onset C gesture across lexical tones. Violin plots show raw data distribution; dots show model-estimated means (± 1 SE).

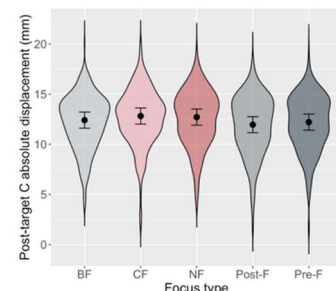


Fig. 5. Absolute displacement of post-target onset C gesture across focus types. Violin plots show raw data distribution; dots show model-estimated means (± 1 SE).

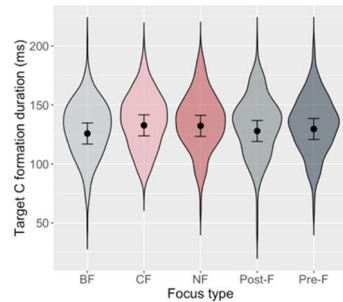


Fig. 2. Formation duration of target onset C gesture across focus types. Violin plots show raw data distribution; dots show model-estimated means (± 1 SE).

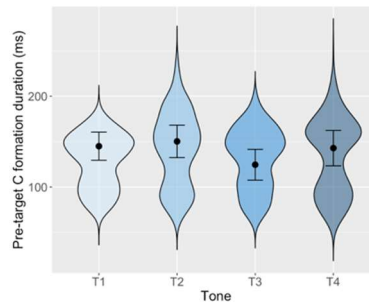


Fig. 4. Formation duration of pre-target onset C gesture across lexical tones. Violin plots show raw data distribution; dots show model-estimated means (± 1 SE).

References

- [1] S. Chen, B. Wang, and Y. Xu, “Closely related languages, different ways of realizing focus,” in *Interspeech 2009*, Brighton, United Kingdom: ISCA, 2009, pp. 1007–1010.
- [2] Y. Chen, “Post-focus F0 compression—Now you see it, now you don’t,” *Journal of Phonetics*, vol. 38, no. 4, pp. 517–525, Oct. 2010.
- [3] M. E. Beckman and J. Edwards, “Articulatory evidence for differentiating stress categories,” in *Papers in laboratory phonology III: Phonological structure and phonetic form*, 1st ed., P. A. Keating, Ed., Cambridge University Press, 1994, pp. 7–33.
- [4] T. Cho, “Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English,” *The Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3867–3878, Jun. 2005.
- [5] T. Cho, “Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English,” in *Papers in laboratory phonology VIII: Varieties of phonological competence (phonology and phonetics)*, L. Goldstein, D. H. Whalen, and C. T. Best, Eds., Berlin: Mouton de Gruyter, 2006, pp. 519–548.
- [6] A. Katsika, J. Jang, J. Krivokapic, L. Goldstein, and E. Saltzman, “A hierarchy of prominence: The production and perception of focus in American English,” in *Proceedings of the 20th International Congress of Phonetic Sciences*, R. Skarnitzl and J. Volin, Eds., Prague, Czech Republic: Guarant International, 2023, pp. 1696–1700.
- [7] D. Mücke and M. Grice, “The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation?,” *Journal of Phonetics*, vol. 44, pp. 47–61, May 2014.
- [8] S. Roessig and D. Mücke, “Modeling Dimensions of Prosodic Prominence,” *Front. Commun.*, vol. 4, p. 44, Sep. 2019.