

Intonation-Induced F0 Realizations of Rising and Falling Lexical Tones in Mandarin Questions

Xinyuan Shi^{1*}, Honglan Li²

¹School of Chinese Language and Literature, Nanjing Normal University, Nanjing, China

²School of Foreign Studies, Nanjing University of Science and Technology, Nanjing, China

Email: *xinyuan.shi@njnu.edu.cn

How to cite this paper: Shi, X. Y., & Li, H. L. (2025). Intonation-Induced F0 Realizations of Rising and Falling Lexical Tones in Mandarin Questions. *Open Journal of Modern Linguistics*, 15, 668-684.
<https://doi.org/10.4236/ojml.2025.154039>

Received: July 18, 2025

Accepted: August 16, 2025

Published: August 19, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This study examines how Mandarin question intonation differentially modifies the fundamental frequency (F0) of syllables with inherent rising (T2) and falling (T4) lexical tones. By utilizing carefully designed test materials and conducting a production experiment, we measured time-normalized F0 values, maximum F0 values and minimum F0 values, and compared mean F0 values and excursion sizes across syllable positions. Results demonstrate that first, while both tones showed global F0 raising, T2 exhibited raising at penultimate and final syllables, whereas T4 only exhibited final-syllable raising. Second, T2 displayed range expansion contrasting with T4's range compression. Third, range expansion of T2 occurs at peripheral positions whereas range compression of T4 concentrated in sentence-final syllable. Consistent sentence-initial syllable prominence across intonation types was identified. These findings reveal tone-specific intonational manipulations, showing global and local F0 modifications on lexical rising versus falling contours in Mandarin questions.

Keywords

Intonation, F0 Realization, Lexical Tone, Mandarin Chinese

1. Introduction

Compared to statements, Mandarin questions are typically considered a marked sentence type (Chao, 1968) and are generally more difficult to identify (Liu et al., 2022; Xu, 2013), as they can only be recognized if listeners detect the presence of question-specific features (Yuan, 2006, 2011; Yuan & Shih, 2004). Syntactically, Mandarin marks information-seeking yes-no questions by adding final particles

such as *ma*. Phonologically, intonation serves as a crucial mechanism for distinguishing sentence types that share identical lexical sequences. Intonation employs prosodic features, particularly fundamental frequency (F0), to convey sentence-level information such as sentence type (Gussenhoven, 2004; Ladd, 2008). Cross-linguistically, it is well established that languages use high/rising F0 to mark questions (Bolinger, 1982; Cruttenden, 1981; Ohala, 1983). The current study focuses on Mandarin syntactically unmarked yes-no questions, which are implemented by phonological intonation patterns.

The implementation of intonation in tone languages is more constrained than in non-tone languages, as lexical tones rely on F0 to distinguish meanings at the word level. Mandarin Chinese has four lexical tones: T1 (high level), T2 (rising), T3 (low/dipping), and T4 (falling), which differentiate word meanings (e.g., *ma1* “mother”, *ma2* “hemp”, *ma3* “horse”, and *ma4* “scold”, where the italicized letters represent pinyin forms and the numbers indicate Mandarin tone categories) (Duanmu, 2007; Zhang et al., 2021).

The intonation-induced manipulation of F0 in Chinese syntactically unmarked yes-no questions has garnered considerable attention, particularly due to the interaction between lexical tones and intonation. In his pioneering work on Chinese intonation, Chao (1929, 1933, 1968) described the interplay between lexical tones and intonation as “small ripples riding on large waves”. This metaphor illustrates that sentence-level F0 contours are shaped by both local lexical tones and global intonation patterns. In questions, while the canonical F0 contours of lexical tones tend to retain their original shapes (Cao, 2002; He & Jing, 1992; Ho, 1977; Shen, 1989; Wu, 1996; Yang & Kong, 2020; Zhang et al., 2021), a global F0 rise occurs across the entire utterance (Gårding, 1987; Liu & Xu, 2005; Ni & Kawai, 2004; Shen, 1985, 1990, 1994; Yuan et al., 2002). However, Chao’s algebraic sum proposal suggests that the tone contour of the sentence-final syllable undergoes modification. For instance, a high-falling tone in this position may become flattened due to the raised final tone target. Consequently, intonation-induced F0 manipulation appears to be realized not only through global pitch raising but also through local adjustments at specific points within the question (Yuan, 2006).

Numerous productive studies have highlighted localized F0 changes when comparing intonation patterns between unmarked yes-no questions and their statement counterparts. Early research identified a higher starting point in questions as a distinctive feature. For instance, in her instrumental study of trisyllabic utterances, Shi (1980) examined F0 patterns in trisyllabic statements and questions with final syllables carrying the four lexical tones, observing a notable initial F0 rise in questions. Subsequent studies consistently reported this elevated starting point, proposing register differences at utterance onset as a key distinguishing feature between statements and questions (Shen, 1990). Further investigations revealed significant F0 modifications beginning from medial positions. Starting from the middle word, questions exhibited expanded F0 ranges in subsequent syllables, with particularly pronounced elevation of the upper F0 limit (Xie, 2021).

Additional research demonstrated question-induced F0 changes concentrated in final phrases. For example, the most substantial F0 rises and range expansions were observed in the terminal noun phrases of yes-no questions (Lee, 2005), with the critical falling or rising contours distinguishing questions from statements localized to these final phrases (Gårding, 1987). Most notably, studies found F0 modifications became increasingly pronounced toward utterance endings. The sentence-final syllable consistently showed the most dramatic F0 increases, steepest slope changes, and greatest range expansions, confirming the cumulative nature of question intonation effects in Mandarin (Jiang, 2010; Lin, 2006; Wang et al., 2017).

In summary, the intonation-induced manipulation of F0 in Chinese syntactically unmarked yes-no questions has attracted significant research interest, primarily due to the complex interaction between lexical tones and intonation. While intonation induces a global F0 modification evident in the overall raised F0 contour of unmarked yes-no questions, the crucial acoustic cues for question intonation appear to emerge from distinct F0 variations occurring at specific positions within the utterance. Mandarin Chinese presents a particularly interesting case, where T2 (a rising lexical tone) and T4 (a falling lexical tone) interact with question intonation that is characteristically marked by both an overall higher F0 register and a terminal rise in sentence-final syllables. Given that both lexical tone distinctions and intonational meaning are primarily encoded through F0 variations, the current study seeks to investigate two key research questions regarding question-induced F0 patterns of T2 and T4. First, how are the F0 realizations of rising (T2) and falling (T4) lexical tones affected by question intonation? More specifically, how is the intonation-induced global and local rise implemented on syllables whose lexical tones already possess inherent rising or falling contours? Second, particularly in pre-final syllables, does question intonation exert equivalent F0 modifications on rising versus falling lexical tones across different syllable positions? To address these questions, the study employed a production experiment methodology.

2. Methodology

2.1. Participants

Twenty-eight native Mandarin-speaking undergraduates aged from 20 to 21 years old (24 females, $M_{\text{age}} = 20.32$ years, $SD_{\text{age}} = 0.61$ years) participated in the production experiment. All participants reported using Mandarin Chinese as their primary language in daily communication. No participant had any neurological or physical impairment, or speech or hearing impairments.

2.2. Material

The experimental materials were designed to examine F0 modifications induced by question intonation on T2 and T4 in Mandarin Chinese, compared to default statement intonation. We constructed four six-syllable multi-tone sentences, re-

fining materials from previous Mandarin intonation production experiments (Qin, 2017a, 2017b). Each test sentence comprised three words: a disyllabic proper name subject (S), a monosyllabic verb (V), and a trisyllabic common noun object (O). All objects were verb-noun compounds consisting of a monosyllabic verb and disyllabic noun. The four-sentence stimulus set was designed for the following reasons. First, multi-tone sentences better reflect natural speech patterns compared to single-tone utterances. Rather than constructing separate single-tone SVO sentences such as *Wang2hao2 chang2 zha2yu2* “Wanghao tasted fried fish” and *Zhao4hao4 zuo4 dun4rou4* “Zhaohao made meat stew”, we distributed all four lexical tones across corresponding syllable positions within each sentence, ensuring that the four syllables at every syllable position contained all four lexical tones (see **Table 1**). Second, T3 sandhi takes place in single-tone sentence like *Li3bao3 dian3 kao3sun3* “Libao ordered grilled bamboo shoots”, while multi-tone sentences effectively prevented T3 sandhi. In particular, our stimulus set maintains tonal stability with no T3 sandhi occurring either within or across word boundaries.

The test sentences were embedded into four simple conversations, and each conversation comprised two intonation types of one test sentence (see **Table 2** for example). To answer the question raised by Speaker A, Speaker B gave the statement version of the testing sentence. After that, Speaker A gave an echo question, i.e., the question version of the test sentence. The experiment also included four filler sentences. Each filler consisted of a disyllabic subject, a disyllabic adverb, a monosyllabic verb, and a disyllabic object. Thus, four testing sentences and four fillers were embedded in eight conversations (see **Table A1** and **Table A2** in **Appendix**).

Table 1. T2 and T4 embedded in four test sentences.

Tone	Subject	Verb	Object	Translation
T1 T4 T4 T3 T4 T3	<i>zhang1hao4</i>	<i>zuo4</i>	<i>kao3rou4bing3</i>	“Zhanghao made grilled meat pie.”
T2 T3 T2 T4 T2 T2	<i>wang2bao3</i>	<i>chang2</i>	<i>dun4yu2tou2</i>	“Wangbao tasted fish head stew.”
T3 T1 T3 T2 T1 T4	<i>li3chao1</i>	<i>dian3</i>	<i>zha2ji1chi4</i>	“Lichao ordered fried chicken wings.”
T4 T2 T1 T1 T3 T1	<i>zhao4hao2</i>	<i>chi1</i>	<i>shao1sun3gan1</i>	“Zhaohao ate cooked bamboo shoots.”

Table 2. A test sentence with statement and question intonation in a simple conversation.

Speaker	Pinyin form	Translation
A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	“What did you say?”
B	<i>zhang1hao4 zuo4 kao3rou4bing3.</i>	“Zhanghao made grilled meat pie.”
A	<i>zhang1hao4 zuo4 kao3rou4bing3?</i>	“Did Zhanghao make grilled meat pie.”
B	<i>shi4de0.</i>	“Yes.”

a. “0” indicates the neutral tone.

2.3. Procedures

The elicitation was conducted in a remote manner. All participants were instructed to record their speech in quiet rooms using their smartphones or digital recorders. Participants were required to produce statements (marked with periods) or questions (marked with question marks) embedded in the simple conversations. The randomized conversations were given in a PDF file, which also contained written instructions for participants. All participants were instructed to maintain focus-neutral delivery, produce natural intonation patterns, and avoid exaggerated emotional prosody. Each two participants formed a small group, and they read the conversations twice. For the first time, one played as speaker A and the other played as speaker B; while for the second time, they played as the opposite speakers. After collection, all audio recordings were saved as WAV-format files with a 44,100 Hz sampling rate, underwent noise reduction processing using Audacity, and subsequently checked and segmented.

2.4. Data Analysis

Our analysis specifically targeted the rime portions of syllables bearing T2 and T4 tones within the test sentences. Each available recorded utterance was annotated and manually verified using Praat (Boersma & Weenink, 2023). We extracted F0 measurements by applying the ProsodyPro script (Xu, 2013). Specifically, we extracted time-normalized F0 values across 10 sample points, maximum F0, minimum F0 and mean F0, and excursion size. To normalize inter-speaker variations, all original F0 values in Hertz were converted to semitones with a reference of 100 Hz. The excursion size represents the difference between maximum F0 and minimum F0 in semitones, which was also extracted by the ProsodyPro script.

To examine question-induced F0 modifications on T2 and T4, we conducted multiple analyses with these F0 measurements. First, we utilized time-normalized F0, maximum F0, and minimum F0 values to plot F0 contours for each target tone at every syllable position across intonation types, and generate F0 curves representing the upper and lower limits of target tones across experimental conditions. This primarily descriptive approach provided a fundamental acoustic characterization of how rising and falling tones are realized in both statements and questions.

Second, we compared mean F0 and excursion size across conditions. Specifically, statistical analyses were performed using linear mixed-effects regression models with syllable position (6-level: 1st, 2nd, 3rd, 4th, 5th versus 6th syllable) and intonation type (2-level: question versus statement) as fixed factors, and with sentence and participant status as random factors. For mean F0, we were interested in whether the question intonation induced a globally raised F0 with an accelerating local sentence-final F0 rise. For excursion size, we were interested in whether question intonation induced different manipulations of F0 range on rising and falling tones. Statistical analyses were carried out with the packages *lmerTest* (Kuznetsova et al., 2017) and *bruceR* (Bao, 2021) in R version 4.4.1 (R Core Team,

2024), and all effects were reported as significant at $p < .05$.

3. Results

All participants completed the reading task without any confusion. Data from one participant was completely excluded due to low recording quality. After checking the remaining 216 sentences (4 sentences \times 2 intonation types \times 27 participants), six statements and four questions were further excluded because of either low quality or obvious narrow focus or emotion. Thus, 102 statements and 104 questions were segmented and annotated to obtain the mentioned F0 measurements for further analysis. We reported results of both descriptive and statistical analyses. For the main effect of interaction, the results of post-hoc simple comparisons among syllable positions in statement and question were summarized in **Table A3** (for mean F0 of T2), **Table A4** (for mean F0 of T4), **Table A5** (for excursion size of T2), **Table A6** (for excursion size of T4) in **Appendix**.

3.1. F0 Contour and Mean F0 Values

Figure 1 presents the time-normalized realizations of T2 and T4 across syllable position and intonation conditions. First, the contours globally raised in question compared to statement, no matter the lexical tone is T2 or T4. However, within a sentence, the raising degree was not consistent among the six syllable positions. Overall, F0 rise became accelerating at later portion of questions. Second, sentence-final T2 had expanded F0 range and increased slope, while T4 in this position had slightly flattened F0 contour. Third, tone contours of both T2 and T4 preserved in most syllable positions, with relative flattening in penultimate syllables in both intonation types. In addition, sentence-final T2 had a slight initial falling, either in statement or question intonation.

Mean F0 values in semitones were compared at six syllable positions in statements and questions. **Figure 2** illustrates the means and error bars of mean F0 of each tone across conditions. For T2, the statistical analysis revealed a significant main effect of syllable position, $\chi^2(5) = 128.46$, $p < 0.001$, a significant main effect

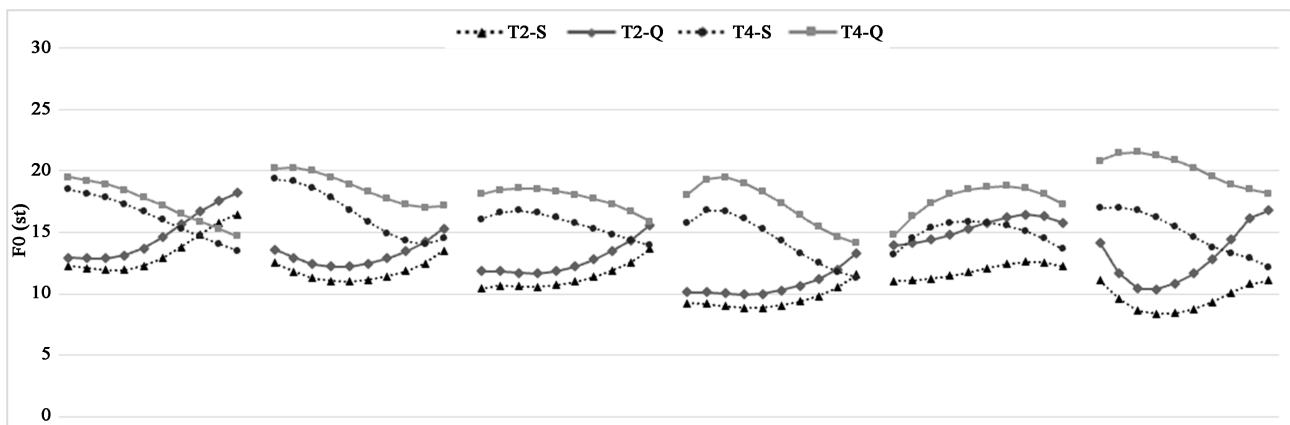


Figure 1. F0 contours of T2 and T4 in different syllable positions within statements (dotted lines) and questions (solid lines). F0 contours are represented by time-normalized 10 equally distanced sample points.

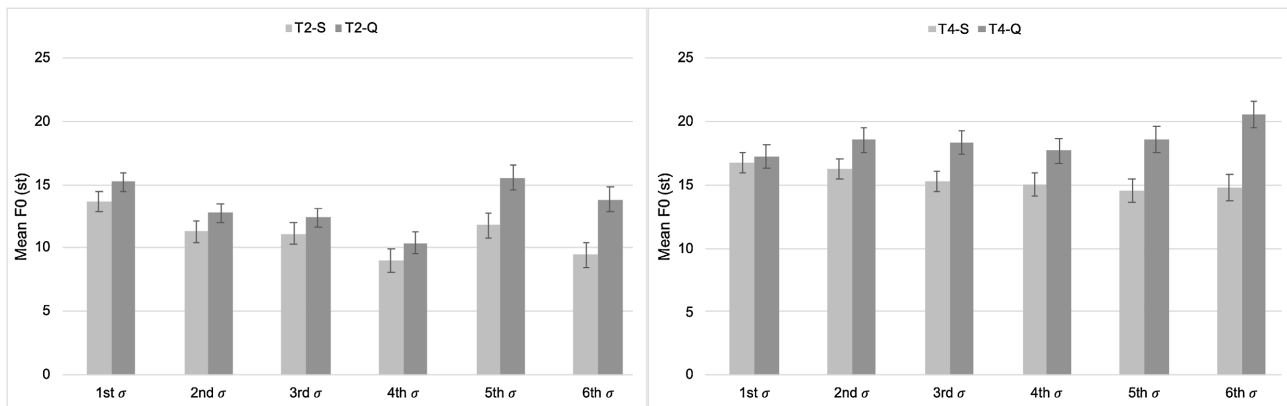


Figure 2. Comparison between mean F0 at each syllable position in questions and statements for T2 (left) and T4 (right). Error bar stands for one standard error.

of intonation, $\chi^2(1) = 147.80$, $p < 0.001$, and a significant syllable position \times intonation interaction, $\chi^2(5) = 51.67$, $p < 0.001$. For the significant interaction effect, we were interested in the F0 changes induced by question intonation. Thus, we compared the mean F0 between two intonation types at each syllable position first. The results revealed that for all syllable positions, the mean F0 values in question were significantly higher than those in statement ($ps \leq 0.004$). We also compared mean F0 among syllable positions in statement and question respectively. In statement, the six syllables could be classified into three groups. The first syllable formed the first group on its own, and it had the highest mean F0 compared with all other syllables ($ps \leq 0.001$). The second group consisted of the second, third and fifth syllables ($ps \geq 0.475$), and the third group contained the fourth and final syllables ($ps = 1.000$). The second group had lower mean F0 than the first group, but higher mean F0 than the third group. In question, the first syllable still had higher mean F0, which was similar to that of the fifth syllable ($p = 1.000$) and final syllable ($p = 0.162$). The final syllable had significantly lower mean F0 than the penultimate syllable ($p = 0.034$). These results indicate that compared with those in preceding syllables, T2 in the penultimate and final syllables revealed significantly greater raising effect.

Results of comparing mean F0 values for T4 were different from those for T2. The statistical analysis revealed a main effect of intonation, $\chi^2(1) = 151.99$, $p < 0.001$, and also a significant syllable position \times intonation interaction, $\chi^2(5) = 29.85$, $p < 0.001$; however, it revealed no main effect of syllable position, $\chi^2(5) = 6.26$, $p = 0.281$. Post-hoc simple comparisons were also conducted. First, the mean F0 value of the first syllable had similar mean F0 in statement intonation and question intonation ($p = .508$), while the mean F0 of each remaining syllable was significantly higher in question than in statement ($ps \leq 0.003$). Second, regarding mean F0, the first two syllables in statement formed a group ($p = 1.000$), the remaining four syllables formed the other group ($ps \geq 0.183$), and the former group had significantly higher mean F0 than the latter group ($ps \leq 0.012$). The syllables in question generally had similar mean F0, and significant differences were only

observed between the third and final syllables ($p = 0.009$), as well as the fourth and final syllables ($p < 0.001$). These results indicate that for T4, the most pronounced raising effect was localized in the final syllable.

3.2. F0 Curves of Upper and Lower Lines and Excursion Size

Figure 3 displays the upper and lower limit curves for T2 and T4 in both questions and statements. First, similar to the mean F0 patterns, both the upper and lower curves of each tone were raised in questions relative to their statement counterparts. Second, we observed an accelerated local F0 rise in sentence-final position for questions, particularly evident in the upper F0 curves of T2 and the lower F0 curves of T4. Third, the divergence between maximum and minimum F0 values of T2 expanded in question-final position, while T4 did not exhibit this pattern.

The results of statistical analysis using linear mixed-effects regression models on excursion sizes of T2 and T4 supported the above observations. The excursion size was calculated as difference between maximum and minimum F0 values in semitones. An increased excursion size indicates an expanded F0 range. **Figure 4**

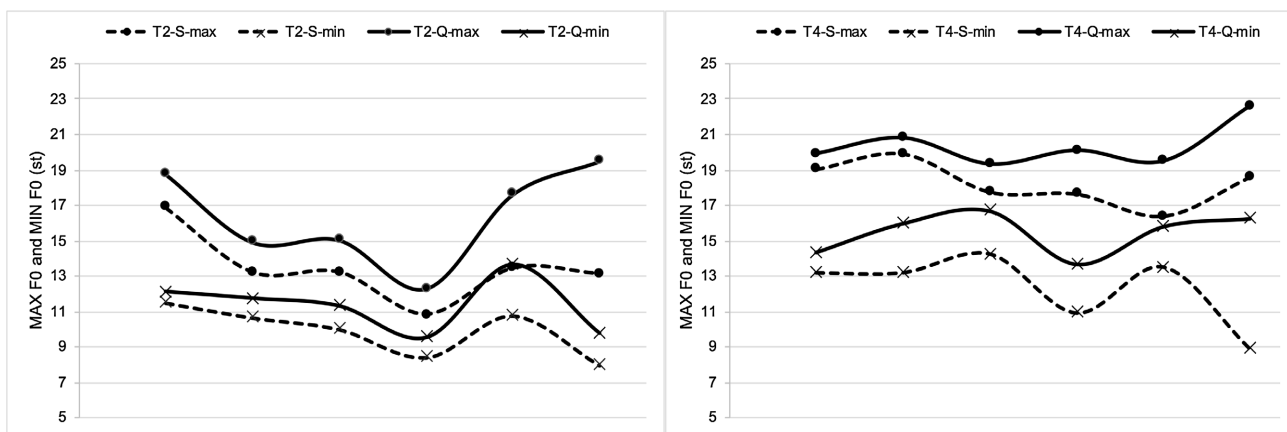


Figure 3. Upper and lower curves of T2 (left) and T4 (right) within six-syllable SVO sentences in current experiment. F0 curves are represented by maximum and minimum F0 values of tones in different syllable positions. Dotted lines indicate upper and lower curves in statement intonation, while solid lines indicate upper and lower curves in question intonation.

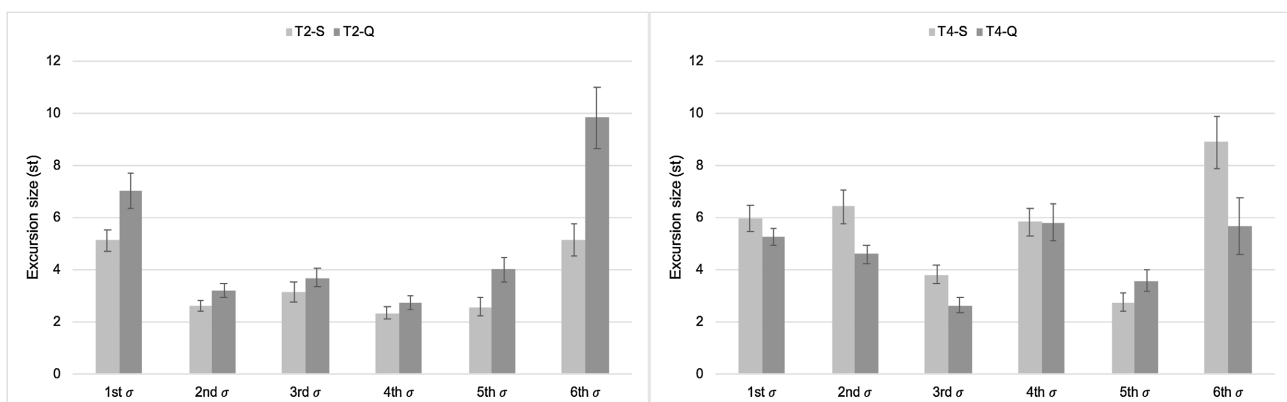


Figure 4. Comparison between excursion size at each syllable position in questions and statements for T2 (left) and T4 (right). Error bar stands for one standard error.

illustrates the means and error bars of excursion size of each tone across conditions. For T2, the statistical analysis revealed a main effect of syllable position, $\chi^2(5) = 186.91$, $p < 0.001$, a main effect of intonation, $\chi^2(1) = 48.15$, $p < 0.001$, and a significant syllable position \times intonation interaction, $\chi^2(5) = 48.82$, $p < 0.001$. First, we compared excursion size between two intonation types at each syllable position, as we were interested in the localized position showing question-induced change of F0 range. Question intonation induced significant change of excursion size on the four peripheral syllable, i.e., the sentence-initial syllable ($p = 0.011$), the second syllable ($p = 0.046$), the penultimate syllable ($p = 0.001$), and the sentence-final syllable ($p < 0.001$). Second, we continued to compare excursion size among the six syllable positions in question and statement, as we were also interested in the expansion of F0 range at different syllable positions induced by question intonation. In question, peripheral syllables and middle syllables had different excursion sizes. Specifically, the former group had significantly larger excursion sizes than the latter group ($ps \leq 0.001$). In addition, the sentence-initial and sentence-final syllables did not show a significant difference in excursion size ($p = 0.122$), and so did comparisons among the middle syllables ($ps \geq 0.055$). Similar patterns were observed in statement. That is, the two peripheral syllables had significantly larger excursion sizes than the middle syllables. These results suggest that for T2, the initial two syllables and the final two syllables expended their F0 ranges significantly in questions.

The results of excursion size of T4 showed different patterns. The statistical analysis revealed a main effect of syllable position, $\chi^2(5) = 42.62$, $p < 0.001$, a main effect of intonation, $\chi^2(1) = 12.59$, $p < 0.001$, and a significant syllable position \times intonation interaction, $\chi^2(5) = 23.09$, $p < 0.001$. Regarding the intonation-induced change of excursion size at different syllable position, post-hoc simple comparisons revealed that question intonation induced significant change of excursion size on the second syllable ($p = 0.003$), the third syllable ($p = 0.002$), the penultimate syllable ($p = 0.033$), and the sentence-final syllable ($p = 0.012$). When the results of comparisons among different syllable positions revealed a relatively consistent small excursion size for middle syllables. In statement, the fifth syllable had similar excursion size to the third syllable ($p = 0.911$), but significantly smaller excursion sizes than the remaining syllables ($ps < 0.001$). However, in question, the third syllables had significantly smaller excursion sizes than the first, second and fourth syllables ($ps \leq 0.017$), whereas there was no significant difference of excursion size among the third, fifth and final syllables. In other words, the excursion size of sentence-final T4 reduced significantly in the question.

4. Discussion

This study investigated F0 modifications induced by question intonation on Mandarin T2 (rising tone) and T4 (falling tone), examining both global and local F0 variations. Through a carefully designed production experiment that incorporated refined stimulus materials and an expanded participant pool compared to

previous Mandarin intonation studies (Qin, 2017a, 2017b), we obtained F0 measurements including time-normalized F0, mean F0, maximum F0, minimum F0, and excursion size. Our analysis focused on comparing two primary intonational effects: the contour-raising effect and range-expansion effect associated with question intonation.

Our descriptive and statistical analyses of time-normalized F0 contours and mean F0 values for T2 and T4 revealed three key patterns in question intonation compared to those in statement intonation. First, word-level canonical lexical contours, i.e., the rising contour of T2 and the falling contour of T4, were preserved at the sentence-level. This observation is consistent with previous finding that the F0 contours of lexical tones remain relatively unchanged from their citation forms (Cao, 2002; He & Jing, 1992; Ho, 1977; Shen, 1989; Wu, 1996; Yang & Kong, 2020; Zhang et al., 2021). The preservation of canonical lexical contours supports that the implementation of intonation in tone languages is more constrained than in non-tone languages, because both lexical tones and intonation utilize F0 to distinguish meanings. Specifically, the former distinguishes meanings at word-level, while the later conveys meanings at sentence level. However, F0 contours in penultimate syllables were relatively flattened in both intonation types, no matter the lexical tone is a rising or falling tone. For them, we attributed such flattened F0 contours to the prosodic status of these penultimate syllables. Unlike other syllables, which stand at boundary of either a prosodic word, or phonological phrase, or an intonational phrase, the penultimate syllable is in the middle of a prosodic word and aligns with the boundary of lower prosodic unit. According to previous studies working on pre-boundary lengthening (Cao, 2005, 2011; Lin, 2000; Wang et al., 2004; Wightman et al., 1992; Yang, 1997), prosodic-word-internal syllable has short duration, which constraints the realization of a full rising or falling. Thus, the penultimate syllables in our current experiment showed flattened F0 contours. In addition, sentence-final T2 had a slight initial falling, either in statement or question intonation. We attributed this slight initial falling in T2's F0 contour to the lexical tone of its preceding syllable. In our recording material, the final T2 followed another T2. In other words, the end point of the first rising tone has a high tone target, and the F0 has to be lowered to initiate the second rising tone. As a consequence, the F0 contour of the second rising tone showed a initial falling contour.

Second, the F0 contour of lexical tone in each syllable was raised. Statistically, the mean F0 was significantly higher in questions than in statements, regardless of the syllable's lexical tone. These results are in accordance with previous finding of a global F0 rise across the entire question (Gårding, 1987; Liu & Xu, 2005; Ni & Kawai, 2004; Shen, 1985, 1990, 1994; Yuan et al., 2002). Given that the canonical F0 contours of these lexical tones are largely maintained when intonation interacts, the current results suggest that question intonation in Mandarin is implemented by raising each lexical contour but not changing tone contours to a rising shape.

Third, the intonation-induced manipulation of F0 was most pronounced in the final portion of questions. Specifically, the increase of mean F0 was significant at penultimate and final syllables for T2, while only at the final syllable for T4. These findings support the local accelerating rise toward the end of question on the one hand (Yuan, 2006; Yuan et al., 2002), and demonstrate the difference of contour-raising effect induced by T2 and T4 on the other hand. In addition, the initial syllables consistently had high mean F0 values in both statements and questions, suggesting that unlike previous proposal (Shen, 1990; Shi, 1980), an initial contour raising may not be a distinctive feature for question.

In addition to F0 contour analysis, we examined F0 range modifications in questions, focusing on maximum F0, minimum F0, and excursion size, i.e., difference between maximum and minimum F0 values. Descriptively, in final position, T2 exhibited expanded F0 range and increased slope, while T4 showed slightly flattened contours. Statistically, excursion size patterns differed. First, T2 displayed enlarged excursions in question-final position, with the most pronounced effects at peripheral syllables; conversely, T4 demonstrated reduced excursions, losing its statement-like final-syllable prominence and exhibiting globally decreased excursion sizes. Second, while our results confirm previous observations of question-final T2 expansion and T4 slope reduction (Yuan, 2006; Yuan et al., 2002), they diverge in revealing that significant range expansion occurred exclusively with T2. In addition, no mid-sentence F0 expansion was observed, contrasting with previous reports of such effects (Xie, 2021).

5. Conclusion

This study investigated how question intonation induces both global and local F0 modifications on syllables bearing inherently rising (T2) and falling (T4) lexical tones in Mandarin, while examining whether these modifications operate equivalently across different syllable positions. Through a production experiment utilizing carefully designed test materials, we analyzed and compared multiple F0 parameters including time-normalized F0, mean F0, maximum and minimum F0 values, and excursion size. Our results reveal distinct intonation-induced changes for T2 and T4 at different positions in questions, demonstrating differential modulation of question intonation through contour-raising and range-expansion. First, although both tones exhibited global F0 raising, T2 showed a significant raising effect at both penultimate and final syllables, whereas T4 displayed this effect only at the final syllable. Second, while both tones underwent F0 range modifications, T2 demonstrated range expansion in contrast to T4's range compression. Third, significant range modification for T2 occurred at peripheral syllables, while for T4 this effect was confined to sentence-final syllables. Additionally, we observed consistent prominence of sentence-initial syllables across intonation types. These findings not only support the coexistence of global and local F0 manipulations in question intonation, but also provide detailed evidence about how Mandarin's rising and falling lexical tones undergo distinct intonation-induced

contour and range modifications.

The present study examined F0 realizations of Mandarin T2 and T4 in question intonation across syllable positions within six-syllable SVO sentences, specifically analyzing tonal variation under intonational modulation. However, several additional factors beyond intonation may influence lexical tone realization. As previously noted, syllables within prosodic words exhibit shorter durations, potentially constraining the full articulation of rising or falling tonal contours. Furthermore, segmental composition affects F0 realization. In particular, other factors being equal, high vowels intrinsically produce higher F0 values than low vowels. To enhance the robustness of our findings, future studies should incorporate systematically varied word lengths to investigate prosodic boundary effects, as well as carefully controlled minimal pairs to isolate segmental influences on tonal realization.

Acknowledgements

We would like to thank Ms. Xinyu Ye for assisting with sentence segmentation. We are deeply grateful to all the participants who took part in our experiment.

Fund

This study was supported by Philosophy and Social Science Research Project grant from Jiangsu Education Department (2023SJYB0257).

Ethics

The experiment was approved by the Institutional Review Boards of the first author's affiliated university (No. NNU202306019).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Bao, H.-W.-S. (2021). *BruceR: Broadly Useful Convenient and Efficient R Functions [Dataset]*. <https://cran.r-project.org/package=bruceR>
- Boersma, P., & Weenink, D. (2023). *Praat: Doing Phonetics by Computer [Computer Software]*.
- Bolinger, D. (1982). Intonation and Its Parts. *Language*, 58, 505-533. <https://doi.org/10.2307/413847>
- Cao, J. (2002). The Relationship between Tone and Intonation in Mandarin Chinese. *Chinese Journal of Phonetics*, 3, 195-202.
- Cao, J. (2005). Types of Segmental Lengthening and Their Prosodic Functions. *Journal of School of Chinese Language and Culture Nanjing Normal University*, 4, 160-167.
- Cao, J. (2011). Prosodic Structure and Speech Sound Variation. *Journal of School of Chinese Language and Culture Nanjing Normal University*, 3, 12-22.
- Chao, Y. R. (1929). Study of Intonation in Beijing Mandarin. In *Linguistic Essays of Y.R. Zhao* (2002 ed., pp. 253-272). Commercial Press.

- Chao, Y. R. (1933). Tone and Intonation in Chinese. *Journal of the Institute of History and Philosophy, Academia Sinica*, 4, 121-135.
- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. University of California Press.
- Cruttenden, A. (1981). Falls and Rises: Meanings and Universals. *Journal of Linguistics*, 17, 77-91. <https://doi.org/10.1017/s0022226700006782>
- Duanmu, S. (2007). *The Phonology of Standard Chinese*. Oxford University Press. <https://doi.org/10.1093/oso/9780199215782.001.0001>
- Gårding, E. (1987). Speech Act and Tonal Pattern in Standard Chinese: Constancy and Variation. *Phonetica*, 44, 13-29. <https://doi.org/10.1159/000261776>
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511616983>
- He, Y., & Jing, S. (1992). Experimental Exploration on Beijing Mandarin Intonation. *Language Teaching and Linguistic Studies*, 2, 71-96.
- Ho, A. T. (1977). Intonation Variation in a Mandarin Sentence for Three Expressions: Interrogative, Exclamatory and Declarative. *Phonetica*, 34, 446-457. <https://doi.org/10.1159/000259916>
- Jiang, H. (2010). *Experimental Study on Mandarin Intonation*. Capital Normal University Press.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/cbo9780511808814>
- Lee, O. J. (2005). *The Prosody of Question in Beijing Mandarin*. Doctor of Philosophy, The Ohio State University.
- Lin, M. (2000). Breaks and Prosodic Phrases in the Utterances of Standard Chinese. *Contemporary Linguistics*, 2, 210-217.
- Lin, M. (2006). Interrogative vs. Declarative Intonation and Boundary Tone in Standard Chinese. *Chinese Language and Literature*, 4, 364-376.
- Liu, F., & Xu, Y. (2005). Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation. *Phonetica*, 62, 70-87. <https://doi.org/10.1159/000090090>
- Liu, M., Chen, Y., & Schiller, N. O. (2022). Context Matters for Tone and Intonation Processing in Mandarin. *Language and Speech*, 65, 52-72. <https://doi.org/10.1177/0023830920986174>
- Ni, J., & Kawai, H. (2004). Pitch Targets Anchor Chinese Tone and Intonation Patterns. In *Speech Prosody 2004* (pp. 95-98). ISCA. <https://doi.org/10.21437/speechprosody.2004-22>
- Ohala, J. J. (1983). Cross-language Use of Pitch: An Ethological View. *Phonetica*, 40, 1-18. <https://doi.org/10.1159/000261678>
- Qin, P. (2017a). The Information Structure and Intonation Patterns of Chinese. *Nankai Linguistics*, 2, 81-91.
- Qin, P. (2017b). The Information Structure and Durational Pattern of Chinese. *Chinese Language Learning*, 6, 78-86.
- Shen, J. (1985). Pitch Range of Tone and Intonation in Beijing Dialect. In *Working Papers in Experimental Phonetics* (pp. 73-130). Peking University Press.
- Shen, J. (1994). The Components and Patterns of Chinese Intonation. *Dialect*, 3, 221-228.

- Shen, X. (1989). Interplay of the Four Citation Tone and Intonation in Mandarin Chinese. *Journal of Chinese Linguistics*, 17, 61-74.
- Shen, X. (1990). *The Prosody of Mandarin Chinese*. University of California Press.
- Shi, P. (1980). Intonation Variations in Four Types of Chinese Sentences. *Language Education and Research*, 2, 71-81.
- Wang, B., Yang, Y., & Lv, S. (2004). Acoustic Analysis on Prosodic Hierarchical Boundaries of Chinese. *Acta Acustica*, 29, 29-36.
- Wang, P., Deng, F., & Shi, F. (2017). The Experimental Study on the Pitch of Declarative Sentences and Interrogative Sentences in Standard Mandarin: Based on a Large Sample. *Applied Linguistics*, 4, 59-68.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental Durations in the Vicinity of Prosodic Phrase Boundaries. *The Journal of the Acoustical Society of America*, 91, 1707-1717. <https://doi.org/10.1121/1.402450>
- Wu, Z. (1996). Contribution of Yuen-Ren Chao on Chinese Tones. *Journal of Tsinghua University*, 11, 58-63.
- Xie, H. (2021). *Comparative Analysis of Volume Ratio with Duration Ratio and Pitch Fluctuation Scale in Several Kinds of Chinese Declarative Sentence and Interrogative Sentence*. Nankai University.
- Xu, B. R. (2013). *Cross-Linguistic Perception of Intonation by Cantonese and Mandarin Listeners*. Master of Philosophy, Chinese University of Hong Kong.
- Xu, Y. (2013). *ProsodyPro—A Tool for Large-Scale Systematic Prosody Analysis* (pp. 7-10).
- Yang, J., & Kong, J. (2020). *A Study on the Perception of Declarative and Interrogative Sentences in Mandarin*. *Chinese Journal of Phonetics*, 13, 1-13.
- Yang, Y. (1997). Prosodic Cues to Syntactic Boundaries. *Acta Acustica*, 22, 414-421.
- Yuan, J. (2006). Mechanisms of Question Intonation in Mandarin. In Q. Huo, B. Ma, E.-S. Chng, & H. Li (Eds.), *Chinese Spoken Language Processing* (pp. 19-30). Springer. https://doi.org/10.1007/11939993_7
- Yuan, J. (2011). Perception of Intonation in Mandarin Chinese. *Journal of the Acoustical Society of America*, 130, 1063-1074.
- Yuan, J., & Shih, C. (2004). Confusability of Chinese Intonation. In *Speech Prosody 2004* (pp. 131-134). ISCA. <https://doi.org/10.21437/speechprosody.2004-31>
- Yuan, J., Shih, C., & Kochanski, G. P. (2002). Comparison of Declarative and Interrogative Intonation in Chinese. In *Speech Prosody 2002* (pp. 711-714). ISCA. <https://doi.org/10.21437/speechprosody.2002-163>
- Zhang, J., Duanmu, S., & Chen, Y. (2021). Prosodic Systems: China and Siberia. In C. Gussenhoven, & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (pp. 332-343). Oxford University Press.

Appendix

Table A1. Four simple conversations embedding test sentences.

No.	Speaker	Pinyin form	Translation
1	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>zhang1hao4 zuo4 kao3rou4bing3.</i>	‘Zhanghao made grilled meat pie.’
	A	<i>zhang1hao4 zuo4 kao3rou4bing3?</i>	‘Did Zhanghao make grilled meat pie.’
	B	<i>shi4de0.</i>	‘Yes.’
2	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>wang2bao3 chang2 dun4yu2tou2.</i>	‘Wangbao tasted fish head stew.’
	A	<i>wang2bao3 chang2 dun4yu2tou2?</i>	‘Did Wangbao taste fish head stew?’
	B	<i>shi4de0.</i>	‘Yes.’
3	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>li3chao1 dian3 zha2ji1chi4.</i>	‘Lichao ordered fried chicken wings.’
	A	<i>li3chao1 dian3 zha2ji1chi4?</i>	‘Did Lichao order fried chicken wings?’
	B	<i>shi4de0.</i>	‘Yes.’
4	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>zhao4hao2 chi1 shao1sun3gan1.</i>	‘Zhaohao ate cooked bamboo shoots.’
	A	<i>zhao4hao2 chi1 shao1sun3gan1?</i>	‘Did Zhaohao eat cooked bamboo shoots?’
	B	<i>shi4de0.</i>	‘Yes.’

a. “0” indicates the neutral tone in Mandarin Chinese.

Table A2. Four simple conversations embedding filler sentences.

No.	Speaker	Pinyin form	Translation
1	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>zhang1chao1 jin1ye4 chi1 shao1ji1.</i>	‘Zhangchao will eat grilled chicken tonight.’
	A	<i>zhang1chao1 jin1ye4 chi1 shao1ji1?</i>	‘Will Zhangchao eat grilled chicken tonight?’
	B	<i>shi4de0.</i>	‘Yes.’
2	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>wang2hao2 zuo2chen2 chang2 zha2yu2.</i>	‘Wanghao tasted fried fish yesterday morning.’
	A	<i>wang2hao2 zuo2chen2 chang2 zha2yu2?</i>	‘Did Wanghao taste fried fish yesterday morninh?’
	B	<i>shi4de0.</i>	‘Yes.’
3	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>li3bao3 mei3wan3 dian3 kao3sun3.</i>	‘Libao ordered bamboo shoot every night.’
	A	<i>li3bao3 mei3wan3 dian3 kao3sun3?</i>	‘Does Lichao order bamboo shoot every night?’
	B	<i>shi4de0.</i>	‘Yes.’
4	A	<i>ni3 gang1gang1 shuo1 shen2me0?</i>	‘What did you say?’
	B	<i>zhao4hao4 hou4tian4 zuo4 dun4rou4.</i>	‘Zhaohao will make meat stew the day after.’
	A	<i>zhao4hao4 hou4tian4 zuo4 dun4rou4?</i>	‘Will Zhaohao make meat stew the day after?’
	B	<i>shi4de0.</i>	‘Yes.’

a. “0” indicates the neutral tone in Mandarin Chinese.

Table A3. Results of post-hoc simple comparisons on mean F0 of T2 among syllables in statement and question.

contrast	df	Statement				Question			
		estimate	SE	t. ratio	p. value	estimate	SE	t. ratio	p. value
syl.1 - syl.2	21	2.388	0.302	7.912	<0.0001	2.467	0.466	5.292	0.0005
syl.1 - syl.3	21	2.525	0.255	9.922	<0.0001	2.835	0.359	7.897	<0.0001
syl.1 - syl.4	21	4.684	0.454	10.311	<0.0001	4.84	0.572	8.455	<0.0001
syl.1 - syl.5	21	1.9	0.391	4.863	0.0012	-0.341	0.393	-0.869	1
syl.1 - syl.6	21	4.231	0.486	8.712	<0.0001	1.389	0.496	2.799	0.1615
syl.2 - syl.3	21	0.138	0.182	0.756	1	0.368	0.31	1.189	1
syl.2 - syl.4	21	2.296	0.299	7.681	<0.0001	2.373	0.307	7.736	<0.0001
syl.2 - syl.5	21	-0.488	0.293	-1.663	1	-2.808	0.487	-5.766	0.0002
syl.2 - syl.6	21	1.843	0.325	5.678	0.0002	-1.077	0.52	-2.072	0.7616
syl.3 - syl.4	21	2.158	0.31	6.961	<0.0001	2.005	0.318	6.306	<0.0001
syl.3 - syl.5	21	-0.625	0.271	-2.302	0.4746	-3.176	0.379	-8.38	<0.0001
syl.3 - syl.6	21	1.705	0.31	5.505	0.0003	-1.445	0.423	-3.417	0.0389
syl.4 - syl.5	21	-2.783	0.369	-7.549	<0.0001	-5.181	0.526	-9.849	<0.0001
syl.4 - syl.6	21	-0.453	0.38	-1.194	1	-3.451	0.58	-5.95	0.0001
syl.5 - syl.6	21	2.33	0.335	6.955	<0.0001	1.731	0.503	3.439	0.0369

Table A4. Results of post-hoc simple comparisons on mean F0 of T4 among syllables in statement and question.

contrast	df	Statement				Question			
		estimate	SE	t. ratio	p. value	estimate	SE	t. ratio	p. value
syl.1 - syl.2	16	0.4693	0.251	1.868	1	-1.3208	0.884	-1.494	1
syl.1 - syl.3	16	1.4098	0.341	4.13	0.0118	-1.1315	0.768	-1.473	1
syl.1 - syl.4	16	1.6219	0.29	5.597	0.0006	-0.4603	0.84	-0.548	1
syl.1 - syl.5	16	2.1774	0.322	6.765	0.0001	-1.3402	0.898	-1.493	1
syl.1 - syl.6	16	1.9945	0.471	4.232	0.0095	-3.3651	1.11	-3.028	0.12
syl.2 - syl.3	16	0.9405	0.179	5.26	0.0012	0.1893	0.247	0.768	1
syl.2 - syl.4	16	1.1526	0.22	5.238	0.0012	0.8605	0.38	2.267	0.5641
syl.2 - syl.5	16	1.7081	0.259	6.593	0.0001	-0.0194	0.466	-0.042	1
syl.2 - syl.6	16	1.5253	0.371	4.113	0.0122	-2.0443	0.624	-3.276	0.0713
syl.3 - syl.4	16	0.2121	0.29	0.73	1	0.6712	0.284	2.361	0.469
syl.3 - syl.5	16	0.7676	0.325	2.359	0.4703	-0.2087	0.438	-0.476	1
syl.3 - syl.6	16	0.5847	0.31	1.887	1	-2.2335	0.526	-4.245	0.0093
syl.4 - syl.5	16	0.5555	0.197	2.826	0.1825	-0.8799	0.377	-2.332	0.4968
syl.4 - syl.6	16	0.3726	0.329	1.133	1	-2.9048	0.489	-5.934	0.0003
syl.5 - syl.6	16	-0.1829	0.382	-0.479	1	-2.0249	0.606	-3.343	0.0619

Table A5. Results of post-hoc simple comparisons on excursion size of T2 among syllables in statement and question.

contrast	df	Statement				Question			
		estimate	SE	t. ratio	p. value	estimate	SE	t. ratio	p. value
syl.1 - syl.2	21	2.5073	0.379	6.619	<0.0001	3.8337	0.527	7.277	<0.0001
syl.1 - syl.3	21	2.0006	0.461	4.338	0.0043	3.3196	0.528	6.289	<0.0001
syl.1 - syl.4	21	2.7848	0.397	7.023	<0.0001	4.2896	0.514	8.347	<0.0001
syl.1 - syl.5	21	2.5381	0.351	7.236	<0.0001	3.0087	0.507	5.933	0.0001
syl.1 - syl.6	21	-0.0245	0.596	-0.041	1.0000	-2.8196	0.964	-2.924	0.1218
syl.2 - syl.3	21	-0.5067	0.283	-1.787	1.0000	-0.5141	0.382	-1.347	1.0000
syl.2 - syl.4	21	0.2775	0.259	1.07	1.0000	0.456	0.201	2.274	0.5037
syl.2 - syl.5	21	0.0309	0.315	0.098	1.0000	-0.825	0.438	-1.886	1.0000
syl.2 - syl.6	21	-2.5318	0.629	-4.024	0.0092	-6.6533	1.1	-6.056	0.0001
syl.3 - syl.4	21	0.7842	0.326	2.402	0.3847	0.97	0.313	3.101	0.0811
syl.3 - syl.5	21	0.5376	0.395	1.362	1.0000	-0.3109	0.302	-1.031	1.0000
syl.3 - syl.6	21	-2.0251	0.665	-3.044	0.0926	-6.1393	0.975	-6.297	<0.0001
syl.4 - syl.5	21	-0.2466	0.278	-0.888	1.0000	-1.2809	0.393	-3.263	0.0557
syl.4 - syl.6	21	-2.8093	0.556	-5.049	0.0008	-7.1093	1.04	-6.804	<0.0001
syl.5 - syl.6	21	-2.5627	0.5	-5.124	0.0007	-5.8284	0.872	-6.685	<0.0001

Table A6. Results of post-hoc simple comparisons on excursion size of T4 among syllables in statement and question.

contrast	df	Statement				Question			
		estimate	SE	t. ratio	p. value	estimate	SE	t. ratio	p. value
syl.1 - syl.2	15	-0.474	0.492	-0.963	1.0000	0.647	0.453	1.43	1.0000
syl.1 - syl.3	15	2.138	0.649	3.293	0.0740	2.608	0.47	5.551	0.0008
syl.1 - syl.4	15	0.132	0.402	0.328	1.0000	-0.562	0.573	-0.981	1.0000
syl.1 - syl.5	15	3.2	0.36	8.879	<0.0001	1.669	0.458	3.647	0.0358
syl.1 - syl.6	15	-2.938	0.939	-3.129	0.1035	-0.439	1.04	-0.421	1.0000
syl.2 - syl.3	15	2.611	0.72	3.628	0.0372	1.961	0.489	4.01	0.0170
syl.2 - syl.4	15	0.606	0.521	1.162	1.0000	-1.209	0.763	-1.586	1.0000
syl.2 - syl.5	15	3.673	0.536	6.857	0.0001	1.022	0.499	2.047	0.8796
syl.2 - syl.6	15	-2.464	0.835	-2.95	0.1489	-1.086	1.000	-1.081	1.0000
syl.3 - syl.4	15	-2.006	0.536	-3.74	0.0296	-3.171	0.746	-4.25	0.0105
syl.3 - syl.5	15	1.062	0.524	2.028	0.9112	-0.939	0.555	-1.692	1.0000
syl.3 - syl.6	15	-5.076	0.949	-5.346	0.0012	-3.047	1.17	-2.601	0.3010
syl.4 - syl.5	15	3.068	0.484	6.339	0.0002	2.232	0.73	3.056	0.1200
syl.4 - syl.6	15	-3.07	0.968	-3.173	0.0946	0.124	1.26	0.098	1.0000
syl.5 - syl.6	15	-6.137	0.914	-6.712	0.0001	-2.108	1.15	-1.828	1.0000