The handle [http://hdl.handle.net/1887/37609](http://hdl.handle.net/1887/37609) holds various files of this Leiden University dissertation

**Author:** Qian Li  
**Title:** The production and perception of tonal variation: evidence from Tianjin Mandarin  
**Issue Date:** 2016-02-10
6.1 Introduction

In speech production, utterances are phrased into constituents of varying sizes, which forms a hierarchical prosodic structure (see e.g., Nespor & Vogel, 1986; Shattuck-Hufnagel & Turk, 1996; Truckenbrodt, 1999; Selkirk, 1995, 2001; Frota, 2012 for comprehensive reviews). There has been abundant cross-linguistic evidence that prosodic boundary plays an important role in conditioning the articulation of segments, in terms of domain-initial strengthening (e.g., Fougeron & Keating, 1997; Fougeron, 1999, 2001; Cho & Keating, 2001; Keating, 2003; Cho & McQueen, 2005), domain-final lengthening (e.g., Wightman et al., 1992; Kuzla et al., 2007), as well as resistance of coarticulation across boundaries (e.g., Egido & Cooper, 1980; Byrd & Saltzman, 1998; Tabain, 2003; Cho, 2004; Pan, 2007).

Much less, however, is known about how the f0 realization of lexical tones is influenced by different prosodic boundaries. With the limited number of studies concerning this issue (e.g., Shih, 1997; Yang & Wang, 2002; Pan & Tai, 2006; Scholz & Chen, 2014a, 2014b), it is clear that prosodic boundary does exert an effect on tonal implementation, which resembles its effect on segmental articulation. Note that all studies, however, have limited their attention to the realization of lexical full tones at prosodic boundaries.

We know that in languages like Standard Chinese, in addition to the lexically distinctive full tones, there exist a number of items under the cover term “neutral tone” (Chao, 1968), which are typically grammatical morphemes or the unstressed final syllable within a disyllabic lexical item. Their surface f0 realization is much less consistent than that of the full lexical tone syllables and therefore shows great variability (Chen & Xu, 2006 and references therein). No study thus far, however, has examined the effect of prosodic boundary on neutral tone realization.

In this study, we aimed to address this gap by exploring how the f0 realization of neutral tone is conditioned by different prosodic boundaries in Tianjin Mandarin (TM). As will become clear below, Tianjin Mandarin exhibits very interesting patterns of neutral tone f0 realization, which calls for further data from well-controlled experiments that can be scrutinized to shed light on the nature of neutral tone and the effect of prosodic boundary on tonal implementation.

Tianjin Mandarin is spoken in the urban area of Tianjin city, which is next to Beijing. Like Standard Chinese, Tianjin Mandarin has four lexical tones. Tone 1 (T1) is a low-falling tone, Tone 2 (T2) a high-rising tone, Tone 3 (T3) a low-dipping tone, and Tone 4 (T4) a high-falling tone. Figure 6.1 illustrates the f0 realization of the four lexical full tones that were produced in isolation (Li & Chen, 2016). The f0 values were normalized with z-score

---

6 A version of this chapter has been submitted for publication as: Qian Li & Yiya Chen. (under revision). Prosodically conditioned neutral tone realization in Tianjin Mandarin. Journal of East Asian Linguistics.
(Rose, 1987) with the formula: \( z = \frac{f_0 - f_{0\text{mean}}}{f_{0\text{SD}}} \) (\( z \): z-score; \( f_0 \): the observed f0 value in Hz; \( f_{0\text{mean}} \): the mean f0 value of the speaker in Hz; \( f_{0\text{SD}} \): the standard deviation of f0 value of the speaker in Hz). The illustrated tonal contours were then based on the mean z-score averaged across 50 samples produced by a male speaker in his 20s.

Figure 6.1 Lexical tones produced in isolation. Lines stand for the mean. Gray areas stand for ±1 standard error of mean. Tone 1 (T1) is illustrated with black line and dark gray area; Tone 2 (T2) with white line and dark gray area; Tone 3 (T3) with black line and light gray area; Tone 4 (T4) with white line and light gray area. Normalized time.

In addition to the four lexical full tones, Tianjin Mandarin has a number of items which are called neutral-tone syllables. These syllables usually follow a full tone syllable within disyllabic lexical items such as grammatical morphemes (e.g., the possessive marker de in wo de ‘mine’), the second syllable of reduplicative forms (e.g., the second ma in ma ma ‘mother’), or the second syllable of disyllabic lexical items (e.g., jin in tian jin ‘Tianjin’). Their distribution is thus very similar to the neutral-tone items in Standard Chinese (see e.g., Lu, 1995; Wang & Jiang, 1997 for the distribution of neutral-tone syllables in Standard Chinese).

Neutral tone in Tianjin Mandarin has been reported to have very different properties in its f0 realization compared to that in Standard Chinese. Specifically, while in most tonal contexts, the f0 realization of TM neutral tone shows similar patterns of f0 variation as that in Standard Chinese, neutral tone before the lexical low-falling tone (i.e., T1) has been reported to surface with a rising f0 contour (e.g., Wang, 2002; Li & Chen, 2011), which is not observed in Standard Chinese. This has led to the proposal that the rising f0 is due to a special rising tonal target of neutral tone before the low-falling T1 (e.g., Wang, 2002). The claimed tonal-context conditioned presence of the rising target of neutral tone in Tianjin Mandarin not only is idiosyncratic but also poses challenges to the established understanding of neutral tone realization based on data from Standard Chinese (Chen & Xu, 2006), which calls for further studies to understand the underlying mechanism of neutral tone f0 realization in general.
It is worth noting that the lexical low-falling tone (T1) has been found to show a considerable raising effect even when the preceding tone is a lexical full tone (Li & Chen, 2016). This suggests the possibility that the rising $f_0$ realization of neutral tone before T1 might be due to a general raising effect of T1 upon its preceding tones, rather than a special rising neutral tone target, which would then explain the so-called context-specific rising neutral tone target. Experimental data on neutral tone realization preceding T1 are thus needed to tap into and verify this possibility. If results of the experiment lend support to this possibility, a follow-up question that arises is how prosodic boundaries of different strength may regulate the $f_0$ raising effect on neutral tone introduced by its following low-falling T1.

The design of our experiment was therefore intended to seek answers to these two research questions, recapitulated in more details in the following.

1) Does Tianjin Mandarin have a special rising neutral tone target as reported in the literature?

We set out to answer this question by first trying to replicate the so-called rising neutral tone $f_0$ realization as reported in the literature. If the rising $f_0$ of TM neutral tone is confirmed, we further ask whether it is due to a special neutral tone target or due to other factors such as the general raising effect of the following T1. The approach we took is to increase the number of neutral tone syllables and examine the corresponding $f_0$ contour changes as a function of the number of neutral tone syllables, as in Chen and Xu (2006). If indeed there is a rising neutral tone target as claimed, we would expect continuous rising neutral tone $f_0$ realization. Otherwise, the so-called “special rising target” would be called into question.

2) How is neutral tone $f_0$ realization conditioned by different prosodic boundaries?

By examining TM neutral tone realization, we also aimed to investigate how in general, the $f_0$ realization of neutral tone is conditioned by different prosodic boundaries. To this end, we were interested in how the $f_0$ realization of neutral tone in Tianjin Mandarin is conditioned by the different prosodic boundaries between neutral tone(s) and the following lexical T1. The comparison was therefore made between a lower-level prosodic boundary vs. a higher-level prosodic boundary. Syntactically, they correspond, respectively, to a boundary within a noun phrase (NP) and a boundary across the subject and the predicate phrase of an utterance. While there has been various debates on how syntactic constituents map onto prosodic domains, there is quite some consensus that these two types of syntactic structures typically correspond to prosodic domains of distinct levels (e.g., Truckenbrodt, 1999). Specifically, we were interested in to what extent the $f_0$ realization of a neutral tone is conditioned by these two different levels of prosodic boundaries.
6.2 Method

6.2.1 Materials

The stimuli were chosen via taking into consideration the following factors.

First, as we know that a neutral-tone syllable in Tianjin Mandarin always follows a full tone syllable and its realization is greatly influenced by the lexical tone of the preceding syllable, all four lexical tones in Tianjin Mandarin were included in the preceding syllable: T1 (low-falling), T2 (high-rising), T3 (low-dipping), and T4 (high-falling).

Second, as neutral tone in Tianjin Mandarin has been reported to show a rising f0 pattern only before T1, the full lexical tone immediately after neutral tone syllable(s) was consistently controlled as T1. We further controlled the lexical tone that follows the T1 as varying between T1 versus non-T1 (i.e., T2). This is because previous studies show that the low-falling T1 is realized differently as a function of the following tonal contexts. T1 is typically realized with a low-falling f0 contour in context (e.g., followed by T2), but when followed by another T1, the first T1 is realized with a rising f0 contour (e.g., Zhang & Liu, 2011; Li & Chen, 2016). As no existing literature provides a clear prediction of whether the T1 contextual variation might affect the neutral tone realization, we systematically controlled this factor so as to make sure whether different f0 realization of T1 would affect its raising effect upon the preceding neutral tone(s) by varying the lexical tone following T1.

Third, as one of the main goals of this study is to understand further the nature of the rising f0 in TM neutral tone within a broader context of searching for the more general mechanism of neutral tone realization, we adopted the design in Chen and Xu (2006) for neutral tone realization in Standard Chinese. We varied the number of the embedded neutral-tone syllables from 1 to 3, so as to investigate the specific domain of realization for such an f0 raising effect. Continuous rising neutral tone f0 within the domain would suggest the presence of an underlying rising neutral tone target, while more localized neutral tone f0 rising right before the following T1 would suggest that the so-called rising neutral tone reported in the literature is to be attributed to contextual tonal variation effect introduced by the following T1.

The fourth factor that we manipulated was aimed to tap into the effect of prosodic boundary on neutral tone realization. The boundary between neutral tone and the following T1 was varied between a low-level boundary (i.e., a Below-NP boundary) and a high-level boundary (i.e., a Subject-Predicate boundary), so that there were two different types of grouping patterns of neutral tone and the following T1. In the Below-NP boundary condition, the neutral tone was grouped together with the next T1, while in the Subject-Predicate boundary condition, the neutral tone was grouped separately from the next T1.

All test materials were embedded in the sentence frame Ta’ shuo’… (“He said …”). The length of each carrier sentence was controlled to be within 12-14 syllables. For example,
Below-NP Boundary:
e.g., 他说他的猫抓住了那只老鼠。
\[ Ta^{1} shuo^{1} ta^{1} de mao^{1} zhua^{4} le na^{4} zhi^{1} lao^{3} shu^{3}. \]
He said he-possessive cat catch-perfective that-classifier mouse.
\[ He said his cat has caught that mouse. \]

Subject-Predicate Boundary:
e.g., 他说姐姐经营了一家餐厅。
\[ Ta^{1} shuo^{1} jie^{3} jie^{3} jing^{1} ying^{2} le yi^{4} jia^{1} can^{1} ting^{1}. \]
He said sister run-present one-classifier restaurant.
\[ He said sister is running a restaurant. \]

Last but not least, we also controlled the information structure of the utterances elicited since focus has been shown to introduce considerable \( \xi \) variation to tonal realization, especially to neutral tone realization (Xu, 1999; Chen & Xu, 2006; Li & Chen, 2011). Specifically, we elicited the utterances as answers to pre-recorded questions, which resulted in two focus conditions. In the on-focus condition, the neutral tone sequence was under focus; in the pre-focus condition, focus was on later parts of the sentence, as shown in the following examples, where focus is indicated by italics.

On-Focus Condition:
**QUESTION** 他说谁的猫抓住了那只老鼠？
\[ Ta^{1} shuo^{1} shui^{3} de mao^{1} zhua^{4} le na^{4} zhi^{1} lao^{3} shu^{3}? \]
He said who-possessive cat catch-perfective that-classifier mouse?  
\[ He said whose cat has caught that mouse? \]
**ANSWER** 他说他的猫抓住了那只老鼠。
\[ Ta^{1} shuo^{1} ta^{1} de mao^{1} zhua^{4} le na^{4} zhi^{1} lao^{3} shu^{3}. \]
He said he-possessive cat catch-perfective that-classifier mouse.
\[ He said his cat has caught that mouse. \]

Pre-Focus Condition:
**QUESTION** 他说他的猫怎么了？
\[ Ta^{1} shuo^{1} ta^{1} de mao^{1} zen^{3} me le? \]
He said he-possessive cat how-perfective?
\[ He said what happened to his cat? \]
**ANSWER** 他说他的猫抓住了那只老鼠。
\[ Ta^{1} shuo^{1} ta^{1} de mao^{1} zhua^{4} le na^{4} zhi^{1} lao^{3} shu^{3}. \]
He said he-possessive cat catch-perfective that-classifier mouse.
\[ He said his cat has caught that mouse. \]

6.2.2 Subjects

A total of fourteen speakers (8 males and 6 females; Mean=24) participated in the experiment. All speakers were born in the 1980s and raised in the urban areas of Tianjin. They were undergraduate or postgraduate students studying in Beijing at the time of the
experiment. None of them had lived out of the Tianjin city before 18. They were paid for their participation but unaware of the purpose of the experiment. All participants provided written informed consent.

6.2.3 Recording

All eliciting questions were recorded beforehand by a female native speaker of Tianjin Mandarin. During the experiment, participants were played one question at a time. They were requested to respond to the question with the sentence presented on the computer screen. All found the task straightforward and followed the same procedure. Recordings were conducted in the Phonetics Lab at Beijing Language and Culture University using an M-Audio® mobile digital audio recorder MicroTrack II with 44.1 kHz sampling rate and 16 bit rate in mono channel. In total, 96 sentences (4 initial tones * 3 neutral tone numbers * 2 boundary types * 2 tones after the immediately following T1 * 2 focus conditions) were elicited from each participant with three repetitions.

6.2.4 f0 measurement & data analysis

The acoustic data were manually segmented in Praat (Boersma & Weenink, 2011). A custom written script was used for f0 extraction and smoothing. f0 contours were obtained by taking 20 points (in Hertz) in the rhyme part for each full tone syllable and 10 points for the neutral tone syllables. To eliminate the pitch range difference due to gender and to better illustrate the neutral tone realization, each speaker's raw f0 data were normalized with z-score (Rose, 1987). The illustrated tonal contours were then based on the mean z-score averaged across speakers and repetitions.

For quantitative analyses, we employed the growth curve analysis (Mirman, 2014) with the package lme4 (Bates et al., 2014) in R (R Core Team, 2014). For the present study, we used only up to second-order polynomials since the most complex f0 contour of lexical or neutral tones in our data had only convex or concave contour shape (i.e., U-shape or reversed).

The f0 realization of neutral tone was therefore analyzed by assessing the intercept, linear and quadratic coefficients in curve fitting. Linear mixed-effect models were then fitted to examine the neutral tone f0 realization as a function of different preceding lexical tones (i.e., T1-T4), following lexical tonal combination (i.e., T1T1 or T1T2), boundary types (i.e., below-NP boundary vs. Subject-Predicate boundary), and focus (i.e., On-Focus vs. Pre-Focus).

To test the effect of the above factors on the overall neutral tone f0 realization, three base models for sentences with different numbers of neutral tone syllables were first established with only time terms in the fixed structure as well as the random intercept and slope of SUBJECT on all time terms. Other fixed factors (i.e., BOUNDARY, PRECEDING TONE, FOLLOWING TONAL COMBINATION, FOCUS, NEUTRAL TONE LOCATION) were added onto each base model in a step-wise fashion. Model fits were tested at each step by assessing whether including one factor improves the goodness
of fit using the likelihood-ratio test. The effect of the factors on each neutral tone syllable was further assessed by establishing separate linear-mixed effect models for each neutral tone syllable. Parameter-specific $p$-values were estimated using the normal approximation (i.e., treating the $t$-value as a $z$-value).

6.3 Results

Results of general model fits suggested that the only factor that did not significantly contribute to a better fit of the models for the $f0$ realization of neutral tones (regardless of the number of neutral tones) was FOLLOWING TONAL COMBINATION (i.e. T1T1 vs. T1T2). This indicates that the lexical tonal sequence T1T1 or T1T2 following neutral tone(s) made no difference in the $f0$ realization of neutral tone. We therefore only plotted data when neutral tone(s) were followed by the T1T1 combination for illustration.

Figure 6.2 shows the mean $f0$ realization of neutral tone(s) after different preceding lexical tones with different prosodic boundaries between the neutral tone(s) and the following under different focus conditions. Each $f0$ contour is an average of 42 repetitions by 14 speakers, and each gap stands for a syllable boundary. The four $f0$ contours in each graph differ in the first lexical tone, as indicated by four different line types, with black thick solid line for T1, black dotted line for T2, gray thick solid line for T3, and black thin solid line for T4. The immediately following full tone was kept constant as T1 in all conditions. The three columns differ in the number of neutral tone syllables (N): one neutral tone - three neutral tones in Column 1-3, respectively. In the upper two rows (A-B), focus was elicited on the phrase consisting of the first full tone and the neutral tone(s), as indicated by “Focus” with brackets. In the lower two rows (C-D), the focus was on the part following neutral tone(s). The prosodic boundary alternates between low-level (Rows A, C) and high-level (Rows B, D) within each focus condition.

6.3.1 Rising $f0$ realization of neutral tone

Our first research question concerned the nature and domain of $f0$ rising in TM neutral tone before T1 as reported in the literature. To answer this question, we first set out to replicate the rising neutral tone $f0$ contour with one neutral tone syllable embedded. As shown in Column 1 in Figure 6.2, the systematic rising $f0$ realization of neutral tone was only observed in Figure 6.2C-1, where neutral tone was pre-focused and followed by a Below-NP prosodic boundary. Second, in all four cases in Column 1, the $f0$ realization of neutral tone varied significantly as a function of different preceding tones ($\chi^2(9)=10183$, $p<0.001$).
The domain of $f_0$ raising is revealed by comparing $f_0$ contours as the number of embedded neutral tone syllables increased up to three, as shown from Column 1 to Column 3 in Figure 6.2. If the rising neutral tone $f_0$ was indeed due to the so-called rising neutral tonal target, we would have observed consecutive rising neutral tone $f_0$ contours over all neutral tone syllables or continuous rising $f_0$ over the string of neutral tone syllables, as the number of neutral tone syllables increased. However, as can be seen from Figure 6.2, none of the graphs shows such continuous/consecutive $f_0$ rising contours. For those which does show rising $f_0$ realization (Row C, Figure 6.2), the rising part was restricted to the last neutral-tone syllable that immediately preceded the Below-NP boundaries, as evident in the second neutral tone in Figure 6.2C-2 and the third neutral tone in Figure 6.2C-3. This suggests that the rising $f_0$ contour of neutral tone was unlikely due to the underlying $f_0$ rising target.

Instead, the mid-low target of neutral tone could be clearly observed when the number of neutral tone increased (Columns 2-3 in Figure 6.2). We can see that, when there are two neutral-tone syllables as in Column 2, at the end of the second neutral tone, there is much less variability despite that the effect of the preceding tones was found to be still significant for both neutral tones ($1^{st}$: $\chi^2(9)=16874$, p<0.001; $2^{nd}$: $\chi^2(9)=4980.4$, p<0.001).
When there are three neutral tones as in Column 3, the convergence of neutral tone f0 realization is even more apparent, despite that the influence from the preceding tones was still noticeable (1st: $\chi^2(9)=16263$, $p<0.001$; 2nd: $\chi^2(9)=13172$, $p<0.001$; 3rd: $\chi^2(9)=213.06$, $p=0.001$). In Figures 6.2A-3, 2B-3 and 2D-3, the f0 of the third neutral tone stayed at a stable mid-low f0 register with very subtle variation. In Figure 6.2C-3, where neutral tone was pre-focused and preceded a Below-NP boundary, the approaching of the mid-low target could be traced by the end of the second neutral tone, followed by a rising f0 pattern at the last neutral tone syllable.

6.3.2 The effect of prosodic boundary

The second goal of our study was to investigate the effect of different sizes of prosodic boundaries on the f0 realization of neutral tone. To this end, we compared the f0 realization of neutral tone preceding a low-level prosodic boundary (i.e., a Below-Phrase boundary as in Rows A and C of Figure 6.2) vs. a high-level prosodic boundary (i.e., a Subject-Predicate boundary as in Rows B and D in Figure 6.2). By comparing Rows A vs. C and Rows B vs. D, we can observe a clear raised f0 contour over neutral tone that immediately preceded a low-level prosodic boundary compared to that preceding a high-level prosodic boundary. This pattern holds across different focus conditions.

To better illustrate this difference, the f0 contours of the neutral tone syllable(s) following a T1 and preceding another T1 were re-plotted in Figure 6.3. Here, the two f0 contours in each graph differ in the level of prosodic boundary following the neutral tone(s), as indicated by solid lines for Subject-Predicate Boundary, and dotted lines for Below-NP Boundary. The three columns differ in the number of neutral-tone syllables (N): one neutral tone - three neutral tones in Column 1-3 respectively. In the upper row, focus was on the phrase consisting of the first full tone and the neutral tone(s), as indicated by “Focus” in brackets. In the lower row, focus was on the part following the neutral tone(s).
Table 6.1 Estimates of BOUNDARY on f0 realization of the immediately preceding neutral tones in different focus conditions; a for On-Focus condition, b for Pre-Focus condition.

<table>
<thead>
<tr>
<th></th>
<th>1st Neutral Tone</th>
<th>2nd Neutral Tone</th>
<th>3rd Neutral Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Slope</td>
<td>Intercept</td>
</tr>
<tr>
<td></td>
<td>( \beta = 0.25 )</td>
<td>( \beta = 0.26 )</td>
<td>( \beta = 0.18 )</td>
</tr>
<tr>
<td></td>
<td>( t = 9.94 )</td>
<td>( t = 3.32 )</td>
<td>( t = 12.14 )</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.001 )</td>
<td>( p &lt; 0.001 )</td>
<td>( p &lt; 0.001 )</td>
</tr>
<tr>
<td></td>
<td>( \text{Quadratic} )</td>
<td>( \text{n.s.} )</td>
<td>( \text{n.s.} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1st Neutral Tone</th>
<th>2nd Neutral Tone</th>
<th>3rd Neutral Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Slope</td>
<td>Intercept</td>
</tr>
<tr>
<td></td>
<td>( \beta = 0.87 )</td>
<td>( \beta = 1.14 )</td>
<td>( \beta = 0.99 )</td>
</tr>
<tr>
<td></td>
<td>( t = 38.15 )</td>
<td>( t = 15.79 )</td>
<td>( t = 57.70 )</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.001 )</td>
<td>( p &lt; 0.001 )</td>
<td>( p &lt; 0.001 )</td>
</tr>
<tr>
<td></td>
<td>( \text{Quadratic} )</td>
<td>( \text{n.s.} )</td>
<td>( \text{n.s.} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1st Neutral Tone</th>
<th>2nd Neutral Tone</th>
<th>3rd Neutral Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Slope</td>
<td>Intercept</td>
</tr>
<tr>
<td></td>
<td>( \beta = 0.87 )</td>
<td>( \beta = 0.78 )</td>
<td>( \beta = 0.18 )</td>
</tr>
<tr>
<td></td>
<td>( t = 38.15 )</td>
<td>( t = 14.48 )</td>
<td>( t = 12.14 )</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.001 )</td>
<td>( p &lt; 0.001 )</td>
<td>( p &lt; 0.001 )</td>
</tr>
<tr>
<td></td>
<td>( \text{Quadratic} )</td>
<td>( \text{n.s.} )</td>
<td>( \text{n.s.} )</td>
</tr>
</tbody>
</table>
Two observations can be made from Figure 6.3. First, neutral tone immediately preceding a Below-NP Boundary (dotted lines) shows a salient raised \( f_0 \) contour, in contrast to the continued \( f_0 \) lowering over the neutral tone syllable that immediately precedes a Subject-Predicate Boundary (solid lines). This was confirmed by the significant statistical results of the slope across both focus conditions in Table 6.1 (see significant results in the 1\textsuperscript{st} Neutral Tone of 1N in Tables 6.1\textsubscript{a}, \textsubscript{b}, the 2\textsuperscript{nd} Neutral Tone of 2N in Tables 6.1\textsubscript{a}, \textsubscript{b}, the 3\textsuperscript{rd} Neutral Tone of 3N in Tables 6.1\textsubscript{a}, \textsubscript{b}).

Second, the magnitude of \( f_0 \) raising was related to the focus conditions of the neutral tone(s). The raising of the neutral tone \( f_0 \) contour is greater in the Pre-Focus condition (Row B, Figure 6.3) than in the On-Focus condition (Row A, Figure 6.3). As can be seen from Row A in Figure 6.3, due to the raising effect, neutral tone immediately before a below-NP boundary surfaced with a relatively leveled-off \( f_0 \) pattern in On-Focus condition. In contrast, this raising effect was magnified in the Pre-Focus condition (Row B, Figure 6.3), and consequently, the neutral tone immediately before the boundary was realized with a salient rising \( f_0 \) pattern.

6.4 Discussion & conclusion

The present study aimed to shed light on two interesting issues that have not been addressed sufficiently on tonal realization in the existing literature. The first concerns with the nature of the weak tonal element in a lexical tone system (i.e., neutral tone), and the second with the effect of prosodic boundary on neutral tone realization.

We set out to investigated the puzzling context-specific (i.e., before T1) rising \( f_0 \) contour of neutral tone reported for Tianjin Mandarin (e.g., Wang, 2002), which posits great challenges to our understanding of the neutral tone realization (Chen & Xu, 2006). We were interested in the underlying mechanism of this rising neutral tone \( f_0 \), either as a special rising neutral tone target as proposed in Wang (2002), or as resulted from the general raising effect introduced by the following low-falling tone (T1), as reported in Li and Chen (2016). We further examined how the \( f_0 \) realization of neutral tone is conditioned by different prosodic boundaries, a question which has not been addressed in the literature.
To investigate the nature of the rising $f0$ over the neutral tone in Tianjin Mandarin, we first replicated the rising neutral tone $f0$ contour before T1 reported in the literature with one neutral tone syllable (with good control over the focus condition of the embedding utterance and the prosodic boundary between the neutral tone and its following lexical T1). We also varied the number of neutral tone syllables (from one to three) to better observe the emergence of the underlying neutral tonal target with an increasingly larger domain of neutral tone realization.

Results showed that, when there was only one neutral tone syllable, the rising neutral tone $f0$ contour as reported in the literature was only observed in the Pre-Focus condition when the neutral tone syllable preceded a low-level prosodic boundary (i.e., a Below-NP Boundary), as shown in Figure 6.2C-1. Furthermore, when the number of neutral tone increased, only the one immediately preceding the Below-NP boundary showed systematically rising $f0$ contours in the Pre-Focus condition (Row C, Figure 6.2).

Our finding thus challenges the received wisdom that there is a special rising neutral tone target before T1 as claimed in Wang (2002) and adopted by virtually all linguists working on TM neutral tone (e.g., Wang & Jiang, 1997; Lu & Wang, 2012). This view can be rejected on two grounds, given our data. First, if the neutral tone before T1 indeed has an underlying rising tonal specification, we should have observed the rising $f0$ tonal contour regardless of the focus or boundary conditions. Second, more importantly, when there are more than one neutral tone syllables, each neutral tone syllable should exhibit rising $f0$ realization so that there would be either consecutive $f0$ rises or a continuously rising $f0$ contour over the sequence of neutral tone syllables. Neither, however, was observed in our data.

Instead, we observed a similar mid-low tonal target as proposed for Standard Chinese (Chen & Xu, 2006). As can be observed from Figure 6.2, regardless of the preceding tonal contexts, the $f0$ realization of neutral tone showed clear tendency of convergence when the number of neutral tone syllables increased. The converging value of neutral tone $f0$ is around the lower mid-level of the speaker’s averaged $f0$ range, except in the Pre-Focus condition before a Below-NP boundary (Row C, Figure 6.2). Furthermore, even when the last neutral tone showed a clearly rising $f0$ contour in Figure 6.2C-3, which was different from other tonal contexts, the approaching of the mid-low target was still observable by the end of the second neutral tone syllable. This indicates that, whatever might have triggered the rising $f0$ contour over the third neutral tone, speakers were clearly aiming to first realize the mid-low target of neutral tone if given some time.

Then the question that arises is what brings about the rising neutral tone $f0$ realization. Previous studies have proposed several other accounts for the $f0$ variability of neutral tone realization, based on the assumption that neutral tone is toneless/targetless. One is the “Tonal Spreading” account proposed by Yip (1980), based on data from Standard Chinese. The surface $f0$ realization of the “toneless” neutral tone was attributed to the tonal spreading from the preceding syllable, given the fact that neutral tone realization in Standard Mandarin seems to be exclusively influenced by the preceding lexical tones (e.g.,
Chao, 1968; Lin & Yan, 1980; Chen & Xu, 2006). This account, however, failed to explain
the converging rising $f_0$ realization after different lexical tones when there are multiple
neutral tone syllables. Others accounted for the neutral tone $f_0$ realization as the transition
(or interpolation) between its preceding and following lexical tonal targets (e.g., Shih, 1987;
vан Santen et al., 1998). Based on our data, this also seems implausible. As the number of
neutral tone syllables increased, the $f_0$ realization of the neutral tone sequence should not
have first reached for a mid-low target and then rose, if the $f_0$ contour over the neutral
tone syllables were simply interpolated between the preceding and following tonal targets.
It is also worth noting that, in the literature, T1 in Tianjin Mandarin is unexceptionally
reported as a low-register tone (see e.g., Chen, 2000; Wee et al., 2005; Hyman, 2007
transcribing T1 as “L”; also see e.g., Wang & Jiang, 1997; Wang, 2002; Ma, 2005 using
“LL”). It would thus make no sense to interpolate to a following L target with the
observed rising neutral tone $f_0$.

We propose that neutral tone does have its own mid-low tonal target as that in
Standard Chinese. What makes different between Tianjin Mandarin and Standard Chinese
is that in Tianjin Mandarin, T1 shows a consistent raising effect over the $f_0$ contour of the
preceding tones (Li & Chen, 2016), which is manifested in terms of either tone sandhi (i.e.,
over T1T1 and T4T1) or anticipatory tonal coarticulation (i.e., over T2T1 and T3T1). It is
therefore plausible that the rising $f_0$ of neutral tone was another manifestation of this T1
raising effect. Together with observations in Chen and Xu (2006) on Standard Chinese, the
cross-linguistic comparison thus lends further support to the vulnerability of prosodically
weak elements produced in contexts, whose acoustic realization is more susceptible to
contextual influence.

6.4.2 Effect of prosodic boundary on neutral tone realization

Our second issue addressed in this study concerns the effect of prosodic boundary on
neutral tone realization. To this end, we manipulated the prosodic boundary between
neutral tone(s) and the following T1. Specifically, we varied the sizes of boundaries from a
low-level boundary (i.e., a below-NP boundary) to a high-level one (i.e., a Subject-Predicate
boundary). In addition, we also controlled the focus status of the neutral tone sequences
(i.e., On-Focus vs. Pre-Focus).

Results showed that the realization of neutral tone(s) was significantly affected by the
size of the following prosodic boundaries, where neutral tone in the below-NP boundary
condition exhibited raised $f_0$ realization compared to that preceding a Subject-Predicate
boundary. Cross-linguistic studies on segmental realization have shown that segments tend
to be resistant to coarticulatory influence across boundaries (e.g., Egido & Cooper, 1980;
Byrd & Saltzman, 1998; Tabain, 2003; Cho, 2004; Pan, 2007), where the bigger the
intervening boundary is, the less contextual coarticulation can be observed. Our results are
thus in line with this observation. The neutral tone preceding a low-level boundary (i.e.,
Below-NP boundary) showed a salient raising effect. When crossing a high-level boundary
(i.e., Subject-Predicate boundary), the raising effect was greatly reduced.

This boundary-regulated T1 raising effect was further modulated by the focus
condition; the raising effect was more clearly observed in the Pre-Focus condition than in the On-Focus condition. This is in line with the general effect of focus on contextual tonal coarticulation. Several studies have shown that the /0 realization of lexical tones was less influenced by its tonal contexts when they were under focus and produced with stronger articulatory force (e.g., Chen & Gussenhoven, 2008; Chen, 2010; Scholz, 2012). This also explains why when the neutral-tone carrying syllable was part of the focused constituent, the raising effect of the following T1 was more suppressed than when the neutral-tone was pre-focus, in which case, the tonal targets were implemented with less articulatory force (Chen & Gussenhoven, 2008) and consequently more prone to the T1 raising effect, leading to greater T1 raising effect.

Our observations thus provided another piece of evidence for the independent implementation of the effects of prosodic boundary and information status, which has been under some debates in the literature. Some studies claimed that these two effects resembled, as the realization of focus is regarded to automatically create prosodic boundaries at either side of the focus domain (e.g., Truckenbrodt, 1999; Gussenhoven, 2008; Büring, 2009; Kabagema-Bilan, López-Jiménez, & Truckenbrodt, 2011), while others have shown evidence suggesting a mismatch between these two effects (e.g., Chen, 2004; Féry, 2007; Féry & Ishihara, 2009; Scholz, 2012; Féry, 2013; Scholz & Chen, 2014a). In our data, it is clear that in both prosodic boundary and focus can significantly influence the neutral tone realization, although with different magnitudes.

In conclusion, by investigating the acoustic realization of the so-called “rising neutral tone” as reported for Tianjin Mandarin, we have shown that the rising neutral tone /0 contour before T1 does not need to be treated as a special tonal target. Rather, we argue that it is due to the raising effect brought about by the following T1. Furthermore, by increasing the number of neutral tone syllables, we observed a mid-low tonal neutral tone target, which is of similar nature to the neutral tone in Standard Chinese. The difference between Standard Chinese and Tianjin Mandarin is therefore not in their neutral tone specification, but rather in the absence vs. presence of the lexical tone /0 raising effect. In Tianjin Mandarin, the low-register falling T1 triggers an anticipatory /0 raising effect (Li & Chen, 2016), which is absent in Standard Chinese. In addition, we observed an interesting effect of prosodic boundaries on neutral tone /0 realization. Specifically, the raising effect due to the following T1 can be blocked by a high-level prosodic boundary (i.e., a Subject-Predicate boundary) compared to a low-level boundary (i.e., a below-NP boundary). Last but not least, we observed an interaction of prosodic boundary and focus, where the effect of boundary was more clearly observed in Pre-Focus condition compared to the On-Focus condition despite the independent implementation of the two effects, arguing further that focus and prosodic boundary are two independent mechanisms that can interact to determine the phonetic implementation of segments as well as lexical tones.