Syllable-based or Word-based? Representation of tones undergoing merger in Hakka

Ming Liu, Jonathan Havenhill
The University of Hong Kong

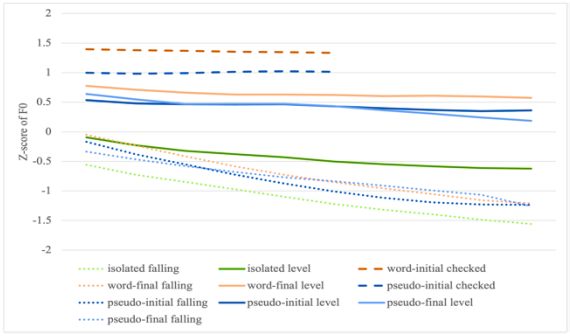
In most studies on Chinese dialects, it is commonly accepted that the syllable serves as the basic unit of tone, and that tonal realizations of multisyllabic words are subject to tone sandhi rules [1]. However, this approach may not be applicable to some southern Chinese dialects, which lack standard dialect-specific writing systems and have complex tone sandhi patterns [2-3]. Moreover, many southern Chinese dialects, such as Hakka, are currently undergoing tone merger triggered by the loss of stop codas. The historical checked tone syllables are partially merging into non-checked tone categories, which results in a more complex tonal system at the syllable level and less transparent tone sandhi rules [4-5]. Hence, it becomes challenging to extract the underlying tonal form and corresponding sandhi rules for each syllable from the surface realizations of words [6-7]. In this situation, do native speakers still acquire syllable-based tonal representations, or do they learn tonal realizations at the word level for these checked syllables undergoing tone merger?

To address this question, we conducted a production experiment focusing on checked tones in Wangmudu Hakka. In this Hakka variety, historical checked tones are at a late stage of merger into non-checked tones [8]. All the checked tone syllables have lost their stop codas and have the same tonal realization as non-checked tones (i.e., mid-level & low-falling) in isolation or in word-final positions. However, in word-initial positions, checked tones maintain a distinct realization in tone sandhi, i.e., a high-level tone with short duration, while non-checked tones vary depending on the following syllable. Nevertheless, this etymon-related distinction is less absolute among younger speakers, suggesting a higher rate of tone merger and greater tone sandhi opacity.

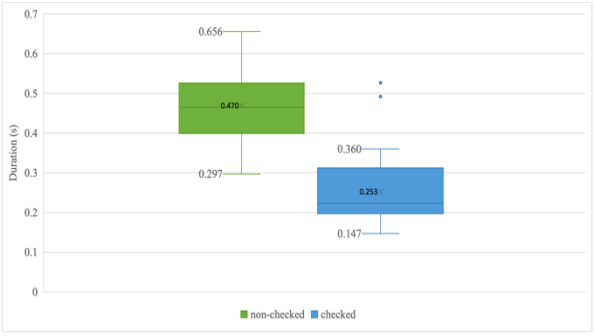
We recruited 10 adult native speakers of Wangmudu Hakka (4 women; mean age 30) to participate in three production tasks of isolated syllables, natural disyllabic words, and disyllabic pseudo-words. In the first task, participants read 50 checked tone syllables. 20 of them are free morphemes, 11 are bound morphemes that appear only in word-final position, 11 only in word-initial position, and 8 in both positions. For the second task, 70 natural disyllabic words containing these syllables were chosen as prompts. In the third task, we created another 70 pseudo-words using these syllables, including 40 items that violate the positional constraints of the target syllable (Table 1). Acoustic analysis of the audio data was performed in Praat [9].

In natural disyllabic words, participants produced averagely 85% of word-initial checked syllables with a high-level pitch contour (Figure 1) and with a significantly shorter duration than the corresponding non-checked realizations in word-final positions (Figure 2, $p < 0.001$). In isolated contexts, most checked syllables were realized as non-checked tones (level or falling), aligning with the trend of tone merger. However, 15% of targets, exclusively bound morphemes, were realized as other non-checked tones or even with different vowels, indicating that participants failed to extract expected underlying forms. In disyllabic pseudo-words, significantly less targets (66%, Figure 4) were realized with checked tones in word-initial position. Additionally, there was substantial individual variation among participants in terms of specific realizations of certain pseudo-words, suggesting a lack of common sandhi rules that can be generalized to new words.

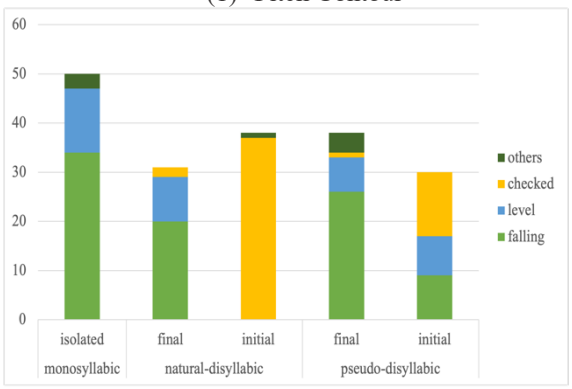
Overall, the production data suggest that, for checked tones undergoing merger, native Hakka speakers have learned tonal realizations at the word level but not at the syllable level, nor have they acquired generalizable tone sandhi rules that may be applied to new words. This finding further leads us to rethink what the underlying tone representation units are for speakers of Chinese dialects with complex tone sandhi patterns. For bound morphemes that have various surface realizations, speakers may not form syllable-based tone representations; instead, the tones could be embedded as a part of the lexicon, with tonal representations being acquired at the word-level.



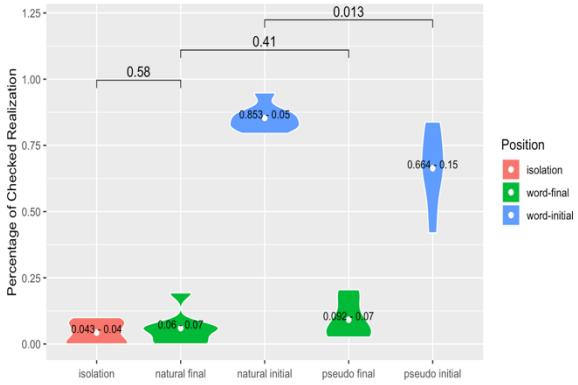
(1) Pitch Contour



(2) Syllable Duration



(3) Distribution of Tones



(4) Percentage of Checked Realizations

Figures: (1) Pitch contours of target syllables under different conditions, plotted using Z-scores of F0 value (log Hz). (2) Duration comparison between checked (high-level) and non-checked (level, falling) tonal realizations. (3) Distribution of different tonal realizations under different conditions for one male speaker. (4) Percentage of checked realizations of each speaker in different tasks.

Table 1: Wordlist Sample **[10]**

syllable	tone	etymon	morpheme	word-initial	word-final	pseudo-initial	pseudo-final
tsʰjæ 七	21	tsʰiet	free	tsʰjæ5 tʰə24	tsʰɵ55 tsʰjæ21	tsʰjæ? tʰə24	tsʰɵ55 tsʰjæ?
tsʰjæ 膝	?	siet	bound	tsʰjæ5 tʰə24			tsʰɵ55 tsʰjæ?
tsʰjæ 戚	?	tsʰɛk	bound		tsʰjen55 tsʰjæ21	tsʰɵ55 tsʰjæ?	

References

[1] Zhang, J. (2014). Tones, tonal phonology, and tone sandhi. The handbook of Chinese linguistics, 443-464.
[2] Li, X., & Xiang, M. (2009). Hanyu fangyanxue jichu jiaocheng [An introductory course on Chinese dialectology]. Peking University Press.
[3] Chen, M. Y. (2000). Tone sandhi: Patterns across Chinese dialects (Vol. 92). Cambridge University Press.
[4] Zhang, J. (2019). Speakers Treat Transparent and Opaque Alternation Patterns Differently—Evidence from Chinese Tone Sandhi. In *Proceedings of the 36th West Coast Conference on Formal Linguistics. Somerville, MA: Cascadilla Proceedings Project*.
[5] Kiparsky, P. (1973). Abstractness, opacity and global rules. Indiana University Linguistics Club.
[6] Zhang, J. (2010). Issues in the analysis of Chinese tone. Language and Linguistics Compass, 4(12), 1137-1153.
[7] Chen, M. Y. (2004). Changting Hakka tone sandhi: Analytical challenges. Yuyan ji Yuyanxue, 5(4), 799-820.
[8] Liu L.X. (2001). Jiangxi kejia fangyan gaikuang [Introduction to Hakka dialects in Jiangxi]. Jiangxi People's Press.
[9] Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. Glot International, 5(9/10), 341-347.
[10] Guo X.L. (1986). Hanzi guyin shouce [The handbook of ancient Chinese]. Peking University Press.