THE ROLES OF DURATION, RHYME STRUCTURE AND FREQUENCY IN MANDARIN ACCIDENTAL GAPS

National Chiao Tung University

Shao-Jie Jin and Yu-An Lu

courtney0419003.fl04@nctu.edu.tw & yuanlu@nctu.edu.tw

ABSTRACT

A corpus study investigating the distribution of accidental gaps (i.e., syllable-tone combinations that could but do not exist) in Mandarin showed that T2 (rising) and T3 (falling-rising) gaps were over-represented, whereas T1 (high level) and T4 (falling) gaps were under-represented. Since the realization of T2/T3 intrinsic contours requires a longer duration, we hypothesized that T2/T3 might be more acceptable in longer durations and complex rhymes. A wordlikeness rating experiment was conducted in which gaps as well as real words in different syllable structures (open vs. closed) were manipulated into two durations (300 and 500 ms). The results revealed that Mandarin listeners accepted T1/T4 gaps more readily than T2/T3 gaps, regardless of duration and rhyme structure. The wordlikeness ratings of gaps were positively correlated with syllable frequencies, a pattern also found for real words. These findings were attributed to the marked status of T2/T3 and to syllable frequency.

Keywords: accidental gaps, frequency, Mandarin, tonotactics

1. INTRODUCTION

Mandarin is a tone language with four phonemic tones (i.e., high-level Tone 1 [X^{55}], rising Tone 2 [X^{35}], falling-rising Tone 3 [X^{214}], and falling Tone 4 [X^{51}]). The maximum syllable structure is (C)(G)V(G)/(C) (c.f., [6]). However, not all syllables are combined with each of the tones. For example, there are words combining the syllable [tsʰu] with T1 ([tsʰu]^{55} "coarse"), T2 ([tsʰu]^{35} "die" in Old Chinese), and T4 ([tsʰu]^{51} "vinegar"), but not with T3. The syllable-tone combination [tsʰu]^{214} does not violate any obvious phonotactic constraints in Mandarin, yet it fails to exist—this is an example of a " accidental gap" [3, 5]. Wang [11] further identified three types of gaps, including tonotactic accidental gaps (the type of gaps mentioned earlier), phonotactic accidental gaps (phonotactically legal syllables but fail to exist), and systematic gaps (phonotactically illegal syllables). In a wordlikeness rating task, Wang elicited wordlikeness judgments from native Mandarin speakers judged tonotactic accidental gap syllables to be more wordlike, followed by phonotactic accidental gap then systematic gap syllables. Myers & Tsay [8] conducted a similar study on Mandarin words and non-words and found three factors that influence wordlikeness judgments: frequency, phonotactics, and neighborhood density. They pointed out that frequency and neighborhood density are only relevant in the judgement of real words, while phonotactics affects both words and non-words.

In order to gain a better understanding of Mandarin tonotactic accidental gaps, we conducted a corpus study and found that T2 and T3 gaps were over-represented among all allowable syllables. Based on this finding and a general typological preference for contour tones with complex/longer rhymes, we hypothesized that, in a wordlikeness rating task, T2/T3 gaps might be judged as more acceptable in syllables with longer durations and complex rhymes. Furthermore, previous findings [5, 7] suggesting that frequency affects Mandarin speakers' judgments of real words led us to examine whether syllable frequency (regardless of tone) would also affect the word judgments of tone-syllable combinations with zero frequency. A wordlikeness rating experiment was conducted to test these hypotheses.

2. CORPUS STUDY

We first investigated the distribution of the accidental gaps among the 388 allowable Mandarin syllables (taken from Appendix B in [6]). We implemented two definitions of accidental gaps: (1) syllable-tone combinations that do not form a word ("broad view"), and (2) syllable-tone combinations that do form words but have zero-frequency in the *Sinica Chinese Spoken Syllable Structure List* [10] ("narrow view"). For example, the syllable [tsʰu] in T2 ([tsʰu]^{35} "die" in Old Chinese), historically exists as a word but is not on the list. As such, this word might be considered a gap by native Mandarin speakers because it is rarely used in spoken Mandarin. This word was thus counted as a gap in the narrow view but not in the broad view.

The results showed that accidental gaps were not evenly distributed across the four tones, as shown in

Figure 1. In the broad view, a one-way chi-square test showed that T2 gaps were over-represented while T4 gaps were under-represented ($\chi^2(3) = 78.25$, $p<.001$). Another one-way chi-square test revealed that both T2 and T3 gaps were over-represented in the narrow view ($\chi^2(3) = 27.19$, $p<.001$).

Figure 1: Mandarin accidental gaps in broad and narrow views.
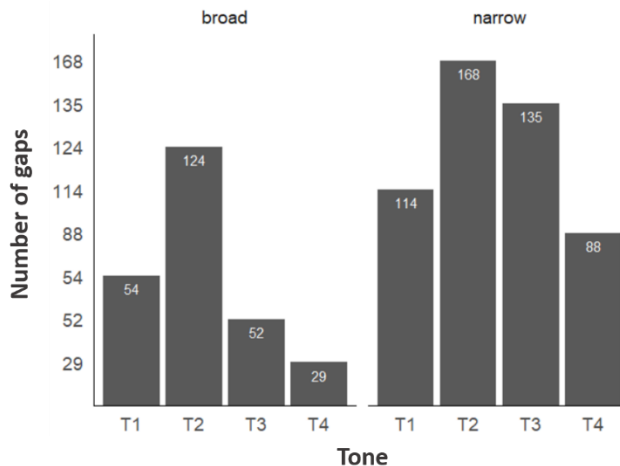
The asymmetrical distribution of accidental gaps may be attributed to *shǎng sheng rù píng*, a historical tone merging process in which a number of *shǎng* tones (i.e., T3) merged into *píng* tones (T4) in Middle Chinese [7]. Secondly, the large percentage of T2 and T3 gaps could be attributed to the markedness of rising and complex contour tones (i.e., T2 and T3) in comparison with level tones and falling tones (i.e., T1 and T4).

We further divided syllables into different types (CV, CGV, CVN, CGVN) to examine if there were fewer T2/T3 gaps in closed syllables which have intrinsically longer durations than open syllables. However, we did not observe any tendencies in this direction, as demonstrated in Figure 2. Instead, the same trend was observed across the different syllable types. For each tone group, there were generally more gaps with CV and CVN structures than with the other syllable types.

Figure 2: Mandarin accidental gaps across different syllable types
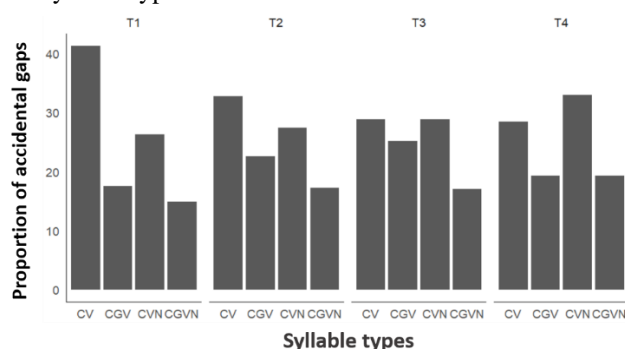
To investigate whether duration, rhyme structures, and frequency, a factor that was shown to affect Mandarin speakers' judgments of real words, influence native Mandarin speakers' word judgments of accidental gaps, we conducted a wordlikeness judgement experiment.

3. EXPERIMENT

3.1. Methodology

3.1.1. Participants

We recruited 22 Taiwan Mandarin native speakers (7M, 15F; age 20-37, *M*=22.05) from National Chiao Tung University. None of the participants reported hearing or speaking deficiencies. All participants were compensated monetarily for their time.

3.1.2. Materials

To examine whether T2/T3 gaps would be judged as more wordlike in complex rhymes, 96 Mandarin accidental gaps were selected with the four Mandarin tones and were counter-balanced across different syllable types (open: CV, CGV; closed: CVN, CGVN). Another 48 Mandarin real words fulfilling the same criteria were selected as filler items. The list of 144 stimuli were produced by a male native Mandarin speaker. To investigate whether the T2/T3 gaps would be judged as more wordlike in longer durations, the stimuli were resynthesized into two durations (300 ms and 500 ms) using the Pitch Synchronous Overlap and Add (PSOLA) algorithm in Praat [2].

3.1.3. Procedure

The 288 resynthesized stimuli (144 x 2 durations) were randomized for each participant and presented in three blocks using E-Prime [9]. The participants were instructed to rate each word on a 7-point scale, with 7 being the most Mandarin-like, and 1 being the least Mandarin-like.

Nine practice trials were presented before the experiment to familiarize participants with the task. The experiment was conducted in a sound-attenuated booth using AKG K240 headphones. The total duration of the experiment was around 15 minutes.

3.2. Results

To interpret the results, linear mixed-effects regression models were fitted in R using the lme4 package [1]. Models were fitted with the participants' wordlikeness ratings on the 7-point scale converted into *z*-scores as the dependent variable. Tone (T1, T2,

T3, T4), Duration (300 ms, 500 ms), and Syllable Type (CV, CGV, CVN, CGVN) were taken as the fixed effects. Participant and Token were taken as the random effects. The full model included the interaction terms for Tone, Duration, and Syllable Type and the intercepts for the random effects. The main effects were obtained by comparing the full model with a simpler model without the target variable. Interactions were obtained similarly by comparing models with and without the interaction term. The results showed a main effect of Tone ($p<.001$). Post-hoc tests were conducted by changing the baselines, and the results showed that T3 and T2 stimuli were rated as less Mandarin-like compared with T1 and T4 stimuli (all $p<.001$), as shown in Figure 4.

**Figure 4**: Wordlikeness ratings for accidental gaps by tones
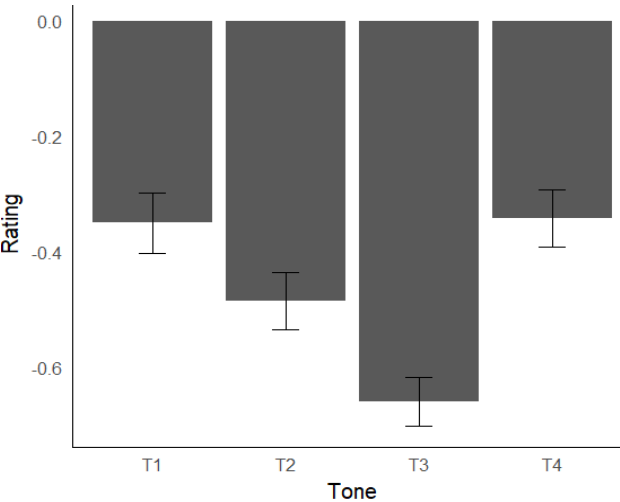
Duration was also shown to have an effect ($p<.001$). Specifically, syllables with shorter durations (300 ms) were overall rated as more Mandarin-like than those with longer durations (500 ms) ($p<.001$). The Duration effect was only found on accidental gaps but not on real words. These results are illustrated in Figure 5. Syllable Type, on the other hand, did not appear to have a significant effect.

Figure 5: Wordlikeness ratings for real words and accidental gaps (with different scale) by Tone and Duration
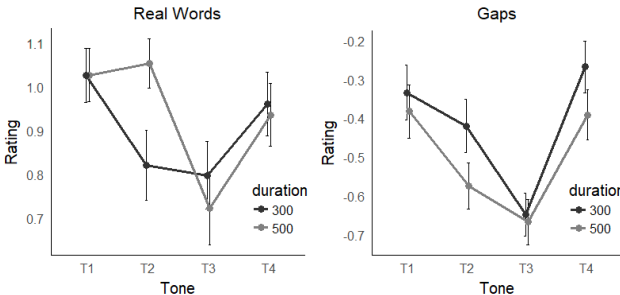

Data were also analyzed to see if there was a frequency effect. First, we determined how many tone-syllable combinations were possible for each syllable. If a syllable could combine with all but one tone, it was counted as "1", all but two tones, "2", and so on. Following this logic, with only one gap in a certain syllable, "1" would be considered more "frequent" while "3", with three gaps, would be considered less "frequent". Note that there is no "4" because in this case all tone-syllable combinations would be illegal. The results (Figure 6) showed a significant difference across syllables with different numbers of tonal gaps ($p<.001$), indicating that when a gap occurs in a syllable involves more tonotactic gaps, it is considered to be less wordlike.

Figure 6: Syllables with different numbers of accidental gaps
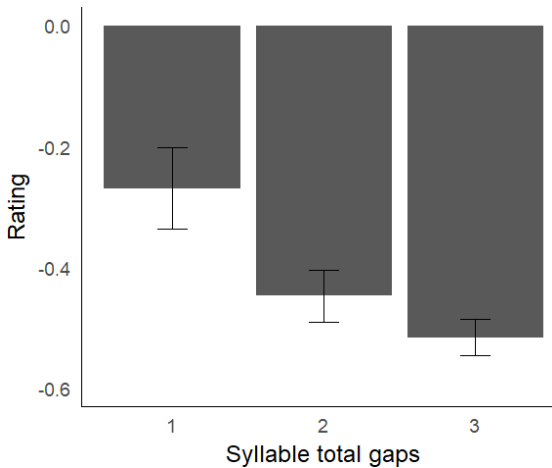
Second, we calculated the overall frequency of each syllable regardless of tone. Based on these frequencies, we found significant positive correlations between the total syllable frequency and the wordlikeness ratings of both real words and gaps. (both $p< .0001$), as shown in Figure 7.

Figure 7: Correlation between wordlikeness ratings and syllable frequency for words and accidental gaps
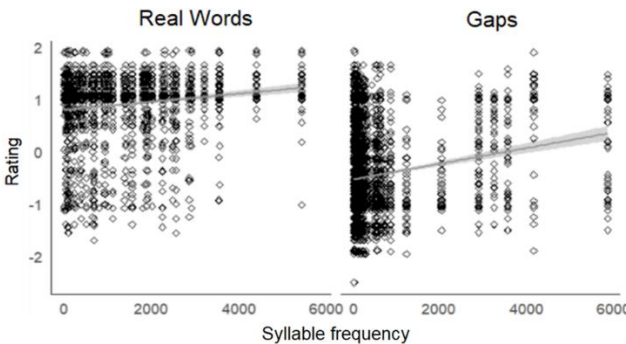
4. GENERAL DISCUSSION

A corpus study investigating the distribution of accidental gaps in Mandarin showed a disproportionally large number of T2 (rising) and T3 (falling-rising) gaps compared with T1 (high level) and T4 (falling) gaps. We hypothesized that T2/T3 gaps might be judged as more wordlike by native Mandarin speakers in longer durations and complex rhymes based on a typological preference for contour tones with complex/longer rhymes. Also, we predicted that accidental gaps would be more acceptable when in syllables with greater frequency.

The results of the wordlikeness rating experiment showed Tone, Duration, and Frequency effects on the speakers' word judgments of accidental gaps. Although Syllable Type interacted with Tone and Duration, it did not pattern with typological preferences nor did it present a pattern that could be generalized.

Regarding the effect of Tone, T2 and T3 gaps generally elicited lower wordlikeness ratings than did T1 and T4 gaps. This correlates with our corpus data and the universally marked status of rising and complex contour tones. Another possible factor is that T2 and T3 alternate in a sandhi process whereby T3 becomes T2 before another T3. The confusability between the two tones may have also attributed to the low wordlikeness ratings of T2 and T3 [4].

The Duration effect indicated that gap words with shorter durations were significantly more acceptable than those with longer durations, regardless of tone. The same effect, however, was not found on real words. This suggests that Mandarin listeners were more reluctant to accept accidental gaps with longer durations across the board, presumably due to clearer audibility.

Finally, contrary to [8], we found a positive correlation between wordlikeness ratings and the frequency of the syllables hosting the accidental gaps. That is, when a gap occurs in a more frequent syllable, it is more likely to be accepted as a Mandarin-like word. This may be due to the fact that the "non-words" in [8] include phonotactically legal accidental gaps [11] as well as illegal non-words. Unlike phonotactic gaps which are illegal syllables and thus have zero syllable frequency, tonotactic accidental gaps could still have frequency effects from the syllable in which they occur. Our results suggest that Mandarin speakers can access the syllable frequency in judging accidental gaps.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] Bates, D., Maechler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* 67, 1–48.

[2] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot Int.* 5:9/10, 341–345.

[3] Duanmu, S. (2007). The phonology of standard Chinese. Oxford University Press.

[4] Hume, E. and K. Johnson (2003). The impact of partial phonological contrast on speech perception. *15th International Congress of Phonetic Sciences*. 2385-2388.

[5] Lai, Y. (2003). A Perceptual Investigation on Mandarin Tonotactic Gaps. MA thesis. National Tsing Hua University.

[6] Lin Y-H. (2007) *The Sounds of Chinese*, Cambridge, UK: Cambridge University Press.

[7] Mei T-L. (1982) On the Rising Tone. *Tseng Hua Journal of Chinese Studies* 14: 233-241.

[8] Myers, J. and Tsay, J. (2004) Exploring performance-based predictors of phonological judgments in Mandarin. Poster presented at Laboratory Phonology 9.

[9] Schneider, W., Eschman, A., & Zuccolotto, A. 2002. *E-Prime User's Guide.* Pittsburgh: Psychology Software Tools Inc.

[10] Tseng, S.-C. 2013. Lexical coverage in Taiwan Mandarin conversation. *International Journal of Computational Linguistics and Chinese Language Processing* 18:1-18.

[11] Wang HS. (1998) An experimental study on the phonotactic constraints of Mandarin Chinese. *Studia Linguistica Serica*: 259-268.