THE EFFECT OF ONSET AND POSITION IN THE REALIZATION OF TONE 1 IN TWO DIALECTS OF TAIWAN MANDARIN

Janice Fon,[1] Hui-ju Hsu,[2] Yi-Hsuan Huang,[1] & Sally Chen[1]

[1]Graduate Institute of Linguistics
National Taiwan University

[2]Dept. Appl. Linguistics & Language Studies
Chung Yuan Christian University

{jfon, r94142011, d93142002}@ntu.edu.tw,[1] hsuhj@cycu.edu.tw[2]

ABSTRACT

This study investigates how onset and sentence positioning affect the realization of Tone 1 in two dialects of Taiwan Mandarin. Results showed that the central dialect was higher in register when placed in isolation, but lower when placed in a sentential context. When there was a tonal mismatch, coarticulatory effects were more robust in the northern dialect, implying that speakers of the central dialect (nonstandard) might be more self-conscious about the standard-vernacular distinction than those of the northern dialect (standard), and overcorrection tended to occur. The effect of onset type was also significant but fairly localized. Obstruent-initial syllables had higher initial pitch than sonorant ones. The declination effect was also significant, the rate of which being higher in the central variety. In addition, sentential stress tended to raise the sentence-final H targets in both varieties. However, the PENTA model was not fully supported.

Keywords: high tone, Taipei Mandarin, Taichung Mandarin, dialectal variation, tonal realization

1. INTRODUCTION

Taiwan Mandarin (TM) is the official language of Taiwan and is genetically related to Mainland Mandarin (MM), the official language of Mainland China. However, due to almost 60 years of political separation between the two places, the two Mandarins have developed independently so that dialectal variations are obvious to speakers of either variety [6].

Political division is not the only cause for the divergence of the two dialects, however. Ethnic distributions are also different. About 73-80% of the population in Taiwan is Southern Min, who speaks a variant of Mainland Southern Min, and this Min is therefore a powerful substrate language for TM [5, 6].

Both Mandarin varieties have four tones, traditionally termed Tone 1 (T1), Tone 2 (T2), Tone 3 (T3), and Tone 4 (T4), which are realized as high level, mid dipping, mid-low dipping, and high falling, respectively, with T2 having an allophonic variant of mid rising in MM, and T3 a mid-low fall in both varieties [7, 16]. Although the phonological categories of the four tones are the same between the two dialects, the phonetic realizations are somewhat different. Specifically, tonal registers of TM T2 and T3 are much lower and narrower than those of the MM variety [8]. This discrepancy is presumably due to the influence from Min, which seems to prefer a lower frequency range [4, 11]. This study thus planned to see if such a lowering effect is also affecting T1, which is a high tone.

2. AIMS OF THE STUDY

There are four specific aims in this study. First of all, we would like to explore possible dialectal differences in T1 realization. If the degree of Min influence is negatively correlated with tonal register [8, 11], one would expect the tonal targets of TM varieties that are more influenced by Min (*i.e.*, the nonstandard varieties) to be lower than those that are not as influenced (*i.e.*, the standard variety).

Secondly, we would like to see if sentential T1 demonstrates a similar interaction with stress as the other tones. Fon & Hsu [8] showed that when T2 and T3 are placed in sentence-final positions, H targets are realized higher and L targets lower than what would be expected from pure declination. We suspected that the exaggeration in the realization of stress might be due to a sentence-final stress rule [2, 13]. Therefore, we would like to see if such a trend could also be observed in T1. If so, then sentence-final T1s should be realized higher than what would be predicted by the neutral topline.

The third aim is to investigate possible effects of syllable structure on T1 realization. According to Hombert, Ohala, and colleagues [9, 10, 12], voiceless obstruents impose a slight pitch-raising effect on the F_0 values. However, this was not found in the realization of T2 and T3 [8]. One possible reason might be the constraints imposed by contour tones. Therefore, we would like to see if level T1 is also impervious to such effects.

Finally, according to the PENTA model [17], the default tonal register in utterance-initial positions should be mid unless otherwise specified (p. 240). Our previous findings [8] could not find affirmative evidence for this claim for TM T2. Thus, we would like to see if such pattern could be observed in T1. The model would predict T1 to be always realized as a mid-to-high rise in isolation and in sentence-initial positions, and as a low-to-high rise in other sentence-internal mismatch positions, as both would be considered as tonal mismatch cases in the PENTA model.

3. METHODS

3.1. Participants

Six subjects between ages 19 and 24 participated in the study. Half were from Taipei (the northern dialect), and half were from Taichung (the central dialect). All of them were ethnically Min, but the Taipei speakers could not speak Min fluently. As this study is still in progress, more subjects will be included when the project is complete.

3.2. Stimuli

27 T1 syllables representative of Mandarin phonotactics were chosen as stimuli, including 6 voiceless obstruent-initials (e.g., [xan] 'charmingly naive'), 15 sonorant-initials (e.g., [la] 'pull'), and 6 vowel-initials (e.g., [u] 'black'). Syllables were also placed in three comparable carrier sentences, so that they occurred in sentence-initial, -medial, and -final positions. Carrier sentences were designed so that syllables immediately before the target in medial and final positions ended mid-low and comparable tonal target clashes would occur in all three positions. In total, 27 (stimuli) × 3 (positions) = 81 sentences were recorded.

3.3. Equipment

Recordings were done using a SONY PCM-M1 Digital Audio Recorder with Maxell R-64 DA 60 min DAT tapes and a SHURE SM10A head-mounted microphone.

3.4. Procedure

Speakers were seated in a quiet room and asked to read out loud the semi-randomized stimuli using natural intonation at a normal rate. The whole process took about 15 minutes. The original recordings had a sampling rate of 48 kHz, which were subsequently downsampled to 16 bit 22050 kHz using Cool Edit Pro 2.00.

3.5. Analyses

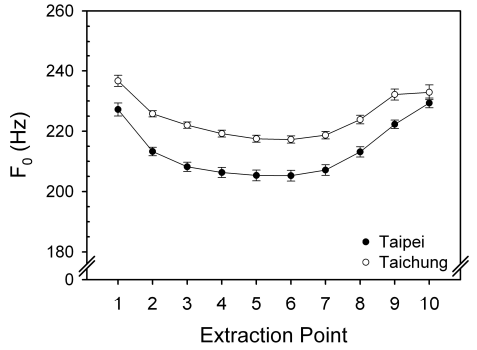
The recordings were hand-labeled using Praat 4.4 [1]. A Praat script was written for automatic pitch extraction on the voiced portion of each syllable, which is considered the measurable domain for tones [2, 3, 14, 15]. For obstruent-initial syllables, the starting point of a tone was determined by the onset of the voice bar after the obstruent, which was voiceless in this study. For the rest, the starting point began from the onset of the syllable, as the whole syllable was voiced. The ending point was always the offset of the voice bar. Occasional syllable-initial or -final glottalized portions caused by voice fry were not included for pitch extraction. Extracted pitch tracks were hand-checked and hand-corrected for doubling and halving through pitch period calculation, and were interpolated and smoothed using Praat functions afterwards. A second Praat script was written to extract pitch reference points at ten equal time points.

4. RESULTS

4.1. T1 in isolation

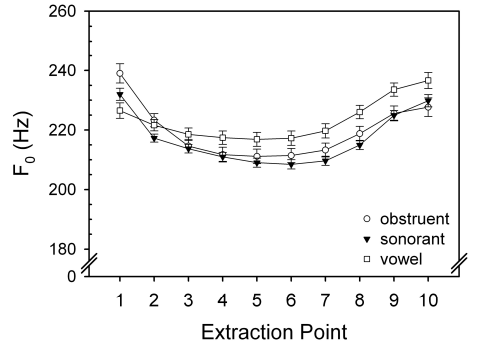
The average F_0 for Taipei speakers was 215.05 Hz while that for Taichung speakers was 225.10 Hz. A Dialect (2) × Onset (3) × Extraction (10) three-way mixed ANOVA was performed to test the effect of dialect and onset. Results showed that all of the main effects were significant [Dialect: $F(1,156) = 22.89$, $p < .0001$, $\hat{\eta}^2 = .13$; Onset: $F(2, 156) = 3.81$, $p < .05$, $\hat{\eta}^2 = .05$; Extraction: $F(2.76, 430.44) = 143.49$, $p < .0001$, $\hat{\eta}^2 = .48$]. Two of the two-way interaction effects involving Extraction were also significant [Dialect × Extraction: $F(2.76, 430.44) = 5.08$, $p < .01$, $\hat{\eta}^2 = .03$; Onset × Extraction: $F(5.52, 430.44) = 8.21$, $p < .0001$, $\hat{\eta}^2 = .10$]. The three-way interaction was not significant (Figures 1 & 2).

Figure 1: Time-normalized F_0 trajectories of isolated T1 in two dialects averaged across onset types.



Post hoc independent *t*-tests regarding the interaction effect of Dialect and Extraction showed that except for the final extraction point, all other extractions were significant ($p < .01$ for Point 1, and $p < .001$ for others). In addition, post hoc pairwise comparisons showed that for northern TM, Points 1, 9, and 10 were the highest in pitch, Points 2 and 8 were the second highest, the remaining points were the lowest ($p < .0001$). For central TM, Points 1, 9, and 10 were still the highest in pitch, Points 2, 3, and 8 were the next highest, and the remaining points were the lowest ($p < .0001$).

Figure 2: Time-normalized F_0 trajectories of T1 in isolation with regards to onset types.

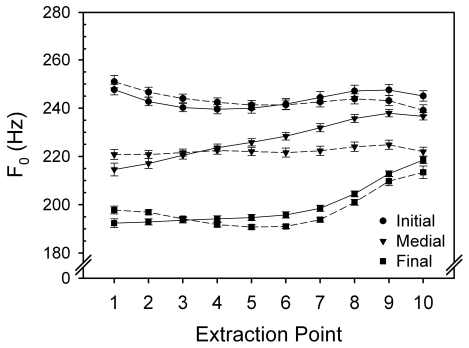


Post hoc one-way ANOVAs regarding the interaction effect of Onset and Extraction showed that Onset was only significantly different at Points 1 and 7 [Point 1: $F(2, 78) = 5.99$, $p < .01$, $\hat{\eta}^2 = .13$; Point 7: $F(2, 78) = 3.47$, $p < .05$. $\hat{\eta}^2 = .08$]. For Point 1, post hoc Tukey-*b* tests showed that obstruent-initial syllables were higher than vowel- and sonorant-initial ones ($p < .05$). For Point 7, post hoc pairwise comparisons showed that vowel-initial syllables were the highest while sonorant-initial ones were the lowest ($p < .05$).

4.2. T1 in context

The average F_0 values for Taipei speakers were 245.65 Hz, 230.20 Hz, and, 200.71 Hz for the three positions, respectively, while those for Taichung speakers were 244.26 Hz, 224.24 Hz, and 198.93 Hz, respectively. A Dialect (2) × Position (3) × Extraction (10) three-way mixed ANOVA was performed to test the effect of dialect and sentential positioning. Since Onset did not seem to have a very robust effect on isolated T1, it was excluded in the following analyses. Results showed that two of the main effects were significant [Position: $F(2, 320) = 730.78$, $p < .0001$, $\hat{\eta}^2 = .82$; Extraction: $F(1.88, 300.22) = 74.60$, $p < .0001$, $\hat{\eta}^2 = .32$]. Two of the two-way interactions involving Extraction were also significant [Dialect × Extraction: $F(1.88, 300.22) = 23.67$, $p < .0001$, $\hat{\eta}^2 = .13$; Position × Extraction: $F(3.36, 537.44) = 73.66$, $p < .0001$, $\hat{\eta}^2 = .32$]. So was the three-way interaction [$F(3.36, 537.44) = 5.80$, $p < .001$, $\hat{\eta}^2 = .03$] (Figure 3).

Figure 3: Time-normalized F_0 trajectories of T1 in context. Solid lines represent Taipei speakers, and dashed lines represent Taichung speakers.



Regarding the declination effect, post hoc pairwise comparisons showed that for northern TM, all sentence-initial extraction points were higher than sentence-medial ones, which were in turn higher than sentence-final ones. However, the difference between the latter two positions were much larger than the former two, especially in the final portions of the tone ($p < .001$ between initial and medial Point 9's, $p < .01$ between initial and medial Point 10's, and $p < .0001$ for others). For central TM, the overall trend was still the same. Sentence-initial extracts were the highest, and sentence-final ones were the lowest. However, the difference between the former two was much larger than that between the latter two for the final portion of the tone ($p < .05$ for medial and final Point 10's, and $p < .0001$ for others).

As for dialectal differences, post hoc pairwise

comparisons showed that for sentence-medial positions, Taipei T1 was significantly higher than that of Taichung starting from Point 6 to the end of the tone ($p < .01$ for Point 6, $p < .001$ for Point 7, and $p < .0001$ for others). For sentence-final positions, Taipei T1 was lower than Taichung T1 at Point 1 ($p < .05$), but was significantly higher from Points 5 to 8 ($p < .05$ for Points 5 & 8, and $p < .01$ for Points 6 & 7). No difference was found in the initial position.

With regards to tonal contours, Taipei and Taichung initial T1s and Taichung medial T1s were fairly level. Taipei medial and final T1s were rising, while Taichung final T1s were dipping.

5. DISCUSSION

Results in this study showed that variations did exist between the two varieties of TM. In terms of pitch register, the direction went as predicted in sentence-medial and -final positions. Taichung T1 was indeed lower, and the rate of declination faster. However, syllables in isolation showed an opposite trend, in which Taipei T1 was lower. Since reading isolated syllables is more unnatural and thus more formal than reading sentences, we hypothesized that the Taichung speakers, speaking a non-standard dialect, might be unconsciously over-correcting themselves in a more formal register, but were unable to do so in a more relaxed one.

The effect of sentential stress in raising sentence-final H tonal targets was also supported, as can be shown by the bigger rise towards the end of the syllable. The effect of sentential stress thus affects not only contour tones, but also level ones.

Different onset types did have an effect on the realization of tones. Obstruent-initial syllables had slightly higher pitch than sonorant-initial ones. However, the effect was fairly small and localized.

Finally, isolated and sentence-initial T1s were not realized as a rise, which contradicted Xu's [17] claim of a default mid tonal register. The only contours that conformed to the PENTA model were the medial and final Taipei T1s. However, Taichung tones did not show this effect.

6. CONCLUSION

This study shows that dialectal differences can affect realization of Mandarin T1. Phonotactic composition, while significant, imposes only minor effects. More studies will be needed in order to understand the actual mechanism underlying tonal realization.

7. REFERENCES

[1] Boersma, P., Weenink, D. 2006. *Praat: doing phonetics by computer* version 4.4.18. http://www.praat.org/.

[2] Chao, Y.R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

[3] Chao, Y.R. 1956. Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese. In: Halle, M. (ed), *For Roman Jakobson*, The Hague: Mouton, 52-59.

[4] Chen, S.H. 2005. The effects of tones on speaking frequency and intensity ranges in Mandarin and Min dialects. *Journal of the Acoustical Society of America* 117(5), 3225-3230.

[5] Chen, Y.-D. 1989. *Taiwan de Kejiaren [Hakka on Taiwan]*. Taipei: Taiyuan Press.

[6] Cheng, R.L. 1985. A comparison of Taiwanese, Taiwan Mandarin, and Peking Mandarin. *Language* 61(2), 352-377.

[7] Fon, J., Chiang, W.-Y. 1999. What does Chao have to say about tones? -A case study of Taiwan Mandarin. *Journal of Chinese Linguistics* 27(1), 15-37.

[8] Fon, J., Hsu, H. 2007. Positional and phonotactic effects on the realization of dipping tones in Taiwan Mandarin. In: Gussenhoven, C. and Riad, T. (eds), *Phonology and phonetics, Tones and tunes: Phonetic and behavioural studies in word and sentence prosody*. Vol. 2, 239-269.

[9] Hombert, J.-M. 1978. Consonant types, vowel quality, and tone. In: Fromkin, V. A. (ed), *Tone: A Linguistic Survey*, New York: Academic Press, 77-111.

[10] Hombert, J.-M., Ohala, J.J., Ewan, W.G. 1979. Phonetic explanation for the development of tones. *Language* 55(1), 37-58.

[11] Hsu, S.-y. 2006. *The tonal variation of Mandarin Tone 2 in Taiwan: A phonetic study on Taiwanese-Taiwan Mandarin bilinguals and Taiwan Mandarin monolinguals*. National Chengchi University.

[12] Ohala, J.J. 1973. The physiology of tone. In: Hyman, L. M. (ed), *Consonant Types and Tone: Southern California Occasional Papers in Linguistics*. Vol. 1, Los Angeles: University of Southern California, 2-14.

[13] Peng, S.-h., Chan, M.K.M., Tseng, C.-y., Huang, T., Lee, O.J., Beckman, E.M. 2005. Towards a pan-Mandarin system for prosodic transcription. In: Jun, S.-A. (ed), *Prosodic Typology: The phonology of intonation and phrasing*, New York: Oxford University Press, 230-270.

[14] Wang, W., S-Y. 1967. Phonological features of tone. *International Journal of American Linguistics* 33(2), 93 - 105.

[15] Xu, Y. 1998. Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica: International Journal of Speech Science* 55(4), 179-203.

[16] Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25(1), 61-83.

[17] Xu, Y. 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46(3-4), 220-251.