The production and perception of tonal variation -: evidence from Tianjin Mandarin

Li, Qian

The Production and Perception of Tonal Variation

Evidence from Tianjin Mandarin

Universiteit Leiden

Cover: Tonal variation in the dark. Photographed and processed by the author.

ISBN: 978-94-028-0015-9
NUR 616

The Production and Perception of Tonal Variation
Evidence from Tianjin Mandarin

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties

te verdedigen op woensdag 10 februari 2016
klokke 10.00 uur

door

QIAN LI

geboren te Nanjing, China
in 1985

Promotor: Prof. dr. N. O. Schiller

Co-Promotor: dr. Y. Chen

Promotiecommissie:

Prof. dr. C. C. Levelt
Prof. dr. C. Gussenhoven (Radboud University Nijmegen)
Prof. dr. M. Oostendorp (Meertens Institute)
dr. A. Y. Ganushchak
dr. E. D. Botma

ACKNOWLEDGEMENT

This study was funded by a PhD scholarship from the Chinese Scholarship Council. I would also like to acknowledge partial financial support from Leiden University Center for Linguistics, Leids Universiteits Fonds, Deutsche Forschungsgemeinschaft (Project SFB 632), Nederlandse Organisatie voor Wetenschappelijk Onderzoek (VIDI grant No. 016084338 to Yiya Chen), and European Research Center (Starting grant No. 206198 to Yiya Chen). In addition, I would like to express my sincere gratitude to many people who have generously offered great help in this study. Without them, I would not have gained so much during these years.

My biggest thanks go to dr. Frank Kügler for offering me a short-term job in the Informational Structure group at Potsdam University in Germany. It was not only an important financial support during the final stage of my PhD study, but more importantly, it was an opportunity for me to step into topics that are highly related but nevertheless not included in my own PhD project. Thanks for the great experiences!

Special thanks to dr. Ming Yan for his inspiration on a reading-aloud study of neutral tone using eye-tracking techniques. Thanks for sharing your scripts and experiences in eye-tracking and statistics.

I am also grateful to the following scholars in China: dr. Ziyu Xiong, for creating an exhaustive word list for language description; Prof. dr. Feng Shi, Prof. dr. Jinsong Zhang, and dr. Wei Zhou, for generously giving me access to their laboratories and research facilities. Without their precious support, it would not be possible to conduct my experiments.

I thank all my labmates in the Phonetics Lab of Leiden University Center for Linguistics, for the stimulating discussions, for all the practical and spiritual support, and for all the fun we have had. In particular, I am grateful to Jos Pacilly for his helps on Praat scripts from time to time and many helpful suggestions on scripting, and to Ting Zou for being a constant consultant of Tianjin Mandarin. I also owe thanks to Marjolein Barendse for kindly allowing me to watch her eye-tracking experiments, and making eye-tracking alive for me. I thank Rongjia Cui for reaching out a warm hand in my early days in Leiden. I also thank Man Wang for the daily Nespresso time when we share our experiences in experiments or data analysis. Thanks Yifei Bi for sitting opposite me. Your mouse clicking and typing sounds make a productive working day. Thanks Min Liu for your all-time encouragement and comfort whenever you see any depression in me.

Further thanks go to all the other members of the Chinese Phonetic group. Thank you Cesko Voeten for the constructive comments on the English summary and also for translating it into Dutch. I also thank dr. Jessie Nixon, dr. Kalinka Timmer, and Menghui Shi, for their comments and suggestions on earlier reports of my studies.

I owe a deep gratitude to dr. Kurt Debono from SR Research for his timely technical support on the use of Experiment Builder and Data Viewer.

Be friends with learned people. This is a typical dream lifestyle of ancient Chinese scholars. I specially thank all my friends in the Chinese linguistic community to make this dream come true: dr. Chin-Hui Lin, dr. Jingwei Zhang, dr. Ao Chen, dr. Liquan Liu, Yan Gu, Jinling Li, Yang Yang, Zhaole Yang, Xin Li, Anqi Yang, and Shuangshuang Hu. Thank you all for the support and encouragement during these years. Special thanks go to Han Hu for her great help in participant recruitment for the eye-tracking experiments in Beijing. Thank you Yang Yang for sharing your knowledge of syntax. I would always remember the warm help and warm dinners from Hui-Ching Lin in those cold Potsdam winter days. I am really grateful to have met you all in this community. We are not fighting a lone battle. I also owe many thanks to dr. Chui-De Chiu, for the enlightening discussions on statistics. Infinite appreciation and admiration go to dr. Kuo-Song Wang for being a stand-by programmer, and for everything else.

Last but not least, I would like to thank Prof. dr. Wen Cao and Rui Wang, for making all these possible.

To do good wherever we can,
To love liberty above all things,
And never deny truth
Though it be at the throne itself.

- Ludwig van Beethoven

To Youth

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

As an important means of our daily communication, spoken language is notoriously rich in its variability. Therefore a word's pronunciation is often different in connected speech from when it is produced in isolation. For example, in the phrase "green beans", the [n] in the word "green" may assimilate its place of articulation to match that of the following consonant, making [n] more similar to the following [b], resulting the word "green" to sound more like "greem" in this context. In the last decades, such segmental variation in connected speech has received much attention in both fields of linguistics and psycholinguistics. Various theories and models have been proposed to account for the underlying mechanisms of how speakers produce and perceive the variation (see Ernestus, 2012 for a review). Variation of suprasegmental properties, however, is much less studied, probably due to the fact that suprasegmental properties such as prosody is considered to play a relatively minor role in distinguishing lexical meanings in Indo-European non-tonal languages such as English (e.g., Cutler, 1986). This has greatly limited the existing speech production and perception theories within a segment-dominant frame in general.

It is known that of all languages in the world, the majority of them are tonal languages, in which pitch patterns over syllables distinguish lexical meanings (see reviews in e.g., Fromkin, 1978; Yip, 2002; Zsiga, 2012). In some of these tonal languages (i.e., the majority of the African languages and certain American-Indian languages), tones are known to correlate with the morphology and syntax of the language; while in others, tones are exclusively used at the lexical level, which can be found in almost all languages in the Sino-Tibetan family as well as many other languages in the Southeast Asia (Wang, 1967). Emerging evidence has shown that in tonal languages (e.g., Chinese), suprasegmental features such as lexical tones play as an important role in constraining lexical meanings as segments do (e.g., Schirmer, Tang, Penney, Gunter, & Chen, 2005; Lee, 2007; Liu & Samuel, 2007). Understanding variation of lexical tones, therefore, is in great need for the construction of more powerful theories of speech production and comprehension.

One well-known example of a lexical tone system is the four full lexical tones in Standard Chinese. When produced in isolation, Tone 1 (T1) is a high level tone as in ma^1 (妈 'mother'), Tone 2 (T2) a rising tone as in ma^2 (麻 'hemp'), Tone 3 (T3) a dipping tone as in ma^3 (马 'horse'), and Tone 4 (T4) is a falling tone as in ma^4 (骂 'to scold'). These different pitch patterns (i.e., high level, rising, dipping, falling) are found to primarily correlate with how the rate of vocal cords vibration (i.e., *f0*) changes over time (e.g., Chao, 1920). Figure 1.1 illustrates the mean *f0* of four lexical full tones in Standard Chinese, based on six tokens of each tone across eight speakers (Xu, 1997). As can be seen in Figure 1.1, the *f0* contour of T1 stays at the upper *f0* range throughout the whole syllable; T2 typically shows a rising *f0* pattern which rises from the mid range to the upper limit; T3 is realized with a concave *f0* shape which falls to the lower end for the first half of the *f0* and then rises back to the mid range; for T4, the *f0* falls from the high end of the *f0* range to the lower end.

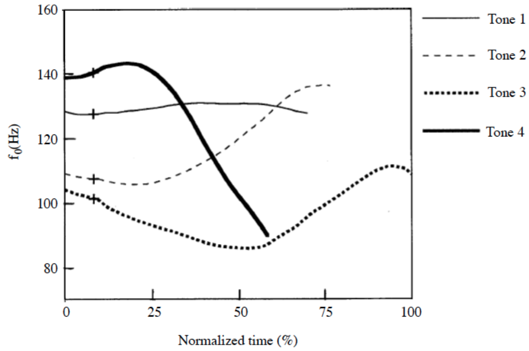

Figure 1.1 *Mean f0 contours of the four Mandarin tones in the syllable /ma/ produced in isolation (Xu, 1997). Time normalized with duration proportional to the mean duration of T3. Averaged across six tokens by eight speakers.*

From the perspective of listeners, the *f0* differences provide important perceptual cues that can be utilized to identify tones (Gandour, 1978; but see e.g., Andruski, 2006; Brunelle, 2009b for tones contrasted in voice quality). These cues include the overall height of the *f0* contour (e.g., Lin & Repp, 1989; Shen, Lin, & Yan, 1993; Francis, Ciocca, & Ng, 2003; Shen et al., 2013), the direction of *f0* change (e.g., Lin & Repp, 1989; Fox & Qi, 1990; Shen & Lin, 1991; Moore & Jongman, 1997; Francis et al., 2003), and the timing of *f0* turning point (e.g., Shen & Lin, 1991; Shen et al., 1993; Moore & Jongman, 1997). For example, as introduced earlier, T2 in Standard Chinese is typically a rising tone. However, due to physiological limitations of the articulators, the *f0* onset of T2 is slightly falling prior to its significant rising *f0* contour. As a result, T2 is very often realized with a concave *f0* pattern similar to that of T3. However, a closer look suggests that the *f0* turning points of the two concave *f0* contours are at different locations, i.e., T2 has an early turning point (if any), while the turning point for T3 is relatively late, as can also be seen in Figure 1.1. Shen and Lin (1991) created a continuum ranging from T2 to T3 in Standard Chinese by manipulating the location of the *f0* turning point. Native listeners were asked to label the stimuli as either T2 or T3 in a forced choice paradigm. Results show that the location of the *f0* turning point, among others, plays an important role in distinguishing T2 vs. T3 in Beijing Mandarin. Stimuli with earlier *f0* turning points were more likely to be labelled as T2 than those with later turning points.

Having known how lexical tones are realized and perceived in isolation is important for our understanding of lexical tones; however, it is far from the whole story. As a matter of fact, in daily speech communication, it is very rare that tones only occur in isolation. Rather, they are more often produced and perceived in contexts, where their *f0* realization usually exhibits extensive variability.

At a local level, the *f0* realization of lexical tones is known to be varied greatly by its local tonal contexts via two different processes. One such local process is known as tone sandhi, which is traditionally defined as the phonological change of lexical tones (see surveys in Wang, 1967; Chen, 2000; also see Zhang, 2010 for a review). An example in Standard Chinese: when two dipping tones (T3) are combined, the first T3 is typically realized with a rising *f0*, resembling that of the rising T2. Although previous impressionistic studies have claimed that the T3 sandhi in Standard Chinese involves the complete tonal change from T3 to T2 (e.g., Chen, 2000; Yip, 2002), this view is challenged by findings from tone production studies, which demonstrate subtle but quite consistent *f0* differences between $T3_{sandhi}$ and the lexical T2 (e.g., Zee, 1980; Peng, 1996; Yuan & Chen, 2014). Furthermore, $T3_{sandhi}$ and the lexical T2 have also been shown to be processed differently during speech encoding (Chen, Shen, & Schiller, 2011; Zhang, Xia, & Peng, 2014). If complete neutralization between $T3_{sandhi}$ and the lexical T2 was indeed involved in tone sandhi, such systematic production differences between the sandhi-derived tone (i.e., $T3_{sandhi}$) and the claimed output tone (i.e., T2) should not have been expected.

Another local contextual process known to introduce tonal variability is tonal coarticulation, which traditionally refers to the phonetic *f0* adjustments to neighboring tonal contexts (e.g., Wu, 1985; Xu, 1997; also see reviews in Xu, 2001; Chen, 2012). Tonal coarticulation has been reported in many East Asian contour-tone languages, such as Thai (e.g., Palmer, 1969; Abramson, 1979; Gandour et al., 1993, 1996; Gandour, Potisuk, & Dechongkit, 1994; Potisuk, Gandour, & Harper, 1997; Zsiga & Nitisaroj, 2007), Vietnamese (e.g., Han & Kim, 1974; Brunelle, 2009a), Standard Chinese (e.g., Shen, 1990; Xu, 1994; Xu, 1997; Kochanski, Shih, & Jing, 2003), Taiwanese (e.g., Peng, 1997; Wang, 2002), and Malaysian Hokkien (e.g., Chang & Hsieh, 2012). Taking an example in Standard Chinese, Figure 1.2 illustrates the mean *f0* realization of the high-level T1 following (Figure 1.2*a*) or preceding different tones (Figure 1.2*b*), based on six tokens produced by eight speakers (Xu, 1997). It is clear in Figure 1.2*a* that, due to different preceding tones, the second high-level T1 is realized with very different *f0* contours in these contexts. The *f0* contours of T1 thus deviate in different magnitudes from the canonical level tonal shape, especially for the first half of the second T1. When T1 is followed by different tones as shown in Figure 1.2*b*, its *f0* also exhibits observable variation, but in a much smaller magnitude (Xu, 1997).

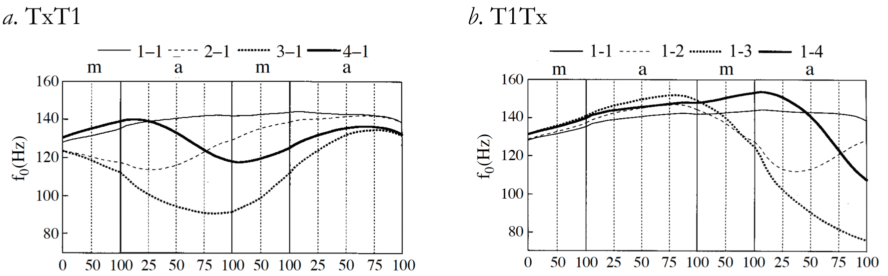

Figure 1.2 *The effect of preceding (a) or following tones (b) on the f0 realization of T1 in /mama/ sequences in Standard Mandarin (Xu, 1997). In (a), the tone in the first syllable is varied from T1 to T4, while the second tone is held constant as T1. In (b), the first tone is held constant, while the second tone is varied from T1 to T4. Averaged across six tokens by eight speakers.*

From the global point of view, the *f0* realization of a lexical tone is further conditioned by the overall prosody of the whole utterance where the tone-bearing lexical item is produced (see reviews in Xu, 2001 and Chen, 2012). Such global influences can come from the grouping of the prosodic units within the utterance (e.g., Shih, 1997; Yang & Wang, 2002; Pan & Tai, 2006; Scholz & Chen, 2014b), the prosodic prominence due to certain information structures (e.g., Xu, 1999; Chen & Braun, 2006; Chen & Gussenhoven, 2008; Chen, 2010; Wang & Xu, 2011), or even sentence types such as declaratives or interrogatives (e.g., Shen, 1991; Liu & Xu, 2005). For example, the *f0* realization of lexical tones in Standard Chinese is found to be subject to different focus statuses of the tone-bearing unit, where tones under focus are usually realized with enhanced distinctiveness of their characteristic *f0* contours compared to those not under focus (e.g., Chen & Braun, 2006; Chen & Gussenhoven, 2008; Chen, 2010).

In addition to variation of the lexical full tones, there exists a special tonal variability phenomenon called "neutral tone" (known in Chinese as 轻声 qing[1] sheng[1] '*light tone*') (Chao, 1920). As illustrated with examples below from Standard Chinese, these syllables are typically grammatical morphemes as in *a*, the unstressed final syllable within a disyllabic lexical item as in *b*, or the final syllable of a reduplicated form as in *c*.

a	我**的**	wo[3] **de**	*possessive marker*
b	玻**璃**	bo[1] **li**	*lexical item*
c	看**看**	kan[4] **kan**	*reduplicated form*

As can be seen from the examples, the neutral tone syllables are never produced independently, but always follow a full tone syllable within a disyllabic domain in Standard Chinese.

Compared to the four full lexical tones, the neutral tone syllables are acoustically reduced and considered to be "unstressed" (see reviews in Liang, 2008, Chen, 2015 and references therein). For example in Standard Chinese, neutral tone syllables are found to be substantially shorter than the full lexical tones (e.g., Lin & Yan, 1980; Shih, 1987). In terms of the *f0* realization, these syllables do not surface with any of the lexical *f0* patterns, but rather are realized with greatly varied *f0* depending on the tone of the preceding full lexical tone syllable (e.g., Shih, 1987; Chen & Xu, 2006 and references therein). Due to the variability, neutral tone has been regarded as "toneless" or "targetless" in most impressionistic studies, and the varied *f0* realization is attributed to the tonal spreading from the preceding lexical tone (e.g., Yip, 1980) or the phonetic interpolation of neighboring lexical tone targets (e.g., Shih, 1987; van Santen, Shih, & Möbius, 1997). However, in a recent study of neutral tone in Standard Chinese, Chen and Xu (2006) investigated the *f0* realization of long stretch of neutral tones. By increasing the number of neutral tone syllables from one to three, the authors show that the *f0* variability of the neutral tone gradually decreases over time and eventually approaches a stable mid-low *f0* range by the end of the third neutral tone syllable. This has thus suggested that neutral tone has its own mid-low tonal target that is independent from its tonal contexts.

It is to be noted that, despite the large repertoire of dialects spoken in China, only Standard Chinese has been studied in depth. Much less well-controlled data have been collected to understand how lexical tones are realized and varied in dialects. For most of the works in Chinese dialects, lexical tone descriptions are done largely within the traditional phonological framework. Variation of tones is mostly investigated on an impressionistic and introspective basis. Detailed variation effects that are not easy to detect with unaided ears are very often ignored. Whenever certain variation is detected, it is very often described based on subjective impression and exclusively attributed to categorical changes (e.g., Chen, 2000). This not only misinforms the theoretical accounts of the variation phenomenon itself, but also greatly restricts our understanding of how listeners produce and recognize tonal variation in general.

To fill this research gap, this dissertation sets out to investigate the tonal variability in Tianjin Mandarin with a series of well-controlled experiments. Tianjin Mandarin serves as a good test case as it is well-known for its tonal sandhi patterns over disyllabic tonal sequences and the iterative and conflicting sandhi applications over trisyllabic tonal sequences, known as the "paradox" in Tianjin tone sandhi (e.g., Chen, 2000). In addition, Wang (2002) reports a special rising *f0* realization of neutral tone for Tianjin Mandarin which challenges the current understanding of neutral tone realization based on evidence from Standard Chinese. It is important to note that although tonal variation in Tianjin Mandarin seems to have been extensively studied in the literature, most previous studies are based on impressionistic data (but see Zhang & Liu, 2011 for an experimental study of tonal coarticulation and tone sandhi in Tianjin Mandarin). Further empirical evidence is thus in great need to shed more light on the underlying mechanisms of these phenomena. Last but not least, an advantage for investigating Tianjin Mandarin is that it is similar with Standard Chinese especially in the segmental aspects as will be shown in Chapter 2, making it easy to compare between the two Mandarin varieties.

Several interesting issues are to be highlighted in this dissertation concerning the production and perception of lexical tones in connected speech. The rest of this chapter briefly introduces these issues and how they have been addressed experimentally in this dissertation.

1.1 Tone sandhi vs. tonal coarticulation

As we have introduced earlier, both tone sandhi and tonal coarticulation give rise to variation in the *f0* realization of a lexical tone, so that the resulting *f0* contour may deviate with varying degrees from the canonical *f0* shape of a lexical tone when produced in isolation. One noteworthy issue is that, although tone sandhi and tonal coarticulation are recognized as two different processes by definition, little is known so far on how to categorize different *f0* variability as being due to tone sandhi or tonal coarticulation.

In the only attempt made so far concerning this issue, Shen (1992) proposes three criteria between the two processes: 1) tone sandhi is constrained by language-dependent morpho-phonemics, while tonal coarticulation is language-independent; 2) tone sandhi could have an assimilatory or dissimilatory nature while tonal coarticulation is only assimilatory; 3) tone sandhi involves the lexical tone change from one to another which is not observed for tonal coarticulation. As will be discussed in Chapter 3, none of these criteria successfully distinguish the two types of tonal variation (also see Chen, 2000 for a criticism). Chen (2000) thus argues that there is "no essential difference" between tone sandhi vs. tonal coarticulation, but defines tone sandhi as the perceptible tonal variation to "trained but unaided ears". Although this definition is shared among almost all studies on tone sandhi in the recent decades, it does not provide us with any real insight into different types of contextual tonal variation, as the impression of individual researchers on what constitutes tone sandhi might differ, which further leads to different theoretical accounts of the tonal alternation phenomenon.

It is notable that tonal sequences that are claimed to undergo tone sandhi by impressionistic studies are very often analyzed separately from tonal coarticulation, leaving no direct comparison of the two types of contextual variation thus far (see e.g., Xu, 1997; Zhang & Liu, 2011 for anticipatory tonal coarticulation analyses excluding tone sandhi sequences). It is therefore yet to be investigated with well-designed experimental data how tone sandhi and tonal coarticulation might individually contribute to the *f0* variability of lexical tones, without which one is unable to objectively attribute different types of *f0* variability to different sources.

Chapter 3 aims to fully understand the various patterns of *f0* deviation due to these two different contextual tonal variation processes. To this end, Chapter 3 investigates all possible combinations of the four lexical tones in Tianjin Mandarin. In particular, this chapter zooms into the effect of the following tones (over the second syllable) on the *f0* realization of the preceding tones (over the first syllable). This is a comparable context to examine both anticipatory tonal coarticulation and tone sandhi as Tianjin Mandarin shows right-dominant tone sandhi patterns. With the technique of Growth Curve Analysis (GCA) (Mirman, 2014), this chapter objectively quantifies whether and how *f0* contours of a lexical tone differ from each other as a function of the following lexical tones. GCA proves to be a powerful method for dealing with the time-varying nature of *f0* contours, which shows a great advantage over traditional statistical analyses of *f0* contours.

Results show that, among the four commonly accepted tone sandhi patterns over disyllabic tonal sequences proposed in the literature (e.g., Li & Liu, 1985; Chen, 2000; Ma, 2005; but see Wee, Yan, & Chen, 2005 for two more disyllabic sandhi sequences), only three have been confirmed. This chapter also yields an interesting anticipatory coarticulatory raising effect of the lexical low-falling tone which has never been reported in the literature. In addition, the confirmed disyllabic sandhi patterns can not be consistently observed within trisyllabic sequences, as claimed in the literature, suggesting a much simpler tone sandhi system of Tianjin Mandarin. Furthermore, by comparing the acoustic realization of the sandhi sequences vs. that of the claimed target patterns according to the literature, Chapter 3 shows that no tonal neutralization can be confirmed over any of the disyllabic or trisyllabic sandhi processes. Rather, based on how sandhi-derived tones resemble their claimed output tones, Chapter 3 further differentiates near-merger sandhi vs. no-merger sandhi, where a near-merger sandhi tone is realized with only subtle *f0* differences from that of the claimed output tone while a no-merger sandhi tone and its claimed output tone show clearer *f0* differences.

1.2 Effect of tonal variability on tone perception

Given the extensive contextual *f0* variation of lexical tones in connected speech, one would further ask how the lexical tones produced in connected speech are perceived. As introduced earlier, listeners are found to primarily rely on the *f0* information during tonal perception. Multiple *f0* cues have been discovered to play crucial roles in the perception of canonical lexical tones that produced in isolation. However, this is insufficient to understand how lexical tones are perceived in connected speech for the following two reasons.

First, the *f0* perceptual cues observed for tones produced in isolation have mainly indicated the most significant *f0* characteristics of the canonical tonal contours. For example in Standard Chinese, a high-level *f0* contour signals a T1 produced in isolation. However, in connected speech, T1 does not always surface with a level *f0* contour. As shown in Figure 1.2*a*, when T1 in Standard Chinese is preceded by four different lexical tones, only the one following T1 is realized with a high-level *f0* contour which is similar to that of the canonical T1 produced in isolation; T1 following T2, T3 or T4 is realized with a salient rising *f0* contour with different rising slopes. Then the question arises how listeners cope with the *f0* variation and identify the tone as T1 in such situation?

Second, it is known that the perception of lexical tones in connected speech is different from the perception of tones produced in isolation, because tonal contexts can greatly influence the identification of the embedded lexical tones (e.g., Lin & Wang, 1985; Fox & Qi, 1990; Xu, 1994; Moore & Jongman, 1997; Francis et al., 2003). For example, in Lin and Wang (1985), it is found that a high-level *f0* contour, which is T1 in Standard Chinese, can be perceived as a rising tone (T2) when it is followed by another tone whose *f0* onset is higher than the high-level tone. It is therefore highly doubtful that to what extent the conclusions made based on lexical tones produced in isolation can be adaptable to tones in connected speech.

Among studies on the perception of lexical tones in connected speech, the main topics discussed in the literature concern 1) how the perception of lexical tones is influenced by the tonal contexts in which they are embedded (e.g., Lin & Wang, 1985; Fox & Qi, 1990; Xu, 1994; Moore & Jongman, 1997; Francis et al., 2003), and 2) whether certain sandhi-derived tones can be distinguished from the claimed output tones according to the tone sandhi rules proposed in the literature (e.g., Wang & Li, 1967; Speer, Shih, & Slowlaczek, 1989; Chen, 2013; Chen, Liu, & Kager, 2015). These studies, though providing us with some insights into how listeners perceive lexical tones in connected speech, have left many issues open for further investigation. For example, one question is how listeners perceive distorted *f0* realization itself without information from the tonal context. Having known that there are various types of *f0* variability when lexical tones are combined so that the *f0* realization is distorted to varying degrees, one further question is whether these different types of tonal variability might affect listeners' perception of lexical tones.

One methodological limitation of previous studies on tone perception is that these studies mainly focused on listeners' end-state performances, as they were required to e.g., make choices among multiple options within a time limit (e.g., Lin & Repp, 1989; Shen et al., 1993; Peng, 1996; Moore & Jongman, 1997; Francis et al., 2003; Chen, 2013; also see McGuire, 2000 for a methodological review on speech perception). Multiple models of speech recognition, e.g., Shortlist Model (Norris, 1994), Cohort Model (e.g., Marslen-Wilson & Tyler, 1980) and TRACE model (McClelland & Elman, 1986), have shown that speech signals are perceived in an incremental fashion, along the unfolding of which listeners actively evaluate the heard signals in real time, and gradually activate the relevant lexical candidates stored in the mental lexicon. Lexical tones are found to be perceived in the similar accumulating way (e.g., Malins & Joanisse, 2010; Shen, Deutsch, & Rayner, 2013). End-state responses, therefore, are only revealing one aspect of this process (see a detailed discussion on this issue in Spivey, 2007).

Chapter 4, via the Visual World Paradigm, investigates the effect of contextual tonal variation on speech recognition in Tianjin Mandarin by monitoring the eye movements of the listeners when presented with different auditory target stimuli. Three types of contextual tonal variability as observed in Chapter 3 have been used as the target stimuli: near-merger sandhi, no-merger sandhi, and no sandhi tonal coarticulation.

The results yield significant perceptual differences among different types of tonal variation, which affect online speech processing differently, as reflected in the different eye movement patterns. No-sandhi coarticulation was easier to recognize than tone sandhi, and between the two sandhi variation types, near-merger sandhi was more difficult to process than no-merger sandhi. This has thus not only shown that in addition to the facilitation of the contextual information in the perception of lexical tones in context, the different degrees of tonal variability exert an effect on tonal recognition, but also suggested the necessity of differentiating tonal variability of different types that have been observed in Chapter 3.

1.3 Nature of tone sandhi

Among various ongoing debates about tone sandhi, one issue further discussed in this dissertation is whether tonal neutralization is involved in the process of sandhi changes. In previous impressionistic studies, tone sandhi in many Chinese dialects has been commonly regarded as the complete tonal change from one lexical tone into another lexical tone within the tonal inventory (e.g., Chao, 1948; Shen, 1992; Yip, 2002; Chen et al., 2015; also see Chen, 2000 for a review). As introduced earlier, it is claimed that in Standard Chinese, when two T3s (dipping tone) are combined, the first T3 is changed into T2 (rising tone) due to the resembling *f0* realization of $T3_{sandhi}$ vs. the lexical T2 (e.g., Chao, 1948). Further evidence from perceptual studies also shows that native listeners cannot reliably distinguish $T3_{sandhi}$ vs. T2 using offline meta-linguistic tonal discrimination tasks (e.g., Wang & Li, 1967; Chen et al., 2015). This has led to the traditional definition of Low tone sandhi as the complete tonal change from T3 to T2. Generally speaking, descriptions of tone sandhi in the literature have been mainly about the categorical change of one tone to another, based on which various tonal alternation theories have been proposed (e.g., Chen, 2000; Wee et al., 2005; Hyman, 2007).

However, as mentioned earlier, such a perception-based conclusion has been challenged by data from production studies. Acoustic studies on Mandarin T3 sandhi have repeatedly observed that there are subtle but consistent differences between a $T3_{sandhi}$ and a lexical T2. For example in Yuan and Chen (2014), data from natural speech corpora show that the *f0* realization of $T3_{sandhi}$ and the lexical T2 are mainly different in the magnitude of *f0* rise and the duration of the *f0* rise. Recent online studies with psycholinguistic approaches further show that the $T3_{sandhi}$ is in fact processed differently from the lexical T2 during production (e.g., Chen et al., 2011; Zhang et al., 2014). If the sandhi-derived tones were completely changed into another lexical tone as claimed, such systematic difference in production would have been unexpected.

Furthermore, if one regards tone sandhi to involve complete tonal neutralization, the potential issue of iterative sandhi applications is expected. The iterativity of sandhi application refers to the phenomenon that the first round of tone sandhi change creates a new context for another round of sandhi application, which is not observed in Standard Chinese but in many other dialects such as Tianjin Mandarin (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Wee et al., 2005). Taking the trisyllabic tonal sequence T3T1T1 in Tianjin Mandarin for example, previous literature has claimed that within T3T1T1, the tonal sequence T1T1 to the right undergoes the first round of tone sandhi change into T3T1, resulting in the intermediate stage T3T3T1, in which the T3T3 sequence to the left further triggers another round of tone sandhi to T2T3, as shown in the claimed rule: T3+<u>T1+T1</u>→<u>T3+T3</u>+T1→T2+T3+T1 (underlines indicating possible sandhi contexts at each stage; e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Wee et al., 2005). There are in total five such iterative application cases claimed for Tianjin Mandarin, which has led to rather complicated tone sandhi application patterns in trisyllabic sequences. If tonal neutralization is indeed not involved in tone sandhi, the iterative tone sandhi application is then called into question.

Both Chapters 3 and 5 tap into this issue. As introduced earlier, the acoustic data in Chapter 3 show significant differences in *f0* realization between the sandhi-derived tones and the claimed output tones in both disyllabic and trisyllabic domains. In addition, Chapter 3 shows no evidence of iterative sandhi applications in the trisyllabic tone sandhi data, further confirming the non-involvement of neutralization in tone sandhi in Tianjin Mandarin. Given the findings in Chapter 3, the question that arises is whether listeners perceive sandhi tones more similarly to the original/underlying tones or to the claimed output tones?

Chapter 5 aims to answer this question by examining the patterns of tone sandhi perception. With the same Visual World Paradigm as in Chapter 4, Chapter 5 investigates the recognition time course of sandhi target stimuli when the targets were presented with different types of competitors. Three competitor types were included for each target which all had segmental overlap with the target: 1) Toneme Competitor, with toneme overlap but *f0* contour difference from the target (Toneme+, Contour-); 2) Contour Competitor, with toneme overlap with that of the claimed output tone, therefore with contour resemblance but toneme difference from the target (Toneme-, Contour+); 3) Segmental Competitor, with no toneme or contour overlap with the target (Toneme-, Contour-). Details aside, the main finding in Chapter 5 is that the eye movement pattern in the Contour Competitor condition exhibits a similar pattern to that in the Segmental Competitor condition during the early stage of processing, suggesting that the claimed output tones are in fact perceived more similarly to unrelated tones. This therefore, from a perceptual perspective, further confirms that tone sandhi does not involve the categorical change from one lexical tone to another in the lexical tonal system.

1.4 Neutral tone & prosodic boundary

While the previous three issues are concerned with the two local contextual tonal variation, i.e., tone sandhi and tonal coarticulation, the current issue explores the effect of a more global factor, i.e., prosodic structure, on the *f0* realization of neutral tone. It has been extensively noted that in speech production, speech utterances are grouped into constituents of various sizes (e.g., Beckman & Edwards, 1994; Ladd, 2008). The organization of these speech constituents is called the prosodic structure; and the boundaries between different constituents are called the prosodic boundaries. Although prosodic structure can sometimes be predicted from the syntactic structure of an utterance, in many cases, it is not isomorphic with the syntactic structure (see reviews in e.g., Shattuck-Hufnagel & Turk, 1996; Frota, 2012).

The role that prosodic boundaries play in speech is not only of chunking the speech, but more of signaling the level that a particular constituent belongs to within the hierarchical system of prosodic structure. The bigger the prosodic boundary, the higher the level in the hierarchy. As a crucial indicator of the prosodic structure of an utterance, prosodic boundaries can be manifested acoustically in various ways. Cross-linguistic evidence has shown that, prosodic boundaries play an important role in conditioning the articulation of segments in ways such as initial strengthening, final lengthening, and coarticulatory resistance. Specifically, the articulation of segments in the initial position within a prosodic domain is often strengthened (e.g., Fougeron & Keating, 1997; Fougeron, 1999; Cho & Keating, 2001; Fougeron, 2001; Keating, 2003; Cho & McQueen, 2005). This strengthening effect is realized in a cumulative way, so that variation due to the boundary effect changes as a function of the hierarchical level of the boundary: the higher the boundary, the greater the strengthening effect. Besides initial strengthening effect, it has also been found in various studies that there is a tendency for segments to be longer before a prosodic boundary than in other contexts, due to the slowing down of the articulation rate before a boundary, known as final lengthening (e.g., Wightman, Shattuck-Hunagel, Ostendorf, & Price, 1992; Kuzla, Cho, & Ernestus, 2007). The degree of final lengthening also varies as a function of the prosodic level: the higher the prosodic boundaries, the greater the degree of final lengthening. Another type of segmental variation brought about by prosodic boundaries is the influence on segmental coarticulation. The extent and magnitude of coarticulation varies as a function of the prosodic boundaries intervening between segments. Evidence from numerous studies has revealed that segments tend to be cumulatively resistant to influences from the contexts. In other words, the higher the level of the prosodic boundary, the less the segmental coarticulation (e.g., Egido & Cooper, 1980; Byrd & Saltzman, 1998; Tabain, 2003; Cho, 2004; Pan, 2007; see also Scholz & Chen 2014a on tonal coarticulation).

Similar with segments, the *f0* realization of lexical tones is also found to be subject to the effect of prosodic boundaries. For example, in Standard Mandarin, it has been found that a rising tone could be realized with a level or even slightly falling *f0* contour between a high and low tonal context when it is located at the middle of the prosodic domain (Xu, 1994). Yang and Wang (2002) show that the *f0* minima of the pre-boundary tone changes as a function of the level of the prosodic boundary: the higher the level of prosodic boundary, the lower the *f0* minima. Furthermore, it has been reported that the Low-tone sandhi (i.e., the realization of a Low tone with a rising pitch contour before another Low tone) in Standard Mandarin could be blocked in certain prosodic groupings (e.g., Shih, 1997; Zhang, 1997; Kuang & Wang, 2006).

Note that in all studies on the effect of prosodic boundaries on lexical tone realization, the subject of study has been limited to the full lexical tones which are realized with lexically distinctive *f0* contours. None of them has examined how neutral tone realization is affected by different prosodic boundaries. This has constrained our understanding of not only the effect of prosodic boundaries on tonal realization, but also the underlying mechanism of neutral tone in various Chinese dialects. It has been reported that the phenomenon of neutral tone exists not only in Standard Chinese, but also in various Chinese dialects (see reviews in Liang, 2008 and Chen, 2015). Many of these dialects have been reported to exhibit interesting neutral tone properties that are different from those of Standard Chinese. For example in Tianjin Mandarin, previous studies have reported consistent neutral tone *f0* realization preceding a lexical low-falling tone (Wang & Jiang, 1997; Wang, 2002; Lu & Wang, 2012; Li & Chen, 2011), which is not observed in Standard Chinese. However, it is noteworthy that most previous studies on neutral tone in dialects are based on impressionistic data, in which the realization of neutral tone is often regarded as the categorical tonal change. As a result, a special rising neutral tone target has been claimed when neutral tone is followed by a lexical low-falling tone in the impressionistic studies of Tianjin neutral tone (Wang & Jiang, 1997; Wang, 2002; Lu & Wang, 2012). This has thus challenged to the mid-low neutral tone target observed in Standard Chinese. Furthermore, studies with only impressionistic observations do not take global prosodic factors (e.g., prosodic boundaries) into account.

To fill this gap, Chapter 6 shows a good example of prosodically-conditioned neutral tone realization in Tianjin Mandarin. With consideration of the effect of prosodic boundaries on neutral tone realization, Chapter 6 shows that the rising neutral tone *f0* contour before T1 does not need to be treated as a special tonal target as claimed, but rather should be due to the raising effect brought about by the following T1 as first proposed in Chapter 3. Neutral tone in Tianjin Mandarin has a mid-low neutral tone target similar to that in Standard Chinese.

In conclusion, while both local and global factors are known to affect both speech production and speech perception, much less has been investigated on how lexical tones are produced and perceived in connected speech. This dissertation aims to gain further understanding of tonal variability through tapping into the afore-mentioned four issues with evidence from Tianjin Mandarin. In the rest of this dissertation, Chapter 2 will first give a comprehensive description of the phonological system of Tianjin Mandarin. Chapter 3 will set out to investigate the *f0* variability induced by tone sandhi vs. tonal coarticulation. The perceptual consequences of the contextual tonal variability will be addressed in Chapter 4. Chapter 5 will set out to further understand the online perception of sandhi tones. In Chapter 6, the effect of prosodic boundaries on neutral tone realization will be investigated. Chapter 7 will summarize the research questions and conclude the main findings throughout the dissertation. Note that Chapters 2-5 are prepared as individual journal articles, and therefore contain some overlapping background information.

CHAPTER 2 LANGUAGE DESCRIPTION OF TIANJIN MANDARIN[1]

2.1 Introduction

Tianjin Mandarin is a member of the northern Mandarin Chinese family. The dialect is spoken in the urban areas of the Tianjin Municipality in People's Republic of China. The city is only about 100 kilometers away from Beijing. While the old variety of Tianjin Mandarin has been reported to differ from Beijing Mandarin in its lack of post-alveolar consonants, the dialect spoken by the younger generation shows basically identical segmental structure to Beijing Mandarin (as also noted in Wee et al., 2005). What differentiates the two dialects is primarily their tonal system. For ease of comparison, the present description on consonants and vowels used the illustrating words from the description of Standard Chinese in Lee & Zee (2003) as much as possible, with deviations mainly due to a different view we have adopted to transcribe the sound system. The sound files illustrated in the present description were produced by a male speaker who was born in the 1980s in Nankai District of Tianjin, one of the oldest districts of Tianjin. The dialect spoken there is regarded as most representative by local people.

2.2 Consonants

There are 25 consonants in Tianjin Mandarin, as shown in Table 2.1. Words illustrating the consonants are shown in Table 2.2 in which most of the syllables have Tone 1, others have Tone 2 or Tone 4. Lexical tones here are marked with a superscript. (See section 2.5 in this chapter for more details on lexical tones and tone sandhi in Tianjin Mandarin.)

Table 2.1 *Overview of consonantal phonemes and their place and manner of articulation.*

	Bilabial	Labio-dental	Denti-alveolar	Post-alveolar	Alveolo-palatal	Velar
Plosive	p pʰ		t tʰ			k kʰ
Affricate			ts tsʰ	t̠ʃ t̠ʃʰ	tɕ tɕʰ	
Nasal	m		n			ŋ
Fricative		f	s	ʃ̠	ɕ	x
Approximant	w			ɹ̠	j ɥ	
Lateral Approximant			l			

[1] A version of this chapter has been submitted for publication as: Qian Li & Yiya Chen. (under revision). Tianjin Mandarin – Illustration of IPA. *Journal of the International Phonetic Association.*

Table 2.2 *Words illustrating consonants.*

p	pɐ^{1} *‘eight’*	t	tɐ^{1} *‘to build’*	tʃ̺	tʃ̺ɐ^{1} *‘residue’*	k	kɤ^{1} *‘song’*
pʰ	pʰɐ^{2} *‘to climb’*	tʰ	tʰɐ^{1} *‘he/she’*	tʃ̺ʰ	tʃ̺ʰɐ^{1} *‘to insert’*	kʰ	kʰɤ^{1} *‘subject’*
m	mɐ^{1} *‘mother’*	ts	tsɐ^{1} *‘to circle’*	ʃ̺	ʃ̺ɐ^{1} *‘sand’*	ŋ	ɑŋ^{2} *‘to raise’*
f	fɐ^{1} *‘to send’*	tsʰ	tsʰɐ^{1} *‘to wipe’*	tɕ	tɕjɐ^{1} *‘to add’*	x	xɤ^{1} *‘to drink’*
w	wɐ^{1} *‘frog’*	n	nɐ^{4} *‘to include’*	tɕʰ	tɕʰjɐ^{1} *‘to nip off’*	j	jɐ^{1} *‘duck’*
		s	sɐ^{1} *‘to cast’*	ɕ	ɕjɐ^{1} *‘shrimp’*	ɥ	ɥe^{1} *‘to restrict’*
		l	lɐ^{1} *‘to pull’*	ɹ̺	ɹ̺ən^{2} *‘person’*		

2.2.1 Plosive

Plosives in Tianjin Mandarin differentiate three different places of articulation: bilabial /p, pʰ/ as in /pɐ^{1}/ *‘eight’* and /pʰɐ^{2}/ *‘to climb’*, denti-alveolar /t, tʰ/ as in /tɐ^{1}/ *‘to build’* and /tʰɐ^{1}/ *‘he/she’*, and velar /k, kʰ/ as in /kɤ^{1}/ *‘song’* and /kʰɤ^{1}/ *‘subject’*. They are contrastive in aspiration; aspirated plosives have longer VOT than their unaspirated counterparts. The contrast holds for all places of articulation, as shown in Table 2.3. Averaged across different places of articulation, the mean VOT for aspirated plosives in our dataset is 102 ms while that for unaspirated plosives is only 23 ms. Furthermore, velar plosives have significantly longer VOT than bilabial and denti-alveolar plosives as revealed by a *post-hoc* Tukey HSD test (velar vs. bilabial: Diff.=20ms, p-adj.<0.001; velar vs. denti-alveolar: Diff.=10ms, p-adj.<0.001), while the bilabial and denti-alveolar plosives are not significantly different.

Table 2.3 *VOT of unaspirated vs. aspirated plosives in three different places of articulation, based on all plosive consonants in a dataset of 3935 monosyllables.*

	Bilabial		Denti-alveolar		Velar	
	unaspirated	aspirated	unaspirated	aspirated	unaspirated	aspirated
Mean	16 ms	91 ms	16 ms	100 ms	36 ms	116 ms
SD	6 ms	16 ms	4 ms	20 ms	13 ms	22 ms
N	121	188	194	145	169	106

2.2.2 Affricate

Affricates in Tianjin Mandarin display the same two-way distinction in aspiration as plosives. They have three places of articulation: denti-alveolar /ts, tsʰ/ as in /tsɐ^1/ *'to circle'* and /tsʰɐ^1/ *'to wipe'*, post-alveolar /tʃ̪, tʃ̪ʰ/ as in /tʃ̪ɐ^1/ *'residue'* and /tʃ̪ʰɐ^1/ *'to insert'*, as well as alveolo-palatal /tɕ, tɕʰ/ as in /tɕjɐ^1/ *'to add'* and /tɕʰjɐ^1/ *'to nip off'*.

The denti-alveolar affricates are produced with the tip of the tongue behind the lower front teeth and the tongue blade against the alveolar ridge. The post-alveolar affricates are pronounced with the tongue tip raised against the post-alveolar region. The alveolo-palatal affricates are pronounced with the tongue tip down behind the lower front teeth and with the dorsum of the tongue against the area between the alveolar ridge and the hard palate.

The alveolo-palatal consonants in Tianjin Mandarin are obligatorily followed by an alveolo-palatal glide before the following vowel if the consonant is not followed by a high vowel, such as in /tɕjɐ^1/ *'to add'*, /tɕʰjɐ^1/ *'to nip off'* and /ɕjɐ^1/ *'shrimp'*. Figure 2.1 shows the spectrogram of /tɕjɐ^1/ *'to add'*, where we see a transition (glide /j/) between the consonant /tɕ/ and the vowel /ɐ/, taking up about ¼ of the total rhyme length. The F1 of the glide starts with a low value (around 500Hz) and gradually increases up to 850Hz as that of the vowel /ɐ/; the F2 of the glide starts around 2000Hz and ends with 1500Hz. The F1 and F2 values of the glide onset therefore resemble that of a front high vowel. Glides in such contexts have been traditionally considered part of the rhyme and transcribed as a vowel (see, e.g., Lee & Zee, 2003 for the transcript of the same syllables in Standard Chinese as /tɕia^1/, /tɕʰia^1/ and /ɕia^1/, respectively). In line with Duanmu (2000) and Lin (2007), we treat them as glides which constitute part of the onset, so as to avoid the ill-motivated idiosyncratic feature of Tianjin Mandarin as having triphthongs.

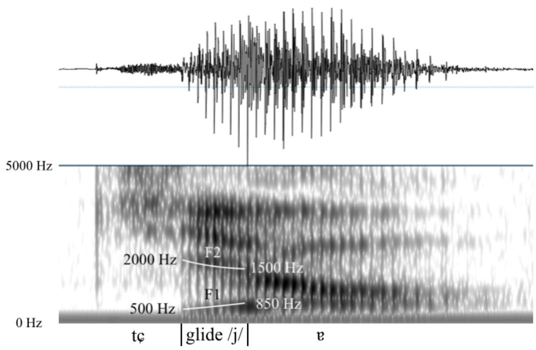

Figure 2.1 *The spectrogram of a /tɕjɐ^1/ syllable.*

It is important to note the difference between Tianjin Mandarin and Shanghai Chinese. Chen and Gussenhoven (in press) have argued that the transition between an alveolo-palatal affricate and its following vowel in Shanghai Chinese is phonetic and predictable from the context, and should not be treated phonemically as /j/. For comparison between the two dialects, we have plotted the spectrograms of three segmentally comparable monomorphemic words /tɑɔ^1/ *'knife'*, /tjɑɔ^1/ *'to hold in the mouth'*, and /tɕjɑɔ^1/ *'to teach'* as those in Figure 2.2 in Chen and Gussenhoven (in press). Different from the observation in Shanghai Chinese where the transition between the alveolo-palatal consonant and the vowel is rather brief, here we see that in Tianjin Mandarin, there is a clear glide-like transition in /tɕjɑɔ^1/ (right in Figure 2.2), which is similar to that in /tjɑɔ^1/ (middle in Figure 2.2) where there is a real glide, and different from that in /tɑɔ^1/ (left in Figure 2.2) where there is only subtle phonetic coarticulation.

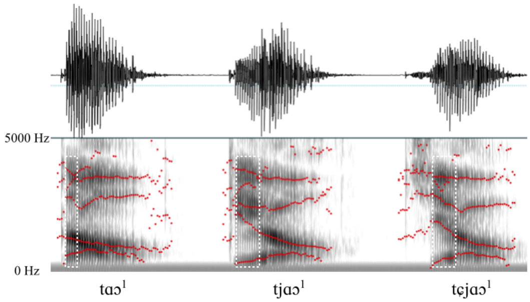

Figure 2.2 *The spectrograms of /tɑɔ^1/, /tjɑɔ^1/ and /tɕjɑɔ^1/. Dashed frames indicate the transitions.*

It is worth noting that alveolo-palatals /tɕ/, /tɕʰ/ and /ɕ/ can directly occur only before the high front vowels /i/ and /y/ (and their corresponding glides /j/, /ɥ/), and are therefore in complementary distribution with the denti-alveolars /ts/, /tsʰ/, /s/, the post-alveolars /tɕ/, /tɕʰ/, /ɕ/, as well as the velars /k/, /kʰ/ and /x/. Their phonemic status has been under some debates, which we will not pursue further here but refer interested readers to Lin (2014) for further details.

2.2.3 Nasal

The nasal series in Tianjin Mandarin contain three different places of articulation: bilabial /m/ as in /mɐ^1/ *'mother'*, denti-alveolar /n/ as in /nɐ^4/ *'to include'*, and velar /ŋ/ as in /ɑŋ^2/ *'to raise'*.

2.2.4 Fricative

Fricatives in Tianjin Mandarin differentiate five different places of articulation: labiodental /f/ as in /fɐ^1/ *'to send'*, denti-alveolar /s/ as in /sɐ^1/ *'to cast'*, post-alveolar /ʃ̺/ as in /ʃ̺ɐ^1/ *'sand'*, alveolo-palatal /ɕ/ as in /ɕjɐ^1/ *'shrimp'*, and velar /x/ as in /xɤ^1/ *'to drink'*, which alternates with the uvular fricative /χ/ when followed by a non-high vowel as in /xaɛ^2/ ([χaɛ^2]) *'child'* and /xɑɔ^3/ ([χɑɔ^3]) *'good'*.

2.2.5 Approximant

Tianjin Mandarin has four central approximants /w, ɹ̠, j, ɥ/ as in /wɐ¹/ *'frog'*, /ɹ̠ən²/ *'person'*, /jɐ¹/ *'duck'* and /ɥe¹/ *'to restrict'*, which mainly differ in their places of articulation, and one lateral approximant /l/ as in /lɐ¹/ *'to pull'*. Among them, /w/, /j/ and /ɥ/ can serve as part of a complex onset. /j/ and /ɥ/ are mostly in complementary distribution except for one context; before the vowel /e/, they are contrastive as in /nje¹/ *'to pinch'* vs. /nɥe⁴/ *'to abuse'*, /tɕje¹/ *'to connect'* vs. /tɕɥe²/ *'to feel'*, /tɕʰje¹/ *'to cut'* vs. /tɕʰɥe¹/ *'to lack'*, and /ɕje¹/ *'to rest'* vs. /ɕɥe¹/ *'boots'*.

2.2.6 Syllabic consonant

As in Beijing Mandarin, Tianjin Mandarin has two syllabic consonants: the denti-alveolar /z̩/ (as in /tsz̩³/ *'son'*) and post-alveolar /ɹ̩/ (as in /t̠ʃɹ̩³/ *'paper'*). They are homorganic with the preceding onset. /z̩/ follows denti-alveolar obstruents /ts, tsʰ, s/ (as in /tsz̩³/ *'son'*, /tsʰz̩²/ *'word'*, /sz̩¹/ *'to think'*), while /ɹ̩/ only following post-alveolar obstruents /t̠ʃ, t̠ʃʰ, ʃ̠/ (as in /t̠ʃɹ̩³/ *'paper'*, /t̠ʃʰɹ̩¹/ *'to eat'*, /ʃ̠ɹ̩¹/ *'poem'*). In addition, /ɹ̩/ can occur by itself as in /ɹ̩⁴/ *'the sun'*. For further details of their acoustic and articulatory realizations, readers are referred to Lee and Zee (2014) and Lee-Kim (2014). While our transcriptions are in line with Chao (1948), it is important to note that the literature witnesses quite some debates in both the phonemic status and notation of the two sounds (see also review in Lee-Lim, 2014). Traditionally, Sinologists refer to them as "apical vowels" (e.g., Karlgren, 1915-1926). They have also been referred to as "fricative vowels" in Ladefoged and Maddieson (1996), syllabic approximant /ɹ̩/ in Lee and Zee (2003, 2014), and syllabic dental approximant /ɹ̪/ and retroflex approximants /ɻ/, respectively in Lee-Kim (2014).

2.3 Vowels

Vowels in Tianjin Mandarin consist of eight monophthongs and four diphthongs. Like consonants, vowels in Tianjin Mandarin are identical to those in Beijing Mandarin. However, our transcription adopted here is different from that in Lee and Zee (2003), mainly due to different treatments (i.e., vowel vs. glide) of the so-called "onglide" and the notations that we chose for the diphthongs. Words with denti-alveolar onset consonants were used to reduce contextual influence from the preceding consonant.

2.3.1 Monophthongs

The eight monophthongs in Tianjin Mandarin as well as their respective allophones are illustrated in Table 2.4. Vowels produced in different contexts are treated as phonemes due to their clearly different vowel qualities although they are not contrastive and can be treated as allophones (e.g., Lin, 2007) or contextual variation as in Lee and Zee (2013).

Table 2.4 *Words illustrating monophthongs.*

	Open Syllable		Closed Syllable			
			before coda /n/		before coda /ŋ/	
i	ti^{1}	*'low'*	lɪn^{2}	*'forest'*	lɪŋ^{2}	*'zero'*
y	ly^{2}	*'donkey'*	tɕʏn^{1}	*'army'*		
e	tje^{1}	*'dad'*	tjɛn^{1}	*'bump'*		
ə	tə	*possessive marker*	tən^{4}	*'to drag'*	təŋ^{1}	*'lamp'*
ɐ	tɐ^{1}	*'to build'*	tan^{1}	*'single'*	tɑŋ^{1}	*'when'*
u	tu^{1}	*'metropolis'*			tʊŋ^{1}	*'east'*
ɤ	tɤ^{2}	*'to get'*				
o	two^{1}	*'more'*				

Figure 2.3*a* shows the mean F1 and F2 values of 50 samples for each monophthong produced in open syllables by measuring the vowel midpoint. /ə/ is not included in Figure 2.3*a* because open syllables with /ə/ can only occur as neutral-tone syllables which cannot be produced in isolation. (See section 5.3 in this chapter for more details.) Monophthongs occurring in closed syllables with denti-alveolar and velar nasal coda are plotted in Figure 2.3*b*.

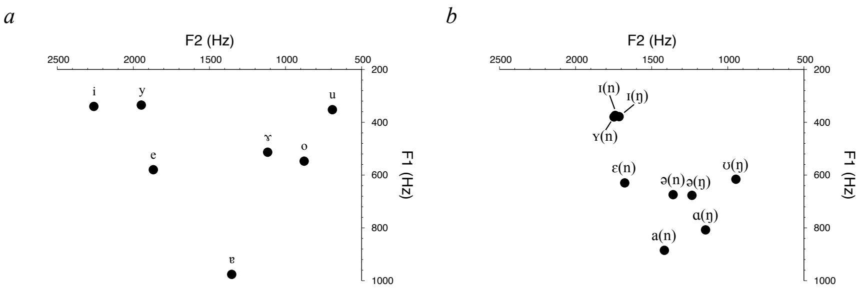

Figure 2.3 *F1 and F2 values of monophthongs produced in open syllables (a) and their respective allophones in closed syllables (b).*

There are three front high vowels in Tianjin Mandarin. /i/ and /y/ (as in /ti^{1}/ *'low'* and /ly^{2}/ *'donkey'*) are front close vowels contrasting only in the rounding of the lips. In closed syllables, /i/ and /y/ are realized as their lax counterparts /ɪ/ as in /lɪn^{2}/ *'forest'* and /lɪŋ^{2}/ *'zero'* and /ʏ/ as in /tɕʏn^{1}/ *'army'*, respectively. When these two high front vowels are followed by nasal codas, an offglide [ə] is inserted as the articulation of the vowel transits to the vocal configuration for the following coda, which is illustrated in Figure 2.4 where the spectrogram of the syllable /t^{h}ɪŋ^{1}/ ([t^{h}ɪ^{ə}ŋ^{1}] *'to listen'*) in Tianjin Mandarin is plotted against that of the syllable /sɪŋ/ *'sing'* in American English (Ladefoged, 1999). In the latter, there is a clear and sharp acoustic boundary between the vowel and the nasal coda without the presence of a transitional schwa.

/e/ is a front close-mid vowel, which can occur in open syllables; while in closed syllables, we only observe the front open-mid vowel /ɛ/. In both syllable types, a preceding glide (i.e., /j/, /ɥ/, /w/) is obligatory as in /tje^{1}/ *'dad'* and /tjɛn^{1}/ *'bump'*. Furthermore, in closed syllables, /ɛ/ only occurs before the denti-alveolar nasal coda /n/.

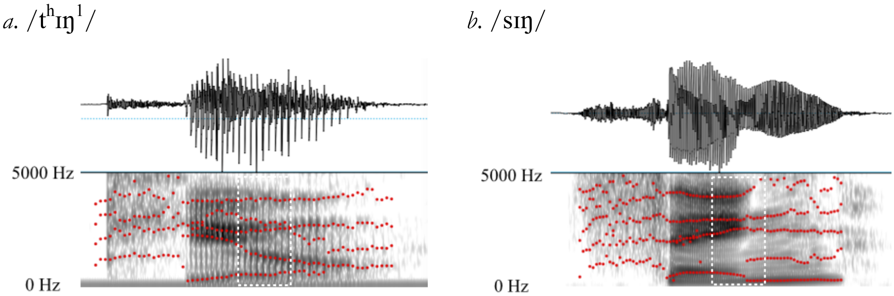

Figure 2.4 *Spectrogram comparison of the syllable /t^{h}ɪŋ^{1}/ in Tianjin Mandarin (a) vs. the syllable /sɪŋ/ in American English (b).*

The central vowel /ə/, commonly known as schwa, can occur in open syllables with only neutral tone as in /tə/ *possessive marker* or closed syllables with both nasal coda /n/ and /ŋ/ (as in /tən^{4}/ *'to drag'* and /təŋ^{1}/ *'lamp'*). As can be seen from Figure 2.3*b*, in closed syllables, the realization of /ə/ is slightly influenced by different following nasal codas due to their closure gestures at different places of articulation. Before the alveolar nasal coda /n/, /ə/ is slightly more frontal than that before the velar nasal coda /ŋ/.

/ɐ/ is a mid low vowel, which is transcribed as the cardinal /a/ in Lee and Zee (2003). /ɐ/ can occur in open syllables as in /tɐ^{1}/ *'to build'*. In closed syllables, /ɐ/ is realized differently according to different nasal codas, as the front low /a/ before denti-alveolar nasal coda /n/ (as in /tan^{1}/ *'single'*) and as the back low /ɑ/ before velar nasal coda /ŋ/ (as in /tɑŋ^{1}/ *'when'*). Lee and Zee (2003) treat the three variants in Beijing Mandarin as the same phoneme. We would like to highlight their salient differences in the vowel quality and opted to transcribe them as three different phonemes.

/u/ is a back close vowel. It can occur in open syllables as in /tu^{1}/ *'metropolis'*; in closed syllables before the velar nasal coda /ŋ/, it is realized as its lax counterpart /ʊ/ as in /tʊŋ^{1}/ *'east'*.

Both /ɤ/ and /o/ are back close-mid vowels differing in the lip shapes as in /tɤ^{2}/ *'to get'* and /two^{1}/ *'more'* where an onglide /w/ is obligatory before /o/.

2.3.2 Diphthongs

Diphthongs in Tianjin Mandarin can only occur in open syllables, as illustrated in Table 2.5. Figure 2.5 shows the mean F1 and F2 values of 50 samples for each diphthong by measuring the respective midpoint of the two parts in the vowel. Arrows in Figure 2.5 demonstrate the trajectories of the gliding.

Table 2.5 *Words illustrating diphthongs.*

eɪ	leɪ^{2}	*'thunder'*
aɛ	taɛ^{1}	*'dull'*
ɑɔ	tɑɔ^{1}	*'knife'*
əʊ	təʊ^{1}	*'all'*

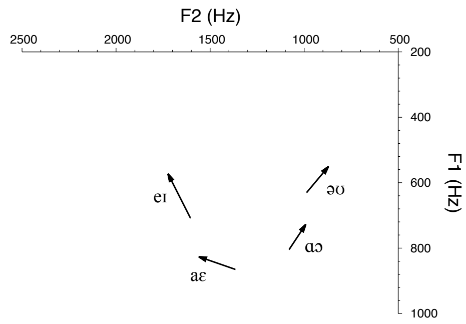

Figure 2.5 *F1 and F2 values of diphthongs. Gliding trajectories are shown with arrows.*

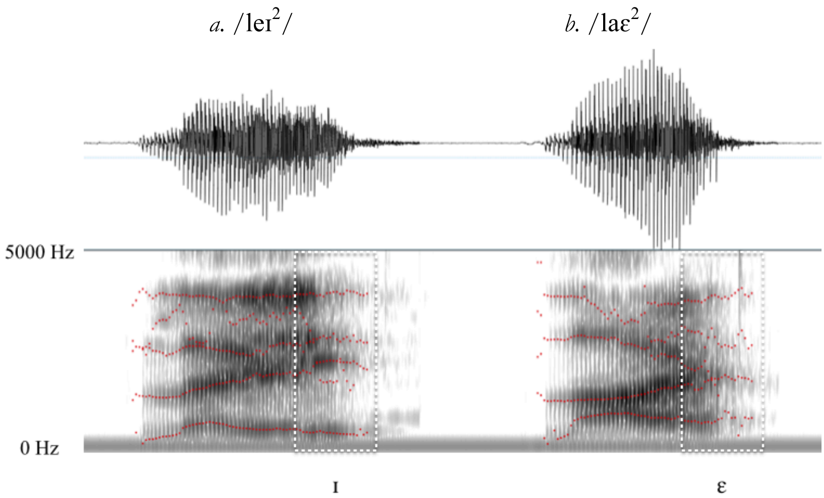

Figure 2.6 *Spectrogram comparison of diphthongs /eɪ/ vs. /aɛ/ as in syllables /leɪ^{2}/ 'thunder' (a) vs. /laɛ^{2}/ 'to come' (b), respectively.*

There are four diphthongs in Tianjin Mandarin, among which /eɪ, aɛ/ gliding towards the front (as in /leɪ^{2}/ *'thunder'* and /taɛ^{1}/ *'dull'*) and /ɑɔ, əʊ/ towards the back (as in /tɑɔ^{1}/ *'knife'* and /təʊ^{1}/ *'all'*). /eɪ/ and /aɛ/ are frequently transcribed as /ei/ and /ai/, both gliding towards the same high front target /i/ (e.g., Lee & Zee, 2003; Lin, 2007). However, this is not confirmed by our data. First, we can see from Figure 2.5 that, neither /eɪ/ or /aɛ/ really reaches the high front region at the offset part; second, as shown in Figure 2.6, the ending point of /eɪ/ vs. /aɛ/ have very different spectrograms especially in terms of F1 and F2, where the offset part of /eɪ/ shows clearly lower F1 but higher F2 values than that of /aɛ/. The same differences can also be observed for /ɑɔ/ vs. /əʊ/, where both have been frequently transcribed to glide towards the back high vowel /u/ (e.g., Lee & Zee, 2003; Lin, 2007).

2.3.3 Rhotic vowel and er-hua

There is a rhotic vowel /ɚ/ in Tianjin Mandarin which is produced as an r-colored schwa, with the tip of tongue up at the end of the vowel. The syllable /ɚ/ can only occur independently in three syllables: /ɚ^2/ *'son'*, /ɚ^3/ *'ear'* and /ɚ^4/ *'two'*. When /ɚ/ is produced with neutral tone, it is used as a diminutive suffix.

When adding the suffix /ɚ/ to a noun, typically, the two syllables are coalesced into one rhoticized syllable in the output form, with the vowel part of the preceding syllable directly rhoticized and only the lexical tone of the preceding syllable is kept, e.g., /nɐ^4/+/ɚ/→/nɐ˞^{4}/ *'there'* (as compared to /nɐ^4/ *'that'*). There are also other rhoticizing processes in the following types of syllable structures:

In open syllables, if the vowel is /aɛ/, only the /a/ part is rhoticized while /ɛ/ is deleted, e.g., /pʰaɛ^2/ + /ɚ^2/ → /pʰa˞^{2}/ *'badge'* (as compared to /pʰaɛ^2/ *'card'*); if the rhyme is /e/, /eɪ/ or syllabic consonants, the entire rhyme part is replaced with /ɚ/, e.g., /peɪ^4/ + /ɚ^2/ → /pɚ^4/ *'very'* (as compared to /peɪ^4/ *'double'*).

In closed syllables, if the coda is /n/, only the vowel is rhoticized, while the coda is deleted, e.g., /pan^4/ + /ɚ^2/ → /pa˞^{4}/ *'partner'* (as compared to /pan^4/ *'campanion'*); if the coda is /ŋ/, the coda is realized with the nasalization of the vowel, and the entire nasalized rhyme is rhoticized, e.g., /xwɑŋ^2/ + /ɚ^2/ → /xwɑ̃˞^{2}/ *'yolk'* (as compared to /xwɑŋ^2/ *'yellow'*). Similar rhoticization processes have been reported in Lee and Zee (2014) for Beijing Mandarin (although the vowels are transcribed differently).

2.4 Syllable structure & phonotactics

The syllable structure in Tianjin Mandarin is (C/G)V(C). C stands for consonants. Except for /ŋ/, all consonants can occur at the initial onset position of a syllable. G stands for glides such as /j/ in /jɐ^1/ *'duck'*, /w/ in /wɐ^1/ *'frog'* and /ɥ/ in /ɥe^1/ *'to restrict'*. V stands for vowels. Coda is optional, and only one nasal consonant (/n/ or /ŋ/) is allowed (as in /tan^1/ *'single'* or /tɑŋ^1/ *'when'*). /n/ is the only consonant that appears both at the beginning and end of a syllable (as in /nɐ^4/ *'to include'* and /tan^1/ *'single'*). Onsetless syllables are also possible, as in /aɛ^1/ (*'sad'*).

In addition, there are some co-occurrence restrictions of consonants and vowels in Tianjin Mandarin:

1) Alveolo-palatals /tɕ, tɕʰ, ɕ/ can only occur before front high vowels /i, y/ (as in /tɕi^1/ *'chicken'*, /tɕʰi^1/ *'seven'*, /ɕi^1/ *'west'*, /tɕy^1/ *'to live'*, /tɕʰy^1/ *'maggot'*, /ɕy^1/ *'needs'*), or their corresponding glides /j, ɥ/ (as in /tɕje^1/ *'to connect'*, /tɕʰje^1/ *'to cut'*, /ɕje^1/ *'to rest'*, /tɕɥe^2/ *'to feel'*, /tɕʰɥe^1/ *'to lack'*, /ɕɥe^1/ *'boots'*).

2) Complementary to 1), front high vowels /i, y/ and the corresponding glides /j, ɥ/ cannot follow denti-alveolar /ts, ts^h, s/, post-alveolar /tʃ̻, tʃ̻^{h}, ʃ̻/, or velar /k, k^h, x/; however, /y/ and /ɥ/ can follow denti-alveolar /n, l/ (as in /ny^3/ *'woman'*, /ly^2/ *'donkey'*, /nɥe^4/ *'to abuse'*, /lɥe^4/ *'to omit'*).

3) Mid-close vowels /e, o/ have to co-occur with glides /j, w/, respectively, as in /tje^1/ *'dad'*, /two^1/ *'more'*; back mid-close vowel /ɤ/ cannot follow labial consonants /p, p^h, m, f/.

4) Glide /w/ cannot be followed by front vowels while /ɥ/ only followed by front vowels (as in /ɥe^1/ *'to restrict'*). /j/, however, can be followed by front, central and back vowels as in /ji^1/ *'one'*, /ja^1/ *'duck'*, /jɑŋ^1/ *'central'*.

2.5 Tones

Tianjin Mandarin differs from Beijing Mandarin mainly in the tonal system. It has a different tonal inventory. Moreover, when tones are combined, Tianjin Mandarin exhibits a complex pattern of tonal variability, different from that in Beijing Mandarin.

2.5.1 Lexical tones in isolation

There are four lexical tones in Tianjin Mandarin. Different *f0* values have been proposed to describe the tonal contours (e.g. Li & Liu, 1985; Shi, 1986; Wee et al., 2005). Figure 2.7 shows the *f0* contours of the four lexical tones elicited in isolation with obstruent onsets. Each tonal contour was obtained by averaging across 50 samples produced by the same male speaker. The *f0* values were normalized so that it can be interpreted into the five-scale pitch system using the T-normalization method developed by Shi (1986). Here the intervals 0-1, 1-2, 2-3, 3-4, and 4-5 correspond to 1–5 in Chao (1920)'s pitch annotation system, respectively.

As illustrated in Figure 2.7, Tone 1 (hereafter referred to as T1) is a low-falling tone, of which pitch contour falls from the mid to the lower end of the speaker's pitch range, as in /lɑɔ^1/ *'to dredge up'*. Tone 2 (T2) is a high-rising tone, whose pitch contour rises from the mid to the upper end of the pitch range, as in /lɑɔ^2/ *'hard-working'*. Tone 3 (T3) is a dipping tone, which falls slightly from the lower pitch range, stays at the bottom and then rises to the mid pitch range of the speaker, as in /lɑɔ^3/ *'old'*. Tone 4 (T4) is a high-falling tone which falls from the upper end to the mid of the pitch range, as in /lɑɔ^4/ *'to flood'*. It is noticeable that T1 and T4 show very comparable falling *f0* patterns but in different registers. T1 was realized in a lower register while T4 in a relatively higher one. Adopting the pitch range scale in Chao (1920), T1 can be transcribed as /31/, T2 as /45/, T3 as /213/, T4 as /53/.

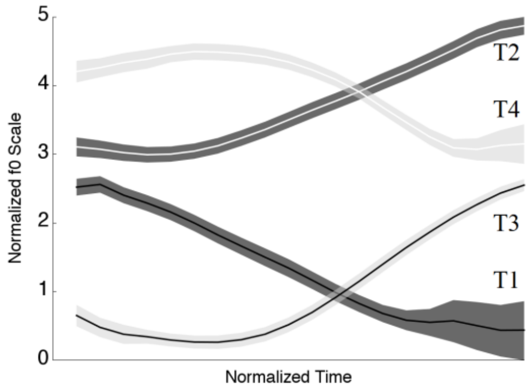

Figure 2.7 *Lexical tones in isolation. Lines stand for the mean. Gray areas stand for ±1 standard error of mean. Tone 1 (T1) is illustrated with black line and dark gray area; Tone 2 (T2) with white line and dark gray area; Tone 3 (T3) with black line and light gray area; Tone 4 (T4) with white line and light gray area. Normalized time.*

It can be noted that, the lexical tones of Tianjin Mandarin are different from that of Beijing Mandarin except for the dipping tone T3. In Beijing Mandarin, T1 is a high level tone. Contrastingly, there is no level tone in Tianjin Mandarin. More importantly, register, among other acoustic cues, seems to play an important role in falling tone discrimination in Tianjin Mandarin, which is not normally found in Beijing Mandarin.

2.5.2 Tone variability in connected speech

When lexical tones are produced in connected speech, the *f0* realization is usually varied either phonetically due to tonal coarticulation or phonologically due to tone sandhi.

2.5.2.1 Tonal coarticulation

Tonal coarticulation in Tianjin Mandarin is bi-directional, including left-to-right carryover effect and right-to-left anticipatory effect. Similar to Beijing Mandarin and most other East-Asian tonal languages, carryover effect in Tianjin Mandarin is assimilatory in nature, while anticipatory effect is dissimilatory. Carryover effect tends to be greater than anticipatory effect both in terms of the magnitude and the temporal extent. In Tianjin Mandarin, carryover effect can be observed in all tonal contexts except when the second tone is the low falling T1, while anticipatory effect is only triggered by low tones (i.e., T1 and T3). Figure 2.8 illustrates the two coarticulatory effects in Tianjin Mandarin. Each tonal contour was obtained by averaging across 12 samples produced by the male speaker. For more details on tonal coarticulation in Tianjin Mandarin, readers are referred to Zhang and Liu (2011) as well as Li and Chen (2016).

As shown in Figure 2.8*a*, the *f0* of a tone can be realized differently due to different preceding tones: when T2 is preceded by a high tone such as T4 (as in /tɕi^4 məʊ^2/

'stratagem'), the onset *f0* realization of the second T2 is clearly higher than that following a low tone such as T1 (as in /kweɪ^{1} mwo^{2}/ *'scale'*). In Figure 2.8*b*, the anticipatory effect is shown as the first tone is realized differently due to different second tones: when T2 is followed by a low tone such as T1 (as in /paɛ^{2} mɑɔ^{1}/ *'white cat'*), the offset *f0* realization of the first T2 shows faster rate of *f0* rising than that before a high tone as T4 (as in /tʃ̺^{h}əʊ^{2} mi^{4}/ *'dense'*).

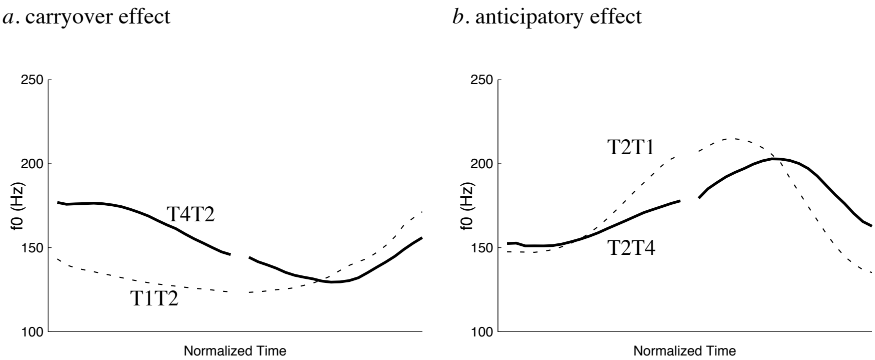

Figure 2.8 *f0 realization of T2 when connected with a high tone (T4) vs. a low tone (T1). T2 as the second tone in a, T2 as the first tone in b. Normalized time.*

2.5.2.2 Tone sandhi

Previous impressionistic studies have reported multiple disyllabic tone sandhi rules in Tianjin Mandarin (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; Wang, 2002; Wee et al., 2005; Hyman, 2007). Three patterns have been confirmed with experimental data: T1T1, T3T3, T4T1 (Li & Chen, 2016), with no complete neutralization with their respective targeted ouputs as proposed in the rules (see also Zhang & Liu, 2011 for similar observations). Figure 2.9 shows the *f0* contours of the three tone sandhi patterns. Each tonal contour was obtained by averaging across 12 samples produced by the male speaker.

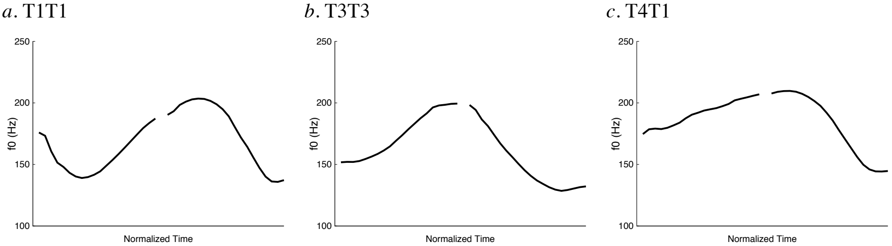

Figure 2.9 *f0 realization of three tone sandhi patterns in Tianjin Mandarin. Normalized time.*

It can be seen from Figure 2.9 that in all the three tonal combinations, the first tone was realized with a drastically different contour from its canonical form (compared to their respective contours in Figure 2.7). In T1T1 (Figure 2.9*a*), the first T1 is not a low falling tone any more as its canonical form. Instead, the offset of the tone was raised to a great extent, as in /tɕjɐ^1 mɑɔ^1/ *'dometic cat'*. In T3T3 (Figure 2.9*b*), the first T3 was realized with a rising contour, which is different from the dipping contour of T3 in isolation, as in /wu^3 ny^3/ *'dancing girl'*. In T4T1 (Figure 2.9*c*), the first tone shows a high rising *f0*, as in /xəʊ^4 mɐ^1/ *'stepmother'*. It is again very different from the high falling contour when the tone was pronounced in isolation. Among the three tone sandhi sequences, T3T3 is near-merged with its so-called sandhi output T2T3 (as in /t^{h}wo^2 njɑɔ^3/ *'ostrich'*), while T1T1 and T4T1 are not merged with any tonal sequence, as claimed as T3T1 (as in /naɛ^3 mɐ^1/ *'nanny'*) and T2T1 (as in /paɛ^2 mɑɔ^1/ *'white cat'*) in the literature.

When these bisyllabic constituents occur in trisyllabic sequences, tone sandhi does not consistently apply as reported in the literature (e.g., Li & Liu, 1985; Chen, 2000; Ma, 2005; Wee et al., 2005). Instead, only T3T3 was found to apply sandhi changes consistently and regardless of its alignment in the trisyllabic sequences, as illustrated with an example in Figure 2.10. In both T3T3T2 (as in /xwo^3 pɐ^3 tɕje^2/ *'torch festival'*) and T2T3T3 (as in /xwɐ^2 pjɑɔ^3 tɕjɑŋ^3/ *'Marble Pillar Award'*) sequences, the first T3 was realized with a rising *f0* contour as found in the disyllabic data (Figure 2.9*b*), indicating the application of tone sandhi in both contexts.

By contrast, the patterns T1T1 and T4T1 can only be applied when the patterns are right aligned in the trisyllabic sequences, as shown with an example for T1T1 in Figure 2.11. In T2T1T1 (Figure 2.11*b*) where the T1T1 sequence is right-aligned as in /xʊŋ^2 təŋ^1 tɕ^{h}y^1/ *'red light district'*, the middle T1 was realized with a rising *f0* contour comparable to that in Figure 2.9*a*, suggesting the application of tone sandhi in this case. When T1T1 is left-aligned (Figure 2.11*a*) as in /kʊŋ^1 ʃ̺ɑŋ^1 tɕy^2/ *'Trade and Industry Bureau'*, tone sandhi does not apply since the first T1 was realized with a falling contour, just like its canonical form as in Figure 2.7. For further details, see Li and Chen (2012a, 2016).

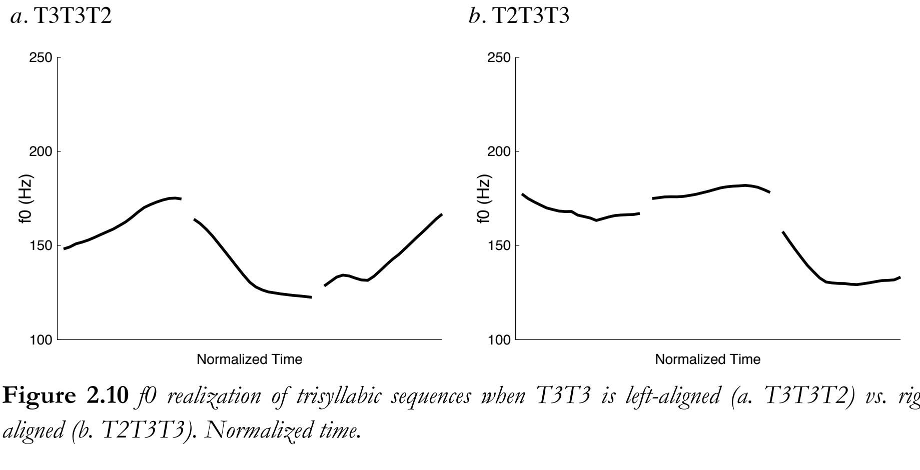

Figure 2.10 *f0 realization of trisyllabic sequences when T3T3 is left-aligned (a. T3T3T2) vs. right-aligned (b. T2T3T3). Normalized time.*

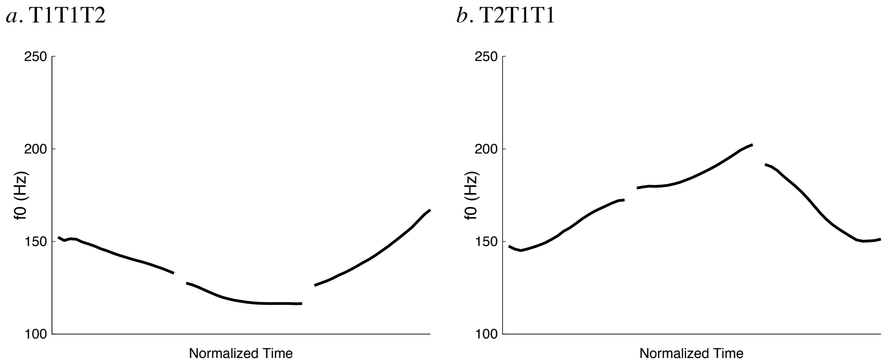

Figure 2.11 *f0 realization of trisyllabic sequences when T1T1 is left-aligned (a. T1T1T2) vs. right-aligned (b. T2T1T1). Normalized time.*

2.5.3 Neutral tone

As in Beijing Mandarin, neutral tone also exists in Tianjin Mandarin. The neutral tone syllables are those that do not surface with any of the lexical tones. As these syllables always occur in the prosodically weak positions like in Beijing Mandarin (Chen & Xu, 2006), they are usually produced with acoustic reduction in the segmental aspects, where the onset consonant of the neutral tone syllable is sometimes voiced, and the vowel might be centralized or even deleted. For example, in the word /kɤ^1 kɤ/ *'elder brother'*, in which the second syllable is a neutral tone syllable, its onset consonant is often voiced, and the vowel can be reduced to a schwa, as in [kɤ^1 gə]. Suprasegmentally, neutral tone syllables are usually produced with short duration (typically about half of the duration for a lexical tone syllable); and their *f0* realization also exhibits much variability (Wang, 2002).

Neutral tone syllables in Tianjin Mandarin never occur independently; instead, they always follow a lexical tone syllable in a disyllabic lexical item. Typical neutral tone syllables include grammatical morphemes (e.g., possessive marker /tə/ in A), the second syllable of a disyllabic lexical item (e.g., /li/ in B), and the final syllable of a reduplicated form (e.g., /kɤ/ in C).

A.	/wo^3 tə/	*'mine'*
B.	/pwo^1 li/	*'glass'*
C.	/kɤ^1 kɤ/	*'elder brother'*

Like in Beijing Mandarin, *f0* realization of neutral tone in Tianjin is much influenced by the preceding lexical tones, due to its prosodic weak position. A stable mid low tonal target can be observed, but only when given enough time as in Chen and Xu (2006).

Previous studies observe special rising neutral tone realization when it is followed by the low falling T1 (Wang, 2002; Li & Chen, 2011) for example in /peɪ^1 t̻ʃə mɐ^1/ *'carrying Mom on the back'*, as illustrated in Figure 2.12. Each tonal contour was obtained by averaging across 6 samples produced by the male speaker.

However, the rising neutral tone realization is due to the general raising effect of T1 upon its preceding tones (Li & Chen, 2012b; cf. Wang, 2002). When there are multiple neutral tone embedded as shown in Figure 2.13*a*, the neutral tone first aims to realize its own mid-low tonal target by the end of the second neutral tone, as in the example sentence /t^{h}ɐ^{1} ʃ̪wo^{1} mɐ^{1} mɐ mən tə mɑɔ^{1} t^{h}i^{1} lwan^{4} lə nɐ^{4} kə ɕjɛn^{4} tɕ^{h}jəʊ^{2}./ *'He said mothers' cats messed up that cotton ball.'* The raising could only be observed over the very last neutral-tone syllable. Importantly, the rising effect can be blocked by a major prosodic boundary as in Figure 2.13*b*, as in the example sentence /t^{h}ɐ^{1} ʃ̪wo^{1} mɐ^{1} mɐ mən tə tsəŋ^{1} tɕjɐ^{1} lə san^{1} paɛ^{3} k^{h}waɛ^{4} tɕ^{h}jɛn^{2}./ *'He said mothers' had increased by 300 yuan.'*

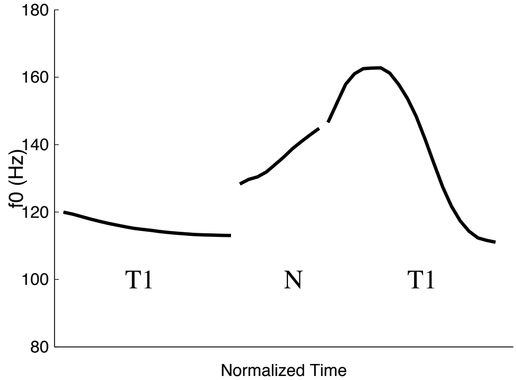

Figure 2.12 *f0 realization of one neutral tone (N) embedded between two T1s. Normalized Time.*

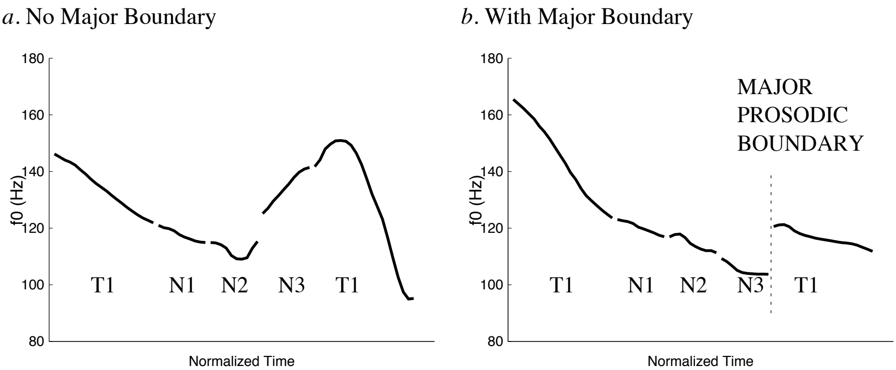

Figure 2.13 *f0 realization of three neutral tones (N1N2N3) embedded between two T1s without (a) or with (b) a major prosodic boundary following the neutral tone sequence. Normalized Time.*

2.6 Transcription of recorded passage 'north wind and the sun'

This passage is transcribed phonemically based on the symbols described in the consonant and vowel part. Tones are marked with numbers instead of tonal values. Tones undergoing sandhi changes are marked with the number of the original tone with brackets. Neutral tone syllables are not marked with tone numbers.

jəʊ^{3} ji$^{(4)}$ tʰjɛn^{1},   peɪ^{3} fəŋ^{1} xɤ^{2} tʰaɛ^{4} jɑŋ   tʃ̪ɹ̪əŋ^{4} tsaɛ^{4} ʃ̪wo^{1} ʃ̪eɪ^{2} tə pən^{3} ʃ̪ɹ̪ tɐ^{4}.
有一天， 北风和太阳 正在说 谁的本事大。

tʃ̪əŋ^{4} xɑɔ˞^{3} jəʊ^{3} kə   tʃ̪ʰwan^{1} təʊ^{3} pʰəŋ tə ɹ̪ən^{2} tsəʊ^{3} kwo^{4} laɛ.
正好 有个 穿斗篷的人 走过来。

tʰɐ^{1} ljɐ^{3} ʃ̪wo^{1},   ʃ̪eɪ^{2} nəŋ^{2} ɹ̪ɑŋ^{4} neɪ^{4} kə ɹ̪ən^{2} tʰwo^{1} tjɑɔ   təʊ^{3} pʰəŋ,
他俩说， 谁能让 那个人 脱掉 斗篷，

tɕjəʊ^{4} swan^{4} ʃ̪eɪ^{2} li^{4} xaɛ.
就算 谁厉害。

ɹ̪an^{2} xəʊ^{4},   peɪ^{3} fəŋ^{1} tɕjəʊ^{4} pʰin^{1} lə miŋ^{4} tə tʃ̪ʰweɪ^{1}.
然后， 北风 就 拼了命地吹。

tan^{4} ʃ̪ɹ̪   tʰɐ^{1} tʃ̪ʰweɪ^{1} tə   ɥe^{4} li^{4} xaɛ,
但是 他吹得 越厉害，

nɐ^{4} kə ɹ̪ən^{2} tɕjəʊ^{4} pɐ$^{(3)}$ təʊ^{3} pʰəŋ   kwo^{3} tə ɥe^{4} tɕin^{3}.
那个人 就把斗篷 裹得越紧。

peɪ^{3} fəŋ^{1} meɪ^{2} tʃ̪ɤ^{2},   tʃ̪ɹ̪$^{(3)}$ xɑɔ^{3} fɑŋ^{4} tɕʰi^{4}.
北风 没辙， 只好 放弃。

tɕje^{1} tʃ̪ə,   tʰaɛ^{4} jɑŋ   ʃ̪aɛ^{4} tə   ɹ̪ɤ^{4} xʊŋ$^{(1)}$ xʊŋ^{1} tə,
接着， 太阳 晒得 热烘烘的，

neɪ^{4} kə ɹ̪ən^{2} li^{4} mɐ˞$^{(3)}$ pɐ^{3} təʊ^{3} pʰəŋ   tʰwo^{1} tjɑɔ^{4} lə.
那个人 立马把 斗篷 脱掉了。

ɹ̪an^{2} xəʊ^{4},   peɪ^{3} fəŋ^{1} tʃ̪ɹ̪^{3} nəŋ^{2} ɹ̪ən$^{(4)}$ ʃ̪u^{1}.
然后， 北风 只能 认输。

CHAPTER 3 AN ACOUSTIC STUDY OF CONTEXTUAL TONAL VARIATION IN TIANJIN MANDARIN[2]

3.1 Introduction

It is known that the *f0* realization of lexical tones varies extensively in connected speech. At a global level, *f0* can vary due to the overall discourse prosody in which the tone-bearing lexical item is produced. At a more local level, the *f0* realization of a lexical tone can be further affected by its neighboring tones (see e.g., Chen, 2012 for a review). One well-noted local contextual effect is tone sandhi, which is often described as the phonological change of lexical tones (Chen, 2000; also see Zhang, 2010 for a review). Another local contextual effect is tonal coarticulation, which is traditionally defined as phonetically minor *f0* adjustments to preceding or following tones (see Xu, 2001 for a review). Both processes cause changes in the *f0* realization of a lexical tone, so that the resulting *f0* contour may deviate to varying degrees from the canonical *f0* shape of the tone produced in isolation. While there has been general agreement on the difference between tonal coarticulation and tone sandhi to account for contextual tonal variation, there is little consensus or research effort on how exactly or whether the wide range of *f0* deviations can be categorized as a function of these two distinct processes.

In this study, we aimed to address this issue through directly comparing these two types of tonal variation. Our empirical base is Tianjin Mandarin, a Northern Mandarin dialect spoken in the urban areas of Tianjin, a metropolis in northern China. Tianjin Mandarin presents an interesting test case on tonal variation, further understanding of which will shed light on how different tonal variation processes intertwine to determine the *f0* realization of lexical tones. The dialect is known for its complex tone sandhi patterns over disyllabic constituents. Over trisyllabic constituents, some tone sandhi rules have been argued to apply only to the leftmost two syllables, while others only to the rightmost two syllables, regardless of the morpho-syntactic structures. Furthermore, some sandhi rules are claimed to apply iteratively, i.e., over derived contexts of the first round of sandhi application. These two properties are together known as the "paradox" in Tianjin tone sandhi (e.g., Chen, 2000).

It is important to note that although tonal variation in Tianjin Mandarin has been extensively studied, most were based on impressionistic data. In a recent acoustic study, Zhang and Liu (2011) examined tonal coarticulation of disyllabic items but did not compare them to the previously claimed tone sandhi patterns. This makes it impossible to evaluate directly how tonal variation due to different sources may differ. Furthermore, despite the potentially significant implications of the Tianjin Mandarin tone sandhi "paradox" on our theoretical modelling of tonal variation, no experimental work has been done on trisyllabic tone sandhi in Tianjin Mandarin. Pilot results in Li and Chen (2012)

[2] A version of this chapter has been published as: Qian Li & Yiya Chen. (2016). An Acoustic Study of Contextual Tonal Variation in Tianjin Mandarin. *Journal of Phonetics 54*: 123-150.

raised serious doubts on the tonal variation patterns of trisyllabic sequences reported in the literature (Chen, 2000 and references therein).

The specific goal of this study therefore was to collect data from a well-controlled acoustic experiment and investigate 1) how lexical tones are realized with varying *f0* contours in disyllabic domains without pre-exclusion of any tonal sequence claimed to involve tone sandhi processes; and 2) whether disyllabic tone sandhi alternations can be observed in trisyllabic constituents, with particular focus upon the issue of directionality and iterativity of sandhi application. By doing so, we hope to arrive at a better understanding of tonal realization in Tianjin Mandarin, which, in turn, would help to further refine a general theory of tonal variation.

3.1.1 Lexical tones of Tianjin Mandarin in isolation

Although Tianjin Mandarin has a very similar segmental structure to Standard Chinese, the dialect differs significantly from Standard Chinese in terms of the tonal system (Han, 1993; Wee et al., 2005). There are four lexical tones in Tianjin Mandarin. Figure 3.1 illustrates the *f0* realization of the four lexical tones elicited in isolation. All *f0* contours were plotted based on 50 samples with different segmental structures for each tone, produced by a male native speaker in his 20s at the time of recording (born in 1983). *f0* contours were time-normalized by taking 20 points in the rhyme part of each syllable with equal time interval. The *f0* values were normalized by converting them to speaker-specific z-scores (Rose, 1987) with the formula: $z = \frac{f0_x - f0_{mean}}{f0_{SD}}$. ($f0_x$: the observed *f0* value in Hz; $f0_{mean}$: the mean *f0* value of the speaker in Hz; $f0_{SD}$: the standard deviation of *f0* value of the speaker in Hz.)

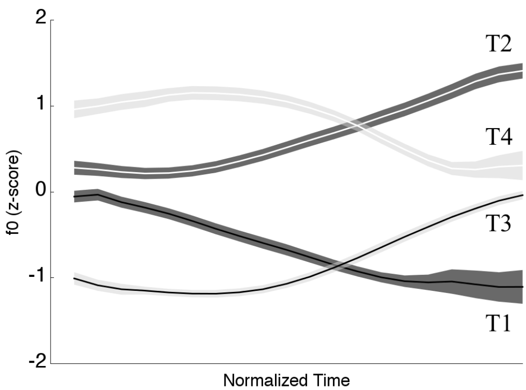

Figure 3.1 *Four lexical tones in Tianjin Mandarin produced in isolation (with normalized time). Lines stand for the mean. Gray areas stand for* ±*1 standard error of the mean.*

As can be seen from Figure 3.1, Tone 1 is a low-register falling tone, with a pitch contour that falls from the middle to the lower end of the speaker's pitch range. Tone 2 is a high-register rising tone, whose pitch contour rises from the middle to the upper end of the pitch range. Tone 3 is a low-register dipping tone, which falls slightly from the lower pitch range, stays at the bottom and then rises to the mid pitch range of the speaker. Tone 4 is a high-register falling tone which falls from the upper end to the mid pitch range.

Table 3.1 summarizes the annotations adopted in different studies. Strictly speaking, no annotation system listed in Table 3.1 accurately reflects the *f0* contours plotted in Figure 3.1. This discrepancy is in part an indication of the considerable variation in lexical tone production that exists both within and across speakers of the same generation as well as across speakers of different generations. Despite the variability, the basic *f0* patterns of the four lexical tones are rather consistent at a more abstract level.

Table 3.1 *Transcriptions of the four lexical tones in Tianjin Mandarin in different studies.*

T1	T2	T3	T4	
L	H	R(LH)	F(HL)	e.g., Chen, 2000; Wee et al., 2005; Hyman, 2007
LL	HH	LH	HL	e.g., Wang & Jiang, 1997; J. Wang, 2002; Ma, 2005
21	45	213	53	e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Han, 1993
211	455	113	553	e.g., Shi, 1986
41	34	12	52	e.g., Zhang & Liu, 2011

3.1.2 Tonal variation in connected speech in Tianjin Mandarin

In connected speech, tonal variation abounds in Tianjin Mandarin. Of our interest are the two processes that have been described in the literature to extensively influence the *f0* contour of a lexical tone within a local tonal context: tonal coarticulation and tone sandhi.

3.1.2.1 Tonal coarticulation

Tonal coarticulation has been reported in many East Asian contour-tone languages, such as Thai (e.g., Palmer, 1969; Abramson, 1979; Gandour et al., 1993; Gandour et al., 1994; Gandour et al., 1996; Potisuk et al., 1997; Zsiga & Nitisaroj, 2007), Vietnamese (e.g., Han & Kim, 1974; Brunelle, 2009), Standard Chinese (e.g., Shen, 1990; Xu, 1994; Xu, 1997; Kochanski et al., 2003), Taiwanese (e.g., Peng, 1997; Wang, 2002), and Malaysian Hokkien (e.g., Chang & Hsieh, 2012). Despite the differences among languages, two types of tonal coarticulation have been commonly attested: the left-to-right carryover tonal coarticulation and the right-to-left anticipatory tonal coarticulation.

Carryover tonal coarticulation refers to the effect that the production of the first tone has on its following tone in a disyllabic tonal sequence. Carryover effects are generally assimilatory in nature (e.g., Han & Kim, 1974; Abramson, 1979; Shen, 1990; Gandour et al., 1994; Potisuk et al., 1997; Xu, 1997; Brunelle, 2009). Therefore, the high *f0* offset of the first tone raises the *f0* onset of the second tone, while the low *f0* offset lowers the following onset. Such tonal coarticulation patterns have been observed consistently across languages regardless of the identity of the first tone. Due to the carryover effects, the onset of the second tone usually varies greatly. The influence from the preceding tone can be extended well into the first half or the entire second tone, although the magnitude of the influence decreases over time (e.g., Gandour et al., 1993; Gandour et al., 1994; Potisuk et al., 1997; Xu, 1997; Brunelle, 2009).

The cross-linguistic carryover effects have been reported for Tianjin Mandarin in Zhang and Liu (2011). Specifically, they argued that three tones in Tianjin Mandarin are subject to the carryover effects, i.e., T2, T3, and T4. The low-falling tone (T1), however, was claimed to resist the carryover effects from any preceding tone.

In contrast to carryover effects, anticipatory tonal coarticulation refers to the effect of a following lexical tone on the production of its preceding tone. Anticipatory effects are mainly dissimilatory in nature (but see Shih & Kochanski, 2000 for reported assimilatory anticipatory effects in Standard Chinese), and are usually restricted to a limited number of triggering contexts. The magnitude and the temporal extent of their influence are also much less (e.g., Gandour et al., 1994; Potisuk et al., 1997; Xu, 1997; c.f., Chang & Hsieh, 2012 for more powerful anticipatory effects reported for Malaysian Hokkien). In a disyllabic tonal sequence, the anticipatory effects seem to be triggered only by a following low tone (e.g., Gandour et al., 1994 and Xu, 1997 for Thai and Mandarin, respectively). In most cases, anticipatory effects on the first tone are also rather negligible and localized to the offset of the preceding tone (e.g., Potisuk et al., 1997; Xu, 1997).

In Tianjin Mandarin, only a minor anticipatory effect has been reported, which is triggered by the low-register dipping T3 (Zhang & Liu, 2011). It is important to point out that Zhang and Liu (2011) did not examine tonal combinations that were assumed to involve tone sandhi for potential anticipatory effects. This makes our understanding of the anticipatory tonal coarticulation effects in Tianjin Mandarin possibly limited and incomplete. Given that in Tianjin Mandarin, tone sandhi is right dominant and anticipatory coarticulation is also triggered by the second syllable within a disyllabic word, we think it is particularly important that we include all possible tonal combinations to shed further light on the potential differences between tonal coarticulation and tone sandhi.

3.1.2.2 Tone sandhi

Different from tonal coarticulation that is usually considered the phonetic modification of *f0* contours, tone sandhi, in many of the impressionistic reports, is considered to involve the categorical change of the phonological identity of a lexical tone. Much work has been done on tone sandhi in various Chinese dialects (see e.g., Chen, 2000 for an overview), among which Tianjin Mandarin is known for its complex sandhi patterns over disyllabic and trisyllabic constituents.

For disyllabic sequences, most previous studies on Tianjin Mandarin have identified four disyllabic tone sandhi rules (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma & Jia, 2006; Hyman, 2007), as shown in Rules (1) – (4), where the tones were transcribed with features adopted from Chen (2000).

(1) T1+T1 (L+L) → T3+T1 (LH+L)
(2) T3+T3 (LH+LH) → T2+T3 (H+LH)
(3) T4+T4 (HL+HL) → T1+T4 (L+HL)
(4) T4+T1 (HL+L) → T2+T1 (H+L)

Rules (1)-(3) have been attributed to tonal dissimilation due to the Obligatory Contour Principle, while Rule (4) is attributed to Tonal Absorption (Chen, 2000). With sociolinguistic data, Lu (1997) proposes that Rule (1) has undergone some diachronic changes. A new pattern of that sandhi change, i.e., T1+T1 (L+L) → T2+T1 (H+L), has been observed in the speech of the younger generation. In addition, Rule (4) has been reported to become "obsolete" in the speech of the younger generation (Liu & Gao, 2003). More recently, Wee et al. (2005) argued that there are two more sandhi rules in Tianjin Mandarin, where T3 changes to T1 before either T2 or T4 as shown in Rules (5) and (6).

(5) T3+T2 (LH+H) → T1+T2 (L+H)
(6) T3+T4 (LH+HL) → T1+T4 (L+HL)

What is shared among all the impressionistic studies is that tone sandhi involves the categorical change of one lexical tone to another within the native lexical tonal system. However, recent experimental data suggest that in Tianjin Mandarin, tonal neutralization might not be involved at all in disyllabic tone sandhi (Zhang & Liu, 2011; Li & Chen, 2012).

For trisyllabic sequences, the tone sandhi situation becomes more complicated. Of all 64 tonal combinations of trisyllabic sequences, 27 have been claimed to involve sandhi based on the four commonly accepted disyllabic sandhi rules (e.g., Li & Liu, 1985; Chen, 2000; Ma, 2005; but see Wee et al., 2005 for more trisyllabic sequences that undergo sandhi changes).

Among the 27 combinations, 20 have been referred to as simple applications, where only one disyllabic sandhi rule could possibly apply. These disyllabic patterns have been argued to apply consistently within a trisyllabic domain, regardless of the location where the disyllabic combinations occur. For example, T1T1 sandhi is claimed to apply over both T1T1T2 and T2T1T1.

The other seven cases (7)-(13) have been claimed to involve iterativity and directionality issues of sandhi applications (e.g., Li & Liu, 1985; Chen, 2000; Ma, 2005). Specifically, (7)-(11) have been claimed to involve iterative sandhi rule application, due to a newly derived sandhi context after the first round of sandhi application. For example, in the sequence T3T1T1 (10), the first T1 is claimed to change to T3, which results in an intermediate stage T3T3T1 where T3T3 is again a possible sandhi context, leading to the final output as T2T3T1.

The seven cases here have also been argued to exemplify the directionality of tone sandhi changes. (7) and (8) have been argued to undergo sandhi processes from the left edge and (9)-(13) from the right edge. Taking the tonal sequence T1T1T1 (12) as an example. This sequence has been reported to surface as T1T3T1, which has been taken as evidence that the sandhi rule must have been applied from the right edge of the constituent instead of the left edge.

	Underlying Pattern		1st Round of Application		2nd Round of Application
(7)	T3+T3+T3	→	T2+T3+T3	→	T2+T2+T3
	(LH+LH+LH)		(H+LH+LH)		(H+H+LH)
(8)	T4+T4+T1	→	T1+T4+T1	→	T1+T2+T1
	(HL+HL+L)		(L+HL+L)		(L+H+L)
(9)	T4+T4+T4	→	T4+T1+T4	→	T2+T1+T4
	(HL+HL+HL)		(HL+L+HL)		H+L+HL)
(10)	T3+T1+T1	→	T3+T3+T1	→	T2+T3+T1
	(LH+L+L)		(LH+LH+L)		(H+LH+L)
(11)	T1+T4+T4	→	T1+T1+T4	→	T3+T1+T4
	(L+HL+HL)		(L+L+HL)		(LH+L+HL)
(12)	T1+T1+T1	→	T1+T3+T1		
	(L+L+L)		(L+LH+L)		
(13)	T4+T1+T1	→	T4+T3+T1		
	(HL+L+L)		(HL+LH+L)		

Various attempts have been made to explain the complex sandhi application in Tianjin Mandarin (e.g., Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; Ma, 2005; Wee et al., 2005), but no consensus has been reached. The issues of iterative and directional applications of tone sandhi in Tianjin Mandarin not only have presented great challenges to phonological theories of tonal alternation, but also have raised interesting questions regarding theories of allotone processing and learnability.

Before any attempt to address these theoretical issues is made, two problems need to be tackled with regard to the proposed trisyllabic sandhi rules. First, since the disyllabic combinations have been shown to be realized differently from what has been claimed in the literature, it is very likely that the actual realization of trisyllabic combinations might also differ from those described in previous studies. Second, the so-called iterative sandhi application proposed for some trisyllabic combinations is based on the assumption that sandhi-derived tones are lexical tones within the system and there is complete neutralization of the derived and the base tones. If indeed tone sandhi in Tianjin Mandarin does not involve lexical tone neutralization, any iterative tonal change is then called into question. At this stage, we believe that well-controlled experimental data are not only needed but also crucial for our further understanding of these issues.

3.2 Method

Disyllabic tonal sequences with full combination of the four lexical tones ($T_{1\text{-}4}T_{1\text{-}4}$) were collected to allow direct comparison of the two possible types of tonal variation, i.e., tonal coarticulation and tone sandhi in the language. To investigate further the iterativity and directionality of tone sandhi application, relevant trisyllabic sequences were also included.

Subsets of the data will also provide us with evidence to shed light on the issue of tonal neutralization under sandhi changes.

3.2.1 Materials

We used real words and phrases because most of our speakers reported to have difficulties in producing nonsense sequences (see also Zhang & Liu, 2010 for different tonal realization of nonsense words compared to real words). Two reading lists were included in this study, one for the disyllabic sequences and the other for the trisyllabic sequences.

The disyllabic reading list consisted of all 16 possible tonal combinations of the four lexical tones in Tianjin Mandarin. All stimuli are common collocations, which are arguably lexicalized in the language. For possible influence from their internal morpho-syntactic structures, we limited the structures to three different combinations: Compound (e.g., 爹妈 die^{1}ma^{1} "*parent*"), Modifier+Head (e.g., 鸵鸟 tuo^{2}niao^{3} "*ostrich*"), and Verb+Object (e.g., 推理 tui^{1}li^{3} "*inference*"). These morpho-syntactic structures had a balanced number for each tonal combination. The choice of the stimuli was further constrained by their segmental composition. Only sonorant onsets were used for the second syllables in order to better observe the effect of contextual tonal variation and also for the ease of segmentation.

The trisyllabic reading list included all 64 combinations of the four lexical tones. For each trisyllabic tonal combination, the sequence was either a mono-morphemic word (σσσ, as mostly seen in loan words such as 乌拉圭 wu^{1}la^{1}gui^{1} "*Uraguay*"; hereafter, underlines indicate the constituent structure of the sequence), or a compound or phrase with two different internal grouping patterns: σσ+σ (e.g., 拉丁舞 la^{1}ding^{1} wu^{3} "*Latin dance*") or σ+σσ (e.g., 高工资 gao^{1} gong^{1}zi^{1} "*high salary*"). For each structure, we included two different items. Due to the limited availability of syllables with sonorant onsets, items with obstruent onsets at the second and third syllable were also used. Most test items are very common words and phrases with an exception of a few, which nevertheless are still relatively common.

Each test item was embedded in a unique carrier sentence with the copular verb "… shi^{4} … ('… *is* …')." The length of each carrier sentence was within the range of 9-11 syllables as illustrated in the following example, where the target stimulus is 密码 (mi^{4} ma^{3}, "*password*").

密码	是	种	保密的	手段。
Mi^{4} ma^{3}	shi^{4}	zhong^{3}	bao^{3} mi^{4} de	shou^{3} duan^{4}.
Password	BE	Classifier	encrypting	method.

Password is an encrypting method .

Data from Chen and Gussenhoven (2008) and Chen (2010) for Standard Chinese as well as Scholz (2012) and Scholz and Chen (2014) for Wenzhou Chinese show that lexical tones are less influenced by adjacent tonal targets when produced under focus, compared to their post-focus counterparts. (Readers are also referred to Chen, 2012 for further discussions on the effects of on-focus vs. post-focus on tonal realization.) To control the effect of focus, we elicited the utterances as answers to specific questions, resulting in two focus conditions, as illustrated below. In the On-Focus condition, focus was on the target phrase. In the Pre-Focus condition, focus was on later parts of the sentence. In this way, we excluded the possibility that speakers might randomly produce the stimulus sentences with varying information structures. For this particular study, we were not interested in the effect of focus per se and had therefore placed the target stimuli in the sentence-initial subject position, where we knew that the effects of on-focus and pre-focus are closer than in other positions (Xu, 1999).

On-Focus Condition:

QUESTION:

什么	是	种	保密的	手段？
Shen^2 me	shi^4	zhong^3	bao^3 mi^4 de	shou^3 duan^4?
What	BE	Classifier	encrypting	method?

What *is an encrypting method?*

ANSWER:

密码	是	种	保密的	手段。
Mi^4 ma^3	shi^4	zhong^3	bao^3 mi^4 de	shou^3 duan^4.
Password	BE	Classifier	encrypting	method.

Password *is an encrypting method.*

Pre-Focus Condition:

QUESTION:

密码	是	种	**什么？**
Mi^4 ma^3	shi^4	zhong^3	**shen^2 me**?
Password	BE	Classifier	**what**?

What *is password?*

ANSWER

密码	是	种	**保密的**	**手段。**
Mi^4 ma^3	shi^4	zhong^3	**bao^3 mi^4 de**	**shou^3 duan^4**.
Password	BE	Classifier	**encrypting**	**method**.

Password is an ***encrypting method****.*

3.2.2 Subjects

Six speakers (3 males and 3 females; Mean=25) participated in the experiment. All were local college students, born in the 1980s and raised in the urban areas of Tianjin. None of them had lived outside of Tianjin. They were paid for the participation but unaware of the purpose of the experiment. All participants provided written informed consent.

3.2.3 Recording

All questions were recorded beforehand in the Phonetics Lab at Leiden University by a native speaker of Tianjin Mandarin. During the experiment, subjects were requested to respond to the question with the targeted response sentence presented on the computer screen. All found the task straightforward and followed the same procedure. Recordings were conducted in a quiet room in Tianjin using an M-Audio® mobile digital audio recorder MicroTrack II with 44.1 kHz sampling rate and 16 bit rate in mono channel.

96 disyllabic items (4 initial tones * 4 final tones * 3 structures * 2 focus conditions) were elicited from each of the six participants, each with two repetitions. 768 trisyllabic items (4 initial tones * 4 middle tones * 4 final tones * 3 structures * 2 words * 2 focus conditions) were also elicited from each participant. However, one third of the items, with fricative onsets at the second or third syllable, were produced only once. The other two thirds, with non-fricative onsets, were produced twice.

3.2.4 *f0* 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 of each syllable or in the whole syllable for the second syllable within disyllabic constituents where the syllable onsets were all sonorant. To eliminate the pitch range difference due to gender and to better illustrate the tonal variation patterns, each speaker's raw *f0* data were normalized by converting them to speaker-specific z-scores (Rose, 1987). The plotted tonal contours were then based on the mean z-scores averaged across speakers.

3.2.5 Growth curve analysis

To better understand the varying nature of *f0* over the time-course of the tone-bearing syllable, we employed the growth curve analysis (Mirman, 2014) with the package *lme4* (Bates, Maechler, Bolker & Walker, 2014) in R (R Core Team, 2014). The growth curve analysis is a multilevel regression technique which is more appropriate and powerful for analyzing non-linear time-varying data (e.g., *f0*). The advantage of this technique is that the overall *f0* contours are taken into consideration. Moreover, the experimental manipulations are better examined by taking into consideration of the random variation due to different subjects.

In growth curve analysis, a non-linear curve is typically fitted with high order orthogonal polynomials, such as, $y = a + bx + cx^2 + dx^3 + ex^4 \dots$ (x for time, y for *f0*). The characteristics of a curve could be independently reflected by the coefficients for different time terms in the polynomials. Among them, the intercept *a* usually indicates the overall mean of a curve; the linear term *b* indicates the general direction of the curve change (slope); the higher order terms generally reflect the curve complexity (e.g., the quadratic term *c* for a U-shape curve, cubic *d* for an S-shape curve, and quartic term *e* for a W-shape curve).

In this study, we used up to second-order polynomials since the most complex *f0* contour of one lexical tone in our data has only a convex or concave contour shape. Tonal variation due to different contextual conditions was therefore analyzed by assessing the intercept, linear, and quadratic terms in the fitted model. If two contours differ from each other, we should expect statistical significance in at least one aspect of their *f0* contours. This includes the overall *f0* mean indexed by the intercept, the direction of *f0* change such as rising vs. falling indexed by the slope, and the steepness of *f0* rising or falling indexed by the quadratic term.

To carry out the growth curve analyses, separate linear mixed-effects models were built for subsets of the data for specific hypothesis testing. For both disyllabic and trisyllabic datasets, we have controlled the tone category, consonant type, vowel height and syllable structure for each syllable within one stimulus. We also controlled the structure (i.e., morpho-syntactic structure for disyllables and grouping structure for trisyllables), focus, repetition and relative familiarity of each stimulus. To determine which controlled factor should be considered in the models, each base model was first established containing only the time terms in the fixed factor structure and the random slope of subject on all time terms. Additional controlled factors were then added in the models in a stepwise fashion. Model fits were tested at each step via model comparisons based on the change in the log-likelihood ratios. Controlled factors that significantly improved the model fits were kept in the fixed factor structure of the models. Parameter-specific *p*-values were estimated using the normal approximation (i.e., treating the *t*-value as a *z*-value). Multiple comparisons with Bonferroni adjustment were conducted with the function *glht* in package *multcomp* (Hothorn, Bretz & Westfall, 2008) whenever necessary.

3.3 Results & discussion I: Tonal variation in disyllabic domain

3.3.1 Results

We were interested in the potential differences in tonal variation due to coarticulation and tone sandhi. Given that tone sandhi of disyllabic sequences in Tianjin Mandarin involves the modification of the first tone as a function of its following tone, we zoomed into the *f0* contours of each first tone followed by different second tones (i.e., the anticipatory effects), as plotted in Figure 3.2. For our planned comparisons, the whole disyllabic dataset was separated into four subsets according to the tonal category of the first tone. The statistical results were summarized in Table 3.2.

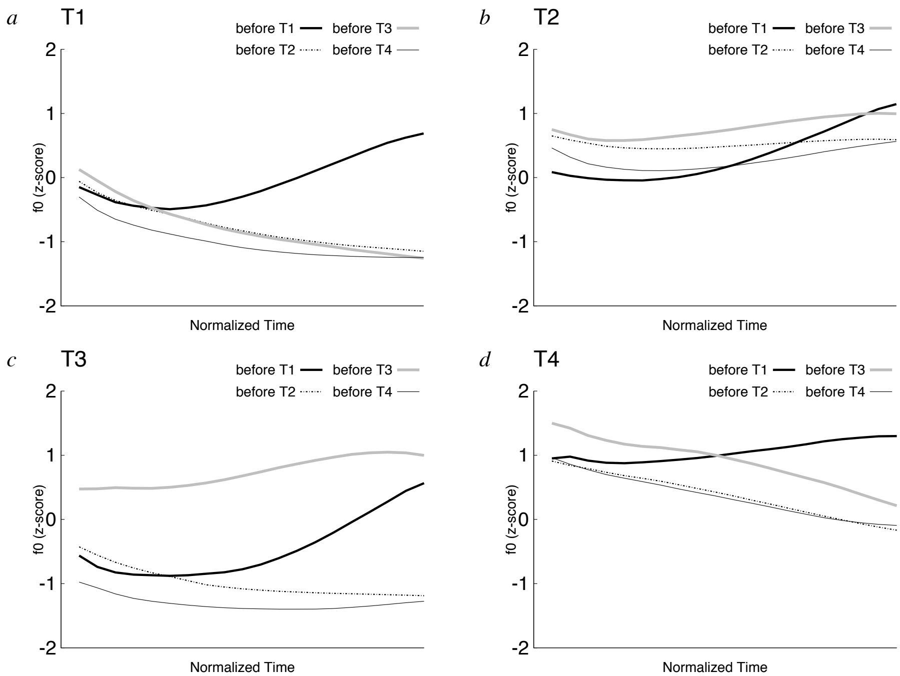

Figure 3.2 *The f0 variation of the four target lexical tones due to different following tones. T1 in a, T2 in b, T3 in c and T4 in d. Lines indicate the mean of the normalized f0 contour across speakers. Normalized time.*

Table 3.2 *Results of pairwise comparison on the variability in the first tone due to different following tones.*

***a*. Target T1**

beforeT2 vs. beforeT1		beforeT3 vs. beforeT1		beforeT4 vs. beforeT1	
Intercept	$\beta=-0.67$ $z=-6.69$ $p<0.001$	*Intercept*	$\beta=-0.65$ $z=-5.88$ $p<0.001$	*Intercept*	$\beta=-0.89$ $z=-8.86$ $p<0.001$
Slope	$\beta=-2.86$ $z=-11.56$ $p<0.001$	*Slope*	$\beta=-3.18$ $z=-12.83$ $p<0.001$	*Slope*	$\beta=-2.55$ $z=-10.37$ $p<0.001$
Quadratic	$\beta=-0.43$ $z=-4.13$ $p<0.001$	*Quadratic*	n.s.	*Quadratic*	$\beta=-0.33$ $z=-3.24$ $p<0.05$
		beforeT3 vs. beforeT2		beforeT4 vs. beforeT2	
		Intercept	n.s.	*Intercept*	n.s.
		Slope	n.s.	*Slope*	n.s.
		Quadratic	n.s.	*Quadratic*	n.s.

				beforeT4 vs. beforeT3	
				Intercept	n.s.
				Slope	n.s.
				Quadratic	n.s.

***b*. Target T2**

beforeT2 vs. beforeT1		beforeT3 vs. beforeT1		beforeT4 vs. beforeT1	
Intercept	n.s.	*Intercept*	n.s.	*Intercept*	β=-0.43 z=-3.77 p<0.01
Slope	β=-1.58 z=-8.17 p<0.001	*Slope*	β=-0.98 z=-5.04 p<0.001	*Slope*	β=-1.19 z=-6.12 p<0.001
Quadratic	β=-0.45 z=-6.32 p<0.001	*Quadratic*	β=-0.48 z=-6.42 p<0.001	*Quadratic*	β=-0.24 z=-3.28 p<0.05
		beforeT3 vs. beforeT2		beforeT4 vs. beforeT2	
		Intercept	β=0.32 z=3.29 p<0.05	*Intercept*	n.s.
		Slope	β=0.60 z=3.06 p<0.05	*Slope*	n.s.
		Quadratic	n.s.	*Quadratic*	n.s.
				beforeT4 vs. beforeT3	
				Intercept	β=-0.53 z=-5.38 p<0.001
				Slope	n.s.
				Quadratic	β=0.24 z=3.22 p<0.05

***c*. Target T3**

beforeT2 vs. beforeT1		beforeT3 vs. beforeT1		beforeT4 vs. beforeT1	
Intercept	β=-1.04 z=-6.66 p<0.001	*Intercept*	n.s.	*Intercept*	β=-0.95 z=-6.25 p<0.001
Slope	β=-2.54 z=-8.71 p<0.001	*Slope*	β=-0.88 z=-3.04 p<0.05	*Slope*	β=-2.16 z=-7.47 p<0.001
Quadratic	β=-0.66 z=-5.37 p<0.001	*Quadratic*	β=-0.95 z=-7.82 p<0.001	*Quadratic*	β=-0.64 z=-5.37 p<0.001

		beforeT3 vs. beforeT2		beforeT4 vs. beforeT2	
		Intercept	β=1.47 z=9.29 p<0.001	*Intercept*	n.s.
		Slope	β=1.66 z=5.69 p<0.001	*Slope*	n.s.
		Quadratic	n.s.	*Quadratic*	n.s.
				beforeT4 vs. beforeT3	
				Intercept	β=-1.39 z=-8.84 p<0.001
				Slope	β=-1.28 z=-4.42 p<0.001
				Quadratic	n.s.

***d*. Target T4**

beforeT2 vs. beforeT1		beforeT3 vs. beforeT1		beforeT4 vs. beforeT1	
Intercept	β=-0.53 z=-19.64 p<0.001	*Intercept*	n.s.	*Intercept*	β=-0.48 z=-16.03 p<0.001
Slope	β=-2.02 z=-21.61 p<0.001	*Slope*	β=-2.18 z=-23.99 p<0.001	*Slope*	β=-2.03 z=-22.25 p<0.001
Quadratic	n.s.	*Quadratic*	β=-0.40 z=-4.31 p<0.001	*Quadratic*	n.s.
		beforeT3 vs. beforeT2		beforeT4 vs. beforeT2	
		Intercept	β=0.52 z=19.84 p<0.001	*Intercept*	n.s.
		Slope	n.s.	*Slope*	n.s.
		Quadratic	n.s.	*Quadratic*	n.s.
				beforeT4 vs. beforeT3	
				Intercept	β=-0.46 z=-18.07 p<0.001
				Slope	n.s.
				Quadratic	n.s.

3.3.1.1 T1Tx

As shown in Figure 3.2*a*, the first T1 remains as a low-falling tone except when it is followed by another T1, in which case, it surfaces with a rising *f0* contour. This difference was statistically significant, as supported by results in Table 3.2*a*, where the *f0* contour of T1 before another T1 showed a significant difference from that before the other tones. (See significant statistical results of target T1 beforeT2 vs. beforeT1, beforeT3 vs. beforeT1, and beforeT4 vs. beforeT1.) For T1 before T2, T3, and T4, however, we did not see any significant difference. (See the lack of significant results of target T1 beforeT3 vs. beforeT2, beforeT4 vs. beforeT2, and beforeT4 vs. beforeT3.)

3.3.1.2 T2Tx

Figure 3.2*b* shows that, the first T2 remains a high-rising tone regardless of the identity of the following tone. There are, however, some subtle differences restricted to two contexts, i.e., before T1 and T3.

Before T1, there was a slight yet consistent *f0* rise towards the end of the tone-bearing syllable compared to that before the other three tones, confirmed by the significant statistical results in the slope and quadratic components in Table 3.2*b*. (See target T2 beforeT2 vs. beforeT1, beforeT3 vs. beforeT1, beforeT4 vs. beforeT1.) The raising effect before T1 on T2 was somehow different from the raising effect of T3 on T2. When followed by a T3, there was an overall raising effect of the whole *f0* mean contour, which was supported by the significant results in the intercept. (See target T2 beforeT3 vs. beforeT2 and beforeT4 vs. beforeT3.) In addition, there were some slight and less consistent contour differences when T2 was followed by T3, compared to that followed by T2 and T4 (reflected in the significant results of slope in target T2 beforeT3 vs. beforeT2 and quadratic component in target T2 beforeT4 vs. beforeT3). There was no significant difference in the *f0* contours of T2 beforeT2 vs. beforeT4.

3.3.1.3 T3Tx

Figure 3.2*c* shows that the target T3 varies greatly as a function of the following tones, which again were restricted mainly to two contexts, i.e., before T1 and T3. Before T1, there was a consistently greater *f0* rise compared to that before the other three tones, as confirmed by the significant results in the slope and quadratic components in Table 3.2*c*. (See target T3 beforeT2 vs. beforeT1, beforeT3 vs. beforeT1, beforeT4 vs. beforeT1.) This is different from the general T3 raising effect before another T3, which was mainly confirmed by the significant intercept differences. (See target T3 beforeT3 vs. beforeT2 and beforeT4 vs. beforeT3.) This T3 raising effect is similar to what we have observed for the raising effect of T3 on the preceding target T2, but with a much greater magnitude. In addition, probably due to the slightly falling *f0* before T2 and T4, we also observed a significant difference in the slope. (See again target T3 beforeT3 vs. beforeT2, beforeT4 vs. beforeT3.)

3.3.1.4 T4Tx

Figure 3.2*d* shows that the variation of T4 was also restricted to the two following tonal contexts, i.e., T1 and T3, as observed for targets T2 and T3 earlier. T1 again, introduced a significantly greater *f0* rise on its preceding T4 in terms of significant slope differences shown in Table 3.2*d*, in addition to the less consistent intercept and quadratic differences. (See target T4 beforeT2 vs. beforeT1, beforeT3 vs. beforeT1, beforeT4 vs. beforeT1.) Before T3, T4 showed a more general overall *f0* raising effect, as indicated by the significantly different intercepts. (See target T4 beforeT3 vs. beforeT2 and beforeT4 vs. beforeT3.) There was no significant difference in the *f0* contours of T4 before T2 vs. T4.

3.3.2. Discussion

Within disyllabic tonal sequences, we have observed a range of variation in the *f0* realization of four lexical tones due to different following tones. The results were summarized in Table 3.3, where within a disyllabic tonal sequence, the first tone was affected in terms of both the overall *f0* mean as well as the *f0* contour shape.

Table 3.3 *Summary of variability in the first tone due to different following tones.*

1st Tone	2nd Tone			
	T1 Low Falling	T2 High Rising	T3 Low Dipping	T4 High Falling
T1 Low Falling	Greater *f0* Rise	No Effect	No Effect	No Effect
T2 High Rising	Greater *f0* Rise	No Effect	Slightly raised overall *f0* contour	No Effect
T3 Low Dipping	Greater *f0* Rise	No Effect	Greatly raised overall *f0* contour; Small *f0* rise	No Effect
T4 High Falling	Greater *f0* Rise	No Effect	Slightly raised overall *f0* contour	No Effect

With regard to the overall *f0* mean, all tones (except T1) were significantly raised by the following T3 compared to those preceding other tones. It is important to note that the raising over T3 was significantly greater than that over T2 and T4. In terms of *f0* contour change, what we observed was the consistent *f0* rise over all lexical tones that preceded T1. Their magnitudes of *f0* rise were fairly comparable regardless of the identity of the target tones.

As introduced earlier, there is consensus in the literature that T1T1, T4T1, and T3T3 undergo tone sandhi changes (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma & Jia, 2006; Hyman, 2007). T4T3 has been reported to involve anticipatory tonal coarticulation (Zhang & Liu, 2011). Furthermore, it has been proposed in the literature that T3T2, T3T4 and T4T4 also involve tone sandhi (see Wee et al., 2005; Zhang & Liu, 2011 for T3T2 and T3T4 sandhi; see Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma & Jia, 2006; Hyman, 2007 for T4T4 sandhi). Our results showed that there was no significant tonal alternation in T3T2, T3T4, and T4T4, raising questions to the application of tone sandhi over these sequences. Instead, we observed a clear *f0* contour raising effect of T1 on T2 and T3 (Figures 3.2*b* and 3.2*c*) as well as a general raising effect of T3 on T2 (Figure 3.2*b*), which have never been reported in the literature either as a tone sandhi change or as contextual tonal coarticulation.

Traditionally, tone sandhi in Tianjin Mandarin has always been characterized as the categorical change of one lexical tone to another within the lexical tone inventory (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma & Jia, 2006; Hyman, 2007). With the three tonal sequences that have been claimed to undergo tone sandhi (i.e., T1T1, T4T1, T3T3), which did show contextual tonal variation in our data, we further examined whether they involved categorical sandhi changes into T2/T3T1, T2T1 and T2T3, respectively. Our approach was to compare the *f0* contours of the three sandhi sequences to the *f0* realization of their respective targeted outputs claimed in the literature (i.e., T1T1 vs. T2/T3T1, T4T1 vs. T2T1 and T3T3 vs. T2T3).

Plotted in Figure 3.3 are the *f0* contours of the three sandhi sequences (in dark gray) against the *f0* contours of their respective claimed targeted outputs (in light gray). Specifically, T1T1 (against T2/T3T1) is shown in Figure 3.3*a*, T4T1 (against T2T1) in Figure 3.3*b*, and T3T3 (against T2T3) in Figure 3.3*c*. Statistical results of their planned comparisons were summarized in Table 3.4.

Figure 3.3 shows that all the three tonal sequences were produced significantly differently from their claimed targeted outputs, as confirmed by significant statistical results in Table 3.4. Our results thus confirmed the non-neutralization in the three disyllabic tonal sequences as found in Zhang and Liu (2011). Apparently, tonal neutralization cannot be the criterion of what constitutes tone sandhi. The question is then what differentiates tone sandhi from phonetic tonal coarticulation?

Generally speaking, phonetic coarticulation caused by the following tone should not affect the distinctiveness of a lexical tonal contour (e.g., Gandour et al., 1994; Xu, 1997), so that its characteristic *f0* contour can still be expected to accord with its canonical tonal shape. A phonological tone sandhi alternation, however, usually brings about major changes in the *f0* contour of a lexical tone so that its surface *f0* may be unpredictable from the characteristics of the canonical tonal *f0* contour. Among the seven disyllabic cases that exhibited a statistically significant amount of tonal variance as shown in Table 3.3, both falling tones T1 and T4 were realized with rising *f0* contours due to the contour raising effect triggered by the following T1 (i.e., in T1T1 and T4T1); the low-dipping T3 was realized with a high *f0* register due to the general raising effect from the following T3 (i.e., in T3T3). These altered *f0* realizations were unpredictable from its canonical low tones, and therefore seemed justified to be categorized as tone sandhi. In the other four cases, i.e., T3T1, T2T1, T2T3, T4T3, however, the *f0* realization of the first tone was rather recognizable despite variation due to the following tone. It is not clear therefore whether a sandhi rule is necessary for these four cases.

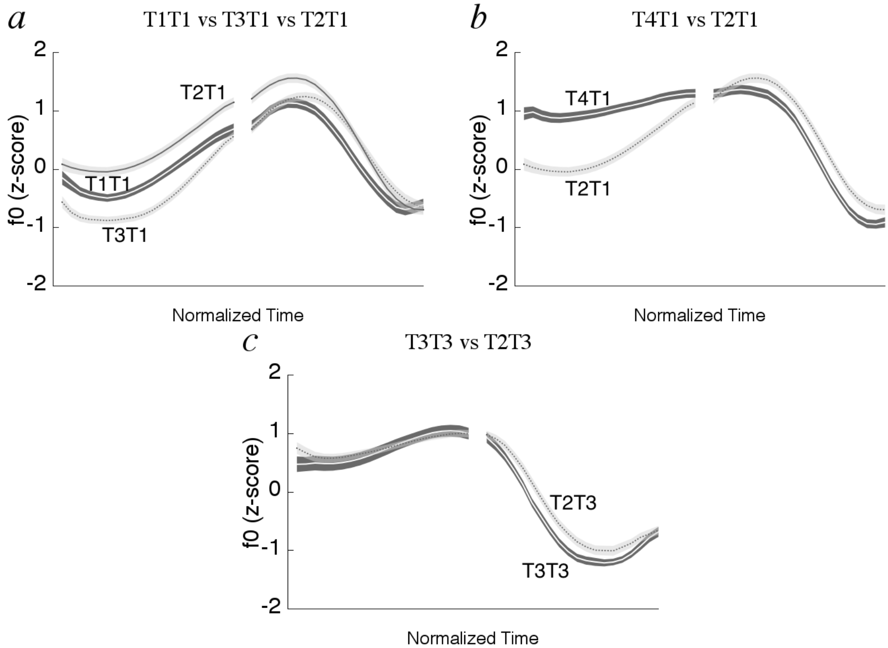

Figure 3.3 *f0 realization of the three confirmed disyllabic sandhi patterns. Thick white lines indicate the mean f0 of the disyllabic sandhi sequences (dark gray areas for ±1 standard error of mean); black lines indicate the mean f0 of the claimed targeted output (light gray areas for ±1 standard error of mean). Normalized time.*

Table 3.4 *Results of growth curve analyses for T1T1 vs.T2T1/T3T1, T4T1 vs. T2T1, T3T3 vs. T2T3. Bolded items for tones under comparison.*

	T1T1		T4T1	T3T3
	T3T1 vs. **T1**T1	**T2**T1 vs. **T1**T1	**T4**T1 vs. **T2**T1	**T3**T3 vs. **T2**T3
Intercept	β=-0.41 t=-12.87 p<0.001	β=0.22 t=8.69 p<0.001	β=0.71 t=19.52 p<0.001	n.s.
Slope	n.s.	n.s.	β=-1.02 t=-11.42 p<0.001	β=0.25 t=2.81 p<0.01
Quadratic	β=0.27 t=2.66 p<0.01	n.s.	β=-0.48 t=-5.37 p<0.001	n.s.

To summarize, we have confirmed three disyllabic tonal sequences to be tone sandhi patterns as reported in the literature, i.e., T1T1, T4T1 and T3T3 (e.g., Li & Liu, 1985;

Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma, 2005; Hyman, 2007). We also found a significant anticipatory raising effect, though with minor modifications, on T2 and T4 before T3, as well as a significant contour raising effect on T2 and T3 before T1. All other tonal sequences were found to induce non-significant (probably perceptually negligible) phonetic changes.

Note that even for the confirmed tone sandhi cases, it is clear that sandhi does not involve categorical change of one lexical tone to another. The lack of tonal neutralization makes it clear that tone sandhi in Tianjin Mandarin should not be characterized as the categorical change of one lexical tone to another. Such a lack of categorical sandhi change also raises serious doubts on the tonal variation patterns of trisyllabic constituents reported in the literature.

3.4 Results & discussion II: Tonal variation in trisyllabic domain

In this section, we examined tonal variation in Tianjin Mandarin over trisyllabic domains. We limited our attention to the three disyllabic sandhi processes confirmed so far and investigated how these sandhi changes are applied in trisyllabic sequences. Each trisyllabic sequence was compared to the targeted output pattern as claimed in the literature. Table 3.5 concludes all the sandhi-involved trisyllabic combinations discussed in the literature based on the three disyllabic processes. (Some combinations might appear twice in this table due to the fact that some disyllabic patterns are involved in more than one trisyllabic combinations.) We separated our analyses according to the classification of simple vs. complex tone sandhi contexts proposed in the literature. In the simple sandhi contexts, the directionality of sandhi application was investigated whereas in the complex sandhi contexts (bolded), we aimed to tap into the iterativity of sandhi application in addition to the directionality.

Table 3.5 *All trisyllabic tone sandhi combinations proposed in the literature involving T1T1, T4T1 and T3T3. The bold items are those possibly involving complex tonal contexts.*

	T1T1		T4T1	
	Left-aligned	Right-aligned	Left-aligned	Right-aligned
T1	**T1T1T1**		**T4T1T1**	T1T4T1
T2	T1T1T2	T2T1T1	T4T1T2	T2T4T1
T3	T1T1T3	**T3T1T1**	T4T1T3	T3T4T1
T4	T1T1T4	**T4T1T1**	T4T1T4	T4T4T1
	T3T3			
	Left-aligned	Right-aligned		
T1	T3T3T1	T1T3T3		
T2	T3T3T2	T2T3T3		

T3	**T3T3T3**	
T4	T3T3T4	T4T3T3

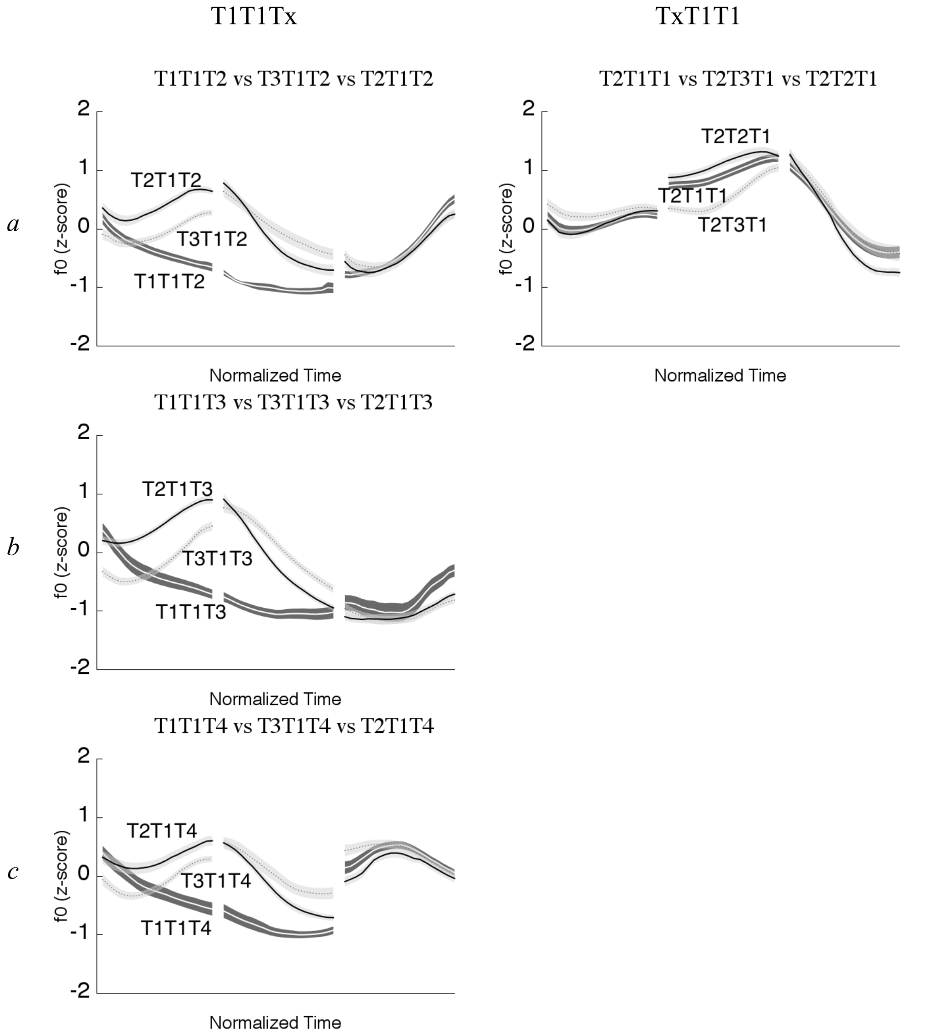

Figure 3.4 *Trisyllabic sandhi contexts containing T1T1. White thick lines indicate the mean f0 patterns of the trisyllabic sandhi sequences (dark gray areas for ±1 standard error of mean); black dotted lines indicate the mean f0 patterns of the claimed targeted output containing T3T1 (light gray areas for ±1 standard error of mean); black thick lines indicate the mean f0 patterns of the claimed targeted output containing T2T1 (light gray areas for ±1 standard error of mean). Normalized time.*

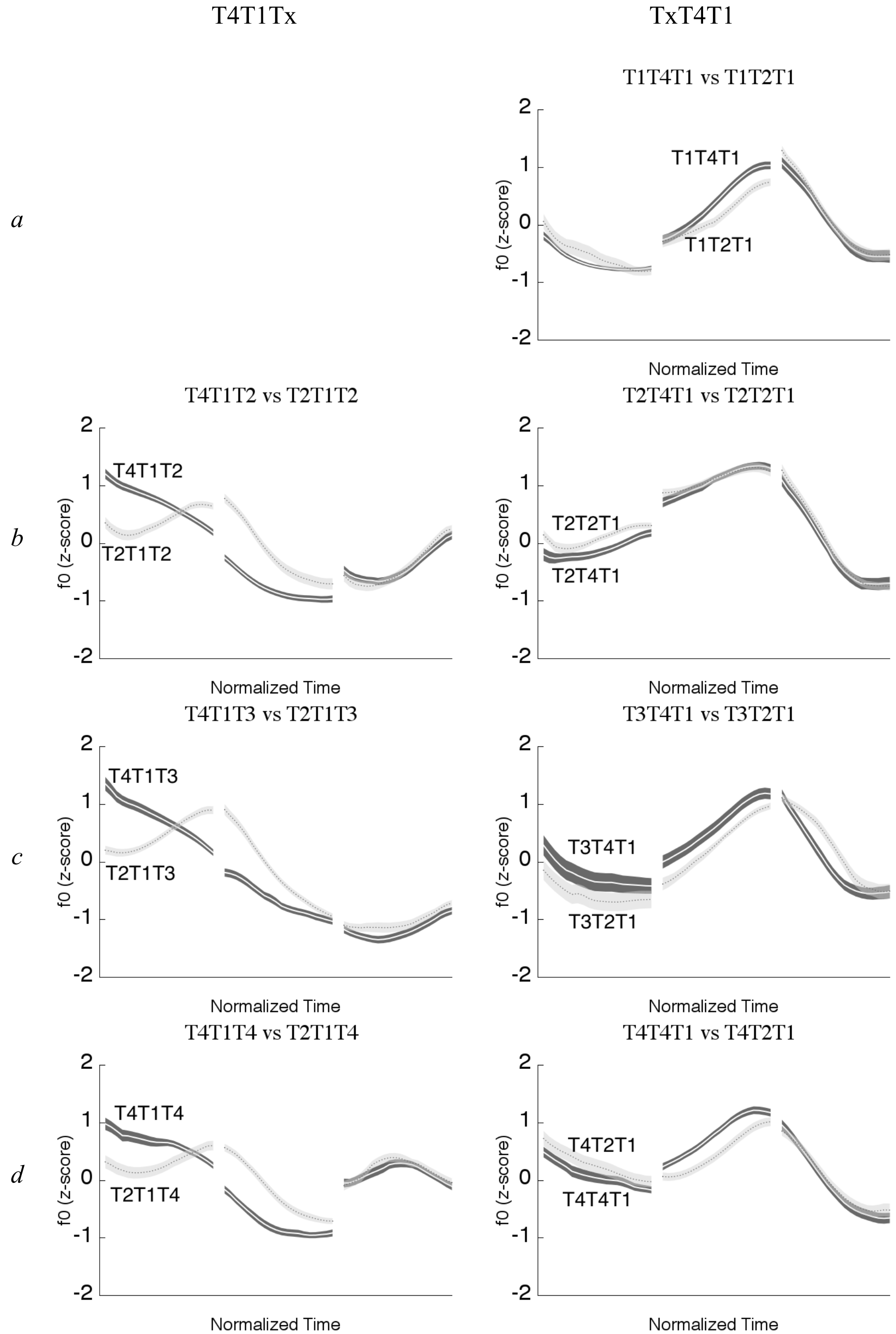

Figure 3.5 *Trisyllabic sandhi contexts containing T4T1. White thick lines indicate the mean f0 patterns of the trisyllabic sandhi sequences (dark gray areas for ±1 standard error of mean) and black dotted lines indicate the mean f0 patterns of the claimed targeted outputs (light gray areas for ±1 standard error of mean). Normalized time.*

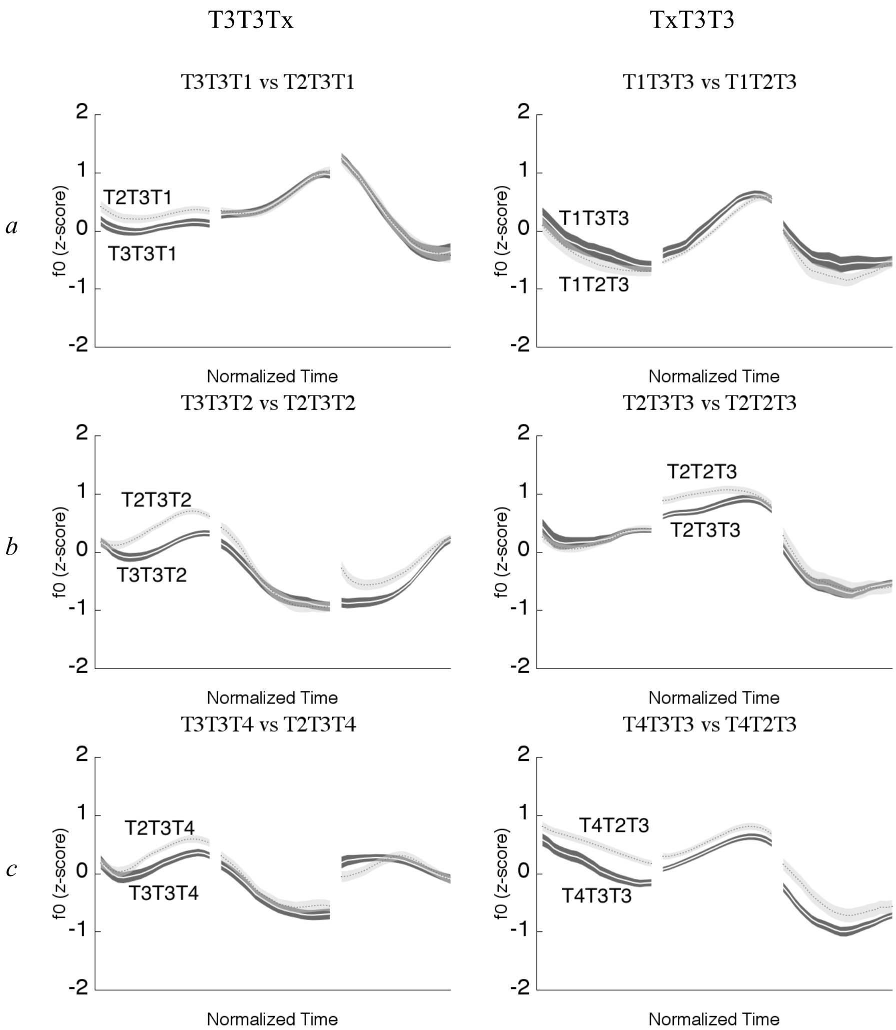

Figure 3.6 *Trisyllabic sandhi contexts containing T3T3. White thick lines indicate the mean f0 patterns of the trisyllabic sandhi sequences (dark gray areas for ±1 standard error of mean) and black dotted lines indicate the mean f0 patterns of the claimed targeted outputs (light gray areas for ±1 standard error of mean). Normalized time.*

Table 3.6 *Results of growth curve analyses for T1T1Tx vs. TxT1T1. Bolded items for tones under comparison.*

	T1T1Tx		TxT1T1	
	T1T1T2 vs. **T2**T1T2	**T1**T1T2 vs. **T3**T1T2	T2**T1**T1 vs. T2**T2**T1	T2**T1**T1 vs. T2**T3**T1
Intercept	β=0.93 t=38.35 p<0.001	β=0.56 t=18.91 p<0.001	n.s.	β=-0.28 t=-13.05 p<0.001

Slope	β=1.78 t=19.69 p<0.001	β=1.73 t=19.06 p<0.001	n.s.	β=0.36 t=4.09 p<0.001
Quadratic	n.s.	n.s.	β=-0.21 t=-2.36 p<0.05	β=0.29 t=3.26 p<0.01
	T1T1T3 vs. **T2**T1T3	**T1**T1T3 vs. **T3**T1T3		
Intercept	β=1.05 t=21.10 p<0.001	β=0.47 t=10.87 p<0.001		
Slope	β=2.48 t=21.15 p<0.001	β=2.63 t=22.46 p<0.001		
Quadratic	n.s.	n.s.		
	T1T1T4 vs. **T2**T1T4	**T1**T1T4 vs. **T3**T1T4		
Intercept	β=0.91 t=26.35 p<0.001	β=0.50 t=11.53 p<0.001		
Slope	β=1.77 t=14.67 p<0.001	β=2.04 t=16.82 p<0.001		
Quadratic	n.s.	n.s.		

Table 3.7 *Results of growth curve analyses for T4T1Tx vs. TxT4T1. Bolded items for tones under comparison.*

	T4T1Tx	TxT4T1
		T1**T4**T1 vs. T1**T2**T1
Intercept		β=0.25, t=7.71, p<0.001
Slope		β=0.48, t=4.83, p<0.001
Quadratic		β=-0.25, t=-2.49, p<0.05
	T4T1T2 vs. **T2**T1T2	T2**T4**T1 vs. T2**T2**T1
Intercept	β=0.17, t=4.40, p<0.001	β=0.05, t=2.42, p<0.05
Slope	β=-2.05, t=-24.07, p<0.001	β=0.25, t=2.89, p<0.01
Quadratic	β=-0.31, t=-3.68, p<0.001	n.s.
	T4T1T3 vs. **T2**T1T3	T3**T4**T1 vs. T3**T2**T1
Intercept	n.s.	β=0.30, t=3.48, p<0.01
Slope	β=-3.00, t=-32.89, p<0.001	n.s.
Quadratic	β=-0.33, t=-3.71, p<0.001	n.s.
	T4T1T4 vs. **T2**T1T4	T4**T4**T1 vs. T4**T2**T1
Intercept	β=0.71, t=17.13, p<0.001	n.s.
Slope	β=-1.41, t=-12.88, p<0.001	n.s.

Quadratic	β=-0.41, t=-3.77, p<0.001	β=-0.36, t=-4.25, p<0.001

Table 3.8 *Results of growth curve analyses for T3T3Tx vs. TxT3T3. Bolded items for tones under comparison.*

	T3T3Tx	TxT3T3
	T3T3T1 vs. **T2**T3T1	T1**T3**T3 vs. T1**T2**T3
Intercept	n.s.	β=-0.55, t=-11.47, p<0.001
Slope	n.s.	n.s.
Quadratic	n.s	β=-0.19, t=-2.45, p<0.05
	T3T3T2 vs. **T2**T3T2	T2**T3**T3 vs. T2**T2**T3
Intercept	β=-0.09, t=-2.87, p<0.01	β=-0.07, t=-2.03, p<0.05
Slope	β=-0.45, t=-5.98, p<0.001	β=0.33, t=4.54, p<0.001
Quadratic	β=0.42, t=5.54, p<0.001	n.s.
	T3T3T4 vs. **T2**T3T4	T4**T3**T3 vs. T4**T2**T3
Intercept	β=-0.29, t=-8.45, p<0.001	β=-0.14, t=-7.35, p<0.001
Slope	β=-0.33, t=-3.57, p<0.001	n.s.
Quadratic	β=0.23, t=2.50, p<0.05	n.s.

3.4.1 Simple sandhi contexts

We set up the simple sandhi contexts as aligning a disyllabic sandhi tonal sequence either to the left or right of a trisyllabic constituent, without involving potentially multiple applications of tone sandhi. In Figures 3.4-3.6, the *f0* contours of the trisyllabic sequences (in dark gray) were compared to the *f0* realization of their respective sandhi targeted outputs as claimed in the literature (in light gray). Note that those plotted on the left have the sandhi contexts in the first two syllables (left-aligned); and those on the right have the sandhi contexts in the last two syllables (right-aligned). Merged *f0* contours are expected if there is neutralization in tone sandhi. Again, for the case of T1T1, both T2T1 and T3T1 were included given that both possibilities have been proposed in the literature. The statistical results were summarized in Tables 3.6-3.8.

3.4.1.1 T1T1Tx vs. TxT1T1

As shown in Figure 3.4, the application of T1T1 sandhi can only be observed when T1T1 is right aligned. However, the sandhi-T1 was not neutralized with either T3 or T2. The non-neutralization pattern confirms what we have observed in disyllabic T1T1 tonal sequences. Specifically, when T1T1 was left aligned (i.e., T1T1T2, T1T1T3, T1T1T4 in Figures 3.4*a-c*, left column), the first T1 was clearly realized with a falling *f0* contour, which was the opposite of that of T2/T3, as confirmed by the significant differences in the slope (see T1T1Tx in Table 3.6). When T1T1 was right aligned as in T2T1T1 (Figure 3.4*a*, right), the middle T1 was realized with a rising *f0* contour, suggesting the application of tone sandhi (Figure 3.3*a*). However, the significant statistical results suggested no tonal neutralization between the middle T1 vs. its claimed output T2 or T3 (see T1T1Tx in Table 3.6).

3.4.1.2 T4T1Tx vs. TxT4T1

As shown in Figure 3.5, T4T1 also undergoes tone sandhi only when it is right aligned, as observed earlier for T1T1. Again, no tonal neutralization was confirmed between T4 and its claimed output T2. Specifically, when T4T1 was left aligned (i.e., T4T1T2, T4T1T3, T4T1T4 in Figures 3.5*a-c*, left column), tone sandhi could not have applied, given its falling *f0* contour. This was confirmed by the significant slope difference between T2 and T4 in all the left-aligned cases (T4T1Tx in Table 3.7). When T4T1 was right aligned (i.e., T1T4T1, T2T4T1, T3T4T1, T4T4T1 in Figure 3.5*a-d*, right), the rising *f0* contour over the T4-bearing syllable suggested the application of tone sandhi (as comparable to that in disyllabic sequences shown in Figure 3.3*b*), but again no neutralization was confirmed. See significant statistical results of the *f0* differences between TxT4T1 vs. TxT2T1, summarized in Table 3.7.

3.4.1.3 T3T3Tx vs. TxT3T3

Figure 3.6 displays the T3T3 sequence in the left-aligned contexts (T3T3T1, T3T3T2, T3T3T4 in Figures 3.6*a-c*, left) and the right-aligned contexts (T1T3T3, T2T3T3, T4T3T3 in Figures 3.6*a-c*, right). Note that T3T3T3 will be discussed in Section 4.2.1 due to its potentially complex tone sandhi application. In both T3T3Tx and TxT3T3, the first T3 showed a rising *f0* contour, suggesting the consistent application of the T3 sandhi, as observed in disyllabic T3T3 sequence (Figure 3.3*c*). However, sandhi-T3 showed significant differences from T2 in all the other tonal contexts (see significant results in both T3T3Tx and TxT3T3 in Table 3.8), except for the sequence T3T3T1. This confirmed again the general lack of complete neutralization in these tone sandhi cases.

3.4.2 Complex sandhi contexts

Four trisyllabic sequences have been claimed to involve potentially more complex sandhi changes, due to the so-called iterative process of tone sandhi application (e.g., Chen, 2000). Note that the iterative process of sandhi application stems from the assumption in the literature that tone sandhi in Tianjin Mandarin involves the categorical change of one lexical tone to another, which leads to iterative sandhi application due to a new sandhi context resulting from the first round of sandhi application.

Given that T3T3 sandhi applies both when left aligned and right aligned, while T1T1 and T4T1 show sandhi change only when right aligned, we examined the possibilities listed in Table 3.9, where the bolded cases have been reported in the literature. Our results should shed light on two issues. One concerns the directionality of tone sandhi application when both directions provide sandhi contexts. The other concerns iterative sandhi application. For ease of exposition, we followed the sandhi rules claimed in the literature, and described the targeted outputs as results of the iterativity of sandhi application where the intermediate step would create a new tone sandhi context (i.e., if tone sandhi indeed involved categorical change of one lexical tone to another). Statistical results were shown in Tables 3.10-3.12.

Table 3.9 *Possible targeted outputs of complex tonal contexts according to the literature.*

	From the left		From the right	
T3T3T3	(1st Round) **T2T3T3**	(2nd Round) **T2T2T3**	T3T2T3	
T3T1T1	No Sandhi		(1st Round) **T3T3T1**	(2nd Round) **T2T3T1**
T1T1T1	No Sandhi		**T1T3T1**	
T4T1T1	No Sandhi		**T4T3T1**	

3.4.2.1 T3T3T3

For the combination T3T3T3, both directions of sandhi application are possible. According to the claimed rules in the literature, both T2T2T3 (from the left) and T3T2T3 (from the right) are possible outputs, where a chain effect is expected if the sandhi rule is applied from the left edge through an intermediate stage T2T3T3. Figure 3.7 compares the three possible outcomes (Figure 3.7*a* from left; Figure 3.7*b* from right). As can be seen from Figure 3.7, T3T3 sandhi was more likely to be applied from the left edge rather than from the right.

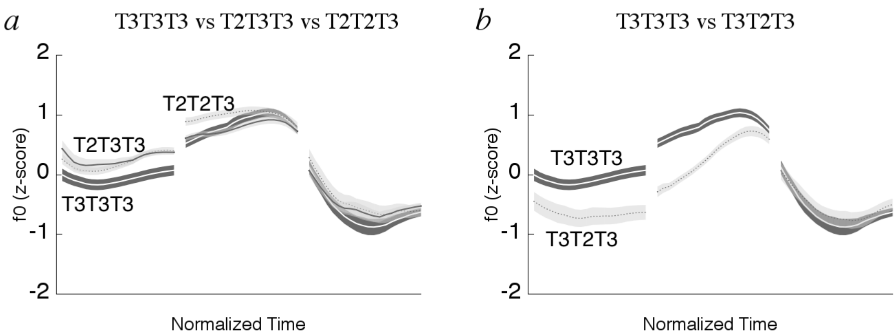

Figure 3.7 *The comparison of two possible directions of tone sandhi application in the sequence T3T3T3. White thick lines indicate the mean f0 patterns of T3T3T3 in both graphs (dark gray areas for ±1 standard error of mean) and black dotted lines indicate the mean f0 patterns of the claimed targeted outputs with different directionalities (light gray areas for ±1 standard error of mean). Graph a indicates the left-initial direction and Graph b the right-initial direction. Normalized time.*

Table 3.10 *Results of growth curve analyses for T3T3T3 vs. T2T3T3/T3T2T3/T2T2T3.*

	a. Left-Aligned	*b. Right-Aligned*	
	T3T3T3 vs. T2T3T3	T3T3T3 vs. T3T2T3	
	1st Tone	1st Tone	2nd Tone
Intercept	β=0.21 t=4.97 p<0.001	β=0.38 t=6.89 p<0.001	β=0.85 t=20.06 p<0.001
Slope	n.s.	β=0.39 t=2.53 p<0.05	β=-0.99 t=-11.61 p<0.001
Quadratic	n.s.	n.s.	n.s.
	T3T3T3 vs. T2T2T3		
	1st Tone	2nd Tone	
Intercept	β=-0.12 t=-4.38 p<0.001	β=0.47 t=7.88 p<0.001	
Slope	β=-0.23 t=-2.49 p<0.05	β=0.52 t=6.26 p<0.001	
Quadratic	n.s.	n.s.	

For the left-aligned sandhi application possibilities (Figure 3.7*a*), T3T3T3 was similar to both T2T3T3 and T2T2T3, but was not completely neutralized with either. T3T3T3 and T2T3T3 were significantly different in terms of the mean *f0* of the first tone. (See significant results for T3T3T3 vs. T2T3T3 in Table 3.10*a*.) T3T3T3 and T2T2T3 differed even more. There were significant differences in the overall *f0* mean and slope of both the first and second tones. (See significant results for T3T3T3 vs. T2T2T3 in Table 3.10*a*.) However, given the similarity of the contour shapes, it is likely that the second T3 changed to T2 in T3T3T3, although it is difficult to pin down that there was sandhi application over the second T3 given that the two contours were nevertheless quite different.

Turning to the right-aligned sandhi possibility (Figure 3.7*b*), T3T3T3 and T3T2T3 were significantly different in the overall *f0* mean and the slope for both the first and second tones. (See significant results of the intercept for T3T3T3 vs. T3T2T3 in Table 3.10*b*.) The significant and salient *f0* differences between the two tonal sequences suggest that T3 sandhi was unlikely to be applied from the right edge.

3.4.2.2 T3T1T1

For T3T1T1, the literature has claimed iterative sandhi application initiated from the right edge through an intermediate stage T3T3T1 which in turn changes to T2T3T1, as shown in the second line of Table 3.9. In Figure 3.8, the surface *f0* pattern of T3T1T1 was compared to that of both potential outputs.

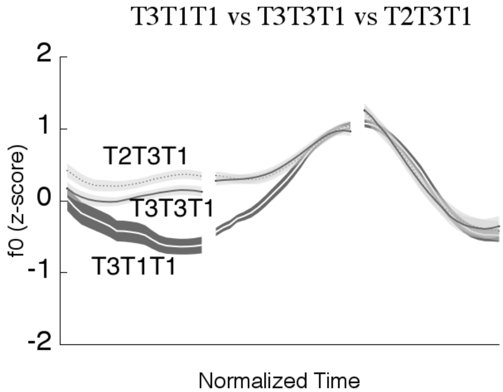

Figure 3.8 *The comparison of the sequences T3T1T1, T3T3T1 and T2T3T1. White thick lines indicate the mean f0 patterns of T3T1T1 (dark gray areas for ±1 standard error of mean), black thick lines indicate the mean f0 patterns of T3T3T1 (light gray areas for ±1 standard error of mean), black dotted lines indicate the mean f0 patterns of T2T3T1 (light gray areas for ±1 standard error of mean). Normalized time.*

Table 3.11 *Results of growth curve analyses for T3T1T1 vs. T3T3T1/T2T3T1.*

	a. Left-Aligned		*b. Right-Aligned*	
	T3T1T1 vs. T3T3T1		T3T1T1 vs. T2T3T1	
	1st Tone	2nd Tone	1st Tone	2nd Tone
Intercept	n.s.	β=0.46 t=14.32 p<0.001	β=-0.72 t=-20.30 p<0.001	β=0.24 t=11.48 p<0.001
Slope	β=0.98 t=7.00 p<0.001	β=-1.20 t=-14.39 p<0.001	β=-0.94 t=-6.70 p<0.001	β=-1.12 t=-13.23 p<0.001
Quadratic	n.s.	n.s.	n.s.	β=0.30 t=3.55 p<0.001

First, in Figure 3.8, the first T3 in T3T1T1 was realized with a low-falling contour, which was significantly different from the rising *f0* over T2 or T3. (See significant results in the slopes of the first tone for T3T1T1 vs. T3T3T1 and T3T1T1 vs. T2T3T1 in Table 3.11.) Second, the middle tone in T3T1T1 was realized with a rising *f0* contour, which indicates the application of tone sandhi here (compared to Figure 3.3*a*). However, the middle T1 was not fully neutralized with the claimed output T3 in either T3T3T1 or T2T3T1. (See significant results for T3T1T1 vs. T3T3T1 and T3T1T1 vs. T2T3T1 in Table 3.11.) This suggests one-step tone sandhi from the right edge for T3T1T1, without any chain effect.

3.4.2.3 T1T1T1 & T4T1T1

For both T1T1T1 and T4T1T1, only one-step sandhi application from the right edge has been proposed in the literature. Figure 3.9 illustrates *f0* realization of the two sequences, compared to their respective claimed outputs T1T3T1 (Figure 3.9*a*) and T4T3T1 (Figure 3.9*b*).

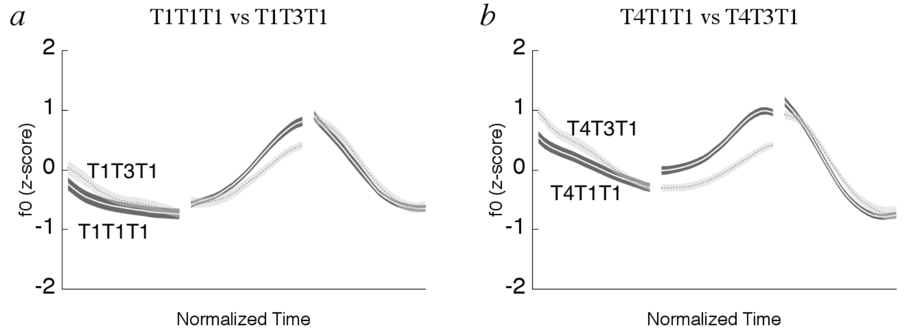

Figure 3.9 *Comparison of sequences T1T1T1 vs T1T3T1 and T4T1T1 vs T4T3T1. White thick lines indicate the mean f0 of the trisyllabic sandhi sequences (dark gray areas for ±1 standard error of mean) and black dotted lines for the mean f0 of the claimed targeted outputs (light gray areas for ±1 standard error of mean). Normalized time.*

Table 3.12 *Results of growth curve analyses for T1T1T1 vs. T1T3T1 and T4T1T1 vs. T4T3T1. Bolded items for tones under comparison.*

	T1**T1**T1 vs. T1**T3**T1	T4**T1**T1 vs. T4**T3**T1
Intercept	β=-0.40, t=-14.23, p<0.001	β=-0.96, t=-16.26, p<0.001
Slope	β=-0.62, t=-6.77, p<0.001	β=-0.54, t=-7.04, p<0.001
Quadratic	n.s.	n.s.

In both cases, sandhi application from the right edge was confirmed, again with no complete neutralization involved. Figure 3.9*a* shows that the middle T1 was realized with a rising *f0*, suggesting the application of T1T1 tone sandhi from the right edge of the trisyllabic constituent. However, it was still not neutralized with its claimed targeted output T3, as suggested by the significant differences between the middle T1 and T3. (See significant results for T1T1T1 vs. T1T3T1 in Table 3.12.) Similarly, tone sandhi in T4T1T1 also seemed to be applied from the right edge (Figure 3.9*b*), without complete tonal neutralization. (See significant results for T4T1T1 vs. T4T3T1 in Table 3.12.)

3.4.3 Discussion

For trisyllabic sequences, we were interested in the issue of directionality and iterativity of tone sandhi application. We separated our data into two sets. One concerns directionality of tone sandhi application in simple contexts where only one round of sandhi application is possible. We varied the position of the sandhi sequences, either to the left or to the right edge of a trisyllabic tone sequence. What has been claimed in the literature is that in such contexts, tone sandhi in Tianjin Mandarin applies consistently regardless of the positional alignment of the tonal sequences (e.g., Li & Liu, 1985; Chen, 2000; Ma, 2005; Wee et al., 2005). Our results, however, did not support this observation. Our data showed that tone sandhi application over T1T1 and T4T1 sequences was only possible when these disyllabic sequences were aligned to the right edge of a trisyllabic constituent; while the T3T3 sequence triggered consistent application of sandhi changes regardless of its position within a trisyllabic utterance.

Our second set of data examined complex sandhi contexts, where the literature has claimed both different directionalities of sandhi application as well as iterative sandhi processes due to the derived sandhi context (i.e., after an initial sandhi application process over the sequence). Two findings are to be noted. One is that we indeed found directional constraints on sandhi application. For T3T3T3 sequences, we observed only left-initiated tone sandhi application, in sharp contrast with the flexibility of T3 sandhi in the simple sandhi contexts. For the T1T1 and T4T1 sequences, we confirmed our observation in the simple contexts where tone sandhi was found to apply only when they were right aligned.

The non-neutralization of sandhi application was further confirmed in our trisyllabic data. We have shown that the *f0* realization of most trisyllabic sandhi combinations was significantly different from the claimed targeted outputs. If categorical change of one lexical tone to another was truly involved in tone sandhi, we would have expected all trisyllabic tone sandhi sequences to be realized exactly as the claimed targeted outputs. This, however, was not borne out. Furthermore, no iterativity of sandhi application was observed in our data. This lends further evidence to the argument that tone sandhi does not involve the categorical change of one lexical tone to another, which otherwise could have created a newly-derived sandhi context for another round of sandhi application.

3.5 General discussion and conclusion

This study illustrated the rich layers of lexical tone variability induced by different local contextual tonal variation processes, with empirical evidence from Tianjin Mandarin. Two specific questions have been addressed concerning the received wisdom of tonal variation in Tianjin Mandarin: (1) How does the *f0* realization of a lexical tone vary as a function of different following tones within a disyllabic domain? (2) How are the disyllabic tone sandhi patterns realized in trisyllabic sequences?

3.5.1 Tonal variation in disyllabic domain

With regard to disyllabic sequences, we aimed to fully understand the various patterns of *f0* deviation due to different contextual tonal variation processes, namely tonal anticipatory coarticulation and tone sandhi. To this end, we investigated all possible combinations of the four lexical tones in the language. In particular, we zoomed into the effect of the following tones (on the second syllable) on the *f0* realization of the preceding tone (on the first syllable), because this is a comparable context to examine both tonal coarticulation and tone sandhi as Tianjin Mandarin shows right-dominant tone sandhi patterns.

With the technique of Growth Curve Analysis (GCA), we were able to objectively quantify whether and how *f0* contours differ from each other in terms of the overall *f0* means, the general *f0* direction and the *f0* contour shape, as a function of the following lexical tones. GCA with linear mixed-effects modelling also helped to take better care of not only the time-varying nature of *f0* contours but also variation due to individual participants. The methodology thus enabled us to better observe the *f0* variation and gain further understanding of the different sources of *f0* variation.

Two patterns emerged from the data, as summarized in Table 3.13. The first is that all significant *f0* variation observed over the first tone is triggered by a low-register lexical tone in Tianjin Mandarin (i.e., T1 or T3). T1 induces a consistent contour raising effect, while T3 exerts a more general raising effect on the preceding tones (except for T1 followed by T3). This in part confirms the findings of Zhang and Liu (2011), which only examined however the anticipatory tonal coarticulation in T1T3, T2T3, T4T3, and T2T1, and reported a significant raising effect only in T4T3. Our study, with all tonal combinations included and more powerful statistical analyses, showed that there was also a subtle but significant raising effect in T2T1, T3T1 and T2T3.

The question that arises is why both T1 and T3 tend to raise the preceding lexical tones? Within the Sinitic family, Xu (1997) reported a rather subtle anticipatory raising effect due to the low onset of the following tone, with data from Standard Chinese. In the case of Tianjin Mandarin, though a closely related Mandarin dialect, the raising effect seems to be dialect-specific. T1 and T2 are not that different in terms of their initial *f0* height. If anticipatory raising was purely phonetic due to the onset *f0* value of the following tone, we should have observed similar anticipatory raising effects. Our results, however, showed that they behaved completely differently in their effects on preceding tones. It is probably the tonal register (i.e., the low register in T1 and T3) that has played a role in conditioning anticipatory raising, suggesting the possibility of phonologically-conditioned anticipatory raising effects. It is to be noted that significant anticipatory effects have also been reported for Malaysian Hokkien (Chang & Hsieh, 2012), where carryover and anticipatory tonal coarticulation show similar magnitude due to the final prominence of the following syllable. These language-specific patterns contribute to the typology of contextual tonal variation, with data that challenge the general asymmetrical effects of tonal coarticulation (with greater carryover effects than anticipatory effects) observed in the literature (see Chen, 2012 for a review).

Second, we observed that for the sequences T1T1, T4T1 and T3T3, they either changed the *f0* trajectory (in the case of T1T1 and T4T1) or substantially raised the overall *f0* register (in the case of T3T3), which indeed seem appropriate to be characterized as due to tonal sandhi (as claimed in the literature). For the other four sequences (i.e., T2T1, T3T1, T2T3, T4T3), the lexical tones were realized with an *f0* contour that actually resembles the canonical realizations of the tones despite significant *f0* deviation. Therefore it is not clear whether tone sandhi has been applied or whether they are better accounted for as anticipatory tonal coarticulation.

Table 3.13 *Summary of the contexts and tonal variability type of first tone within a disyllabic constituents.*

1st Tone		2nd Tone			
		Low Register		High Register	
		T1	T3	T2	T4
Low Register	T1	Sandhi Rising	No Effect	No Effect	No Effect
	T3	Anticipatory Rising	Sandhi Raising	No Effect	No Effect
High Register	T2	Anticipatory Rising	Anticipatory Raising	No Effect	No Effect
	T4	Sandhi Rising	Anticipatory Raising	No Effect	No Effect

Previous impressionistic studies have claimed that T4T4 also undergoes sandhi change to T1T4 (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma, 2005; Hyman, 2007). Our data, however, did not show any significant *f0* variation over T4 due to the following T4. If T4T4 underwent sandhi change, the first T4 should have surfaced with a low-register falling *f0*. On the contrary, Figure 3.2*d* shows that the first T4 in T4T4 was realized with a high-register falling contour, similar to the T4 in T4T2 and T4T3. This is consistent with findings in recent studies (e.g., Liu & Gao, 2003; Zhang & Liu, 2011), which argued that the T4T4 sandhi process has become "obsolete", possibly due to the influence from Standard Chinese, where there is no tone sandhi over the T4T4 sequence.

When followed by T2 or T4, T3 was realized with a low-falling contour, resembling the canonical realization of T1 in isolation. This has led to the claim that T3T2 and T3T4 also involve tone sandhi change into the lexical T1 (e.g., Wee et al., 2005) or into a half T3 (Zhang & Liu, 2011).[3] Ma and Jia (2006) showed that listeners could reliably distinguish T3T2 from T1T2 and T3T4 from T1T4. Moreover, it is important to note that in our data, neither T2 nor T4 seems to introduce any anticipatory effect on its preceding tones unlike as observed for T1 and T3. We thus adopt the view that the half-realized T3 is due to tonal reduction in connected speech, which seems to be a more general phonetic reduction of dipping tone in connected speech as is also evident in Standard Chinese. However, as pointed out by one reviewer, we cannot rule out the possibility that tonal reduction is phonologized and should be classified as tone sandhi, as proposed in Zhang and Liu (2011).

There have been some attempts in the last decades to establish objective criteria to distinguish between tone sandhi and tonal coarticulation. Shen (1992) proposes a three-way distinction between the two types of tonal variation. 1) Tone sandhi is constrained by language-specific morphophonemics, while tonal coarticulation is a language-independent mechanism. 2) Tone sandhi is either assimilatory or dissimilatory, while tonal coarticulation is only assimilatory. 3) Tone sandhi involves tonal identity change from one tone to another while the tonal identity is maintained in tonal coarticulation. None of these criteria, however, really succeeded in differentiating the two processes, as also discussed in Chen

[3] The falling contour of T3 is also comparable to the half low tone realization in Standard Mandarin (Chao, 1968), where the final rise of the dipping T3 is always absent in non-prepausal positions (but present when followed by a neutral tone as discussed in details in Chen & Xu, 2006 and Prom-on, Liu & Xu, 2012).

(2000). First, tone sandhi and tonal coarticulation could both be affected by morphophonemics as suggested in Wu (1985). Second, cross-linguistic evidence has shown that tonal coarticulation could be either assimilatory or dissimilatory (e.g., Gandour, et al., 1994 for Thai; Xu, 1997 for Beijing Mandarin; Zhang & Liu, 2011 for Tianjin Mandarin; Chang & Hsieh, 2012 for Malaysian Hokkien). Third, tone sandhi does not always involve tonal identity change. For example, in the T1T1 sequence, our results show that the first T1 is clearly not changed from T1 to T3 as claimed in the literature, but the magnitude of the change suggests the possibility of phonologization of such an anticipatory raising effect. These observations led to the proposition in Chen (2000) that there is actually "no essential difference" between the two processes.

As the boundary between tone sandhi and tonal coarticulation is not straightforward and proves difficult to establish, this makes it even more important for us to first objectively observe the patterns of *f0* variation as we have done in the current study. Without this, any discussion over the difference between these two variation types is doomed to be without foundation.

Through the extensive comparisons of the full dataset with all four lexical tones in both the first and second syllables, we can see that the raising effects due to T3 and T1 indeed need to be further differentiated. The T3 category across Mandarin dialects is known for its sandhi process over a T3T3 sequence, regardless of the specific *f0* contour of the lexical tone category (Shih, 1986). This process has been proposed to date back to the 16th century (Mei, 1977; also see discussion in Zhang, 2010). As for T1T1 and T4T1, a significant raising effect was observed, which is in contrast to the more subtle raising effect of T1 over T2T1 and T3T1. We believe that this difference is related to the saliency of cues for *f0* raising before T1. In the case of two falling tones, i.e., T1 and T4, the raising effect can be effectively implemented, which probably led to an enhanced realization of *f0* raising and consequently a phonologized contextual tonal change; while in the case of T2 and T3, both with a rising *f0* contour in the canonical *f0* shape, there is no further room to signal the raising due to T1, which therefore remains a more subtle coarticulatory effect.

The raising effects due to T3 and T1 can be further teased apart in terms of how closely the sandhi outputs resemble a lexical tone in the native system. T3T3 is a near-merger (following Yu, 2007), in that the derived T3 sandhi realization shows an almost merged *f0* pattern to that of a lexical high-register rising tone, i.e., T2T3. This explains why in the literature T3T3 is described as changing into T2T3. T1T1 and T4T1, however, are clearly no-mergers. Their derived sandhi outputs are clearly different from any of the lexical tones in the system, in contradiction to the claim that they change to T3/T2T1 and T2T1 respectively. Our study, thus far, is the first to illuminate this distinction. An interesting question that is currently addressed in a perception study is how such different patterns of tone sandhi affect the way lexical tones are processed in online speech recognition.

3.5.2 Tonal variation in trisyllabic domain

Our disyllabic data established firmly that tone sandhi in Tianjin Mandarin should not be characterized as the categorical change of one lexical tone to another, confirming earlier findings in Zhang and Liu (2011) and Li and Chen (2012). This finding also led to the absolute necessity of further investigation on trisyllabic tone sandhi patterns in Tianjin Mandarin, which have raised numerous critical issues in the theoretical treatment of tone sandhi in general (e.g., Chen, 2000). Our approach was to first investigate the empirical foundation of the reported complexity in Tianjin Mandarin tone sandhi over trisyllabic sequences. To this end, we have analyzed the *f0* realization of trisyllabic tonal combinations with two more specific goals in mind.

The first was on the directionality of sandhi application in simple tonal contexts with only one disyllabic sequence involving potential tone sandhi. The three tone sandhi sequences (i.e., T1T1, T4T1, and T3T3) were embedded in trisyllabic constituents, either left-aligned (e.g., T1T1Tx) or right-aligned (e.g., TxT1T1). Different from the claims in the literature, we observed that while T3T3 sandhi applied in either direction, T1T1 and T4T1 surfaced with their predicted sandhi *f0* contours only when they were right aligned. This is the very first experimental finding on Tianjin trisyllabic tone sandhi. Our finding therefore greatly simplifies the hitherto claimed complex directionality of tone sandhi in Tianjin Mandarin.

Our second goal was to examine further the claimed iterativity of sandhi application, which refers to the application of tone sandhi over derived sandhi contexts due to an earlier round of tone sandhi application. Specifically, we examined the *f0* contours of the underlying sequences of T3T3T3, T3T1T1, T1T1T1, and T4T1T1. Interestingly, we noticed a further directionality restriction over T3T3T3 sequence where T3 sandhi seemed to only apply from the left edge, resulting in the ambiguity between $T3_{sandhi}$T3T3 and $T3_{sandhi}T3_{sandhi}$T3. Either pattern, however, does not have to involve iterative sandhi application. Furthermore, no iterativity of sandhi application in T3T1T1 was found, even though it has been claimed in the literature that the change of T1T1 to T3T1 creates an intermediate stage T3T3T1 which triggers a further sandhi change into T2T3T1. For the other two tonal sequences concerning T1T1 (i.e., T1T1T1 and T4T1T1), our data only showed one-step tone sandhi application from the right-aligned disyllabic sequences. In short, no iterative sandhi application can be confirmed in our data.

Taking together both disyllabic and trisyllabic data, we have observed two types of tone sandhi in Tianjin Mandarin, as summarized in Table 3.14. T3T3 is a near-merge case. Its sandhi output $T3_{sandhi}$T3 shows an almost merged *f0* contour with the lexical tonal sequence of T2T3. In a simple trisyllabic sandhi context, this sandhi process can occur from either direction regardless of its alignment within the constituent. However, in a complex trisyllabic sandhi context, T3T3 surfaces with a pattern that suggests only left-aligned sandhi application, resulting in an *f0* contour that can be characterized as $T3_{sandhi}$T3T3 or $T3_{sandhi}T3_{sandhi}$T3. T1T1 and T4T1 can be classified as no-mergers, given that their sandhi realizations do not resemble the *f0* contours of any of the lexical tones in the language. In contrast to T3T3, tone sandhi in T1T1 and T4T1 is applied only when the disyllabic sequence is right aligned regardless of whether the embedding trisyllabic constituent provides a simple or complex sandhi context. No iterativity of sandhi application was observed in any of these tonal sequences.

Table 3.14 *Summary of two types of tone sandhi in Tianjin Mandarin.*

	Sandhi Variation of Disyllabic Sandhi	Directionality of Trisyllabic Sandhi	
		Simple Contexts	Complex Contexts
T3T3	Near-Merger	From both directions	From the left edge
T1T1	No-Merger	From the right edge	
T4T1	No-Merger		

This study therefore provides a substantial amount of experimental data that greatly simplify the complex and paradoxical tone sandhi patterns reported in the literature (e.g., Li & Liu, 1985; Hung, 1987; Tan, 1987; Zhang, 1987; Chen, 2000; J. Wang, 2002; Ma, 2005; Wee et al., 2005; Zhang & Liu, 2011). We believe that our findings are not only of major empirical significance, but will also contribute to the further development of theories of tonal variation in general, as any good theory needs to be based upon solid empirical evidence. A cautionary note is that our data are representative of speakers born in the 1980s. It is possible that those tone sandhi patterns claimed in the literature can only be observed in older speakers, which then requires relevant experimental data in the future. If any generational difference is verified, it would be important to investigate the mechanisms that explain the changing sandhi patterns observed in old vs. young Tianjin Mandarin speakers as well as the implication of such tonal change mechanisms for language evolution in general.

CHAPTER 4 TONAL VARIABILITY ON THE PERCEPTION OF LEXICAL TONES - EVIDENCE FROM EYE MOVEMENTS[4]

4.1 Introduction

Among all languages in the world, 60-70% of them have lexical tones (Yip, 2002), where pitch changes over a word, together with other acoustic cues, signal lexical contrasts. In many tonal languages, tones are primarily signaled via different *f0* patterns (Gandour, 1978, but see e.g., Andruski, 2006; Brunelle, 2009b for the effect of phonation types on tonal contrasts). A handful of studies have revealed that native speakers of tone languages utilize various types of *f0* cues to identify lexical tones, such as the overall height of *f0* contour (e.g., Lin & Repp, 1989; Shen et al., 1993; Francis et al., 2003; Shen et al., 2013), the direction of *f0* change (e.g., Fox & Qi, 1990; Shen & Lin, 1991; Moore & Jongman, 1997), as well as the timing of *f0* turning point (e.g., Shen & Lin, 1991; Shen et al., 1993; Moore & Jongman, 1997). These findings show the importance of *f0* information in the perception of the canonical realization of lexical tones produced in isolation, but fall short in taking into consideration possible contextual *f0* variation in connected speech, the perception of which on tonal identification is the most practiced behavior during daily speech communication.

When lexical tones are produced in connected speech, the *f0* realization of the tones may deviate greatly from that produced in isolation, due to contextual effects such as tone sandhi and tonal coarticulation. Tone sandhi refers to phonological tonal alternation, which usually causes such a great change in the *f0* contour that the sandhi-derived tones are typically unpredictable or unexpected from the canonical *f0* realization. For instance in Beijing Mandarin, when two Low tones are combined in a disyllabic domain, the first Low tone, which otherwise would be realized with a low *f0* target, would surface with a rising *f0* contour (Chen, 2000 and references therein). Note that even for tones without tone sandhi allophonic variation, they are usually realized with an *f0* contour that is different from their canonical *f0* shape. The deviation, however, is much subtler and with varying degrees of phonetic modification depending upon the tonal contexts. They are nevertheless rather predictable from their canonical *f0* realization despite the phonetic modifications. This is known as tonal coarticulation (see Xu, 2001; Chen, 2012 and references therein for further details on tonal variation).

Despite the extensive contextual variation in the *f0* realization of lexical tones in connected speech and the rich layers of different types of tonal variability, only a small number of perception studies have taken the contextual effect into consideration. This body of work shows consistently that listeners' identification of a lexical tone is affected by its tonal context (e.g., Fox & Qi, 1990; Xu, 1994; Moore & Jongman, 1997; Francis et al, 2003). For example, Xu (1994) investigated the perception of the middle tone within

[4] A version of this chapter has been submitted for publication as: Qian Li & Yiya Chen. (submitted). Tonal variability on the perception of lexical tones - evidence from eye movements. *The Journal of the Acoustical Society of America*.

trisyllabic constituents. The target was embedded in both “compatible contexts” in which its *f0* realization is less deviated, and “conflicting contexts” where its *f0* realization is greatly deviated. Results show that the listeners’ identification of lexical tones produced in both contexts is highly accurate with their respective original contexts (i.e., target produced in compatible contexts was presented in compatible contexts, and target produced in conflicting contexts presented in conflicting contexts). However, if the original tonal contexts were replaced by white noise, only tones in compatible condition can be correctly identified above chance level, while the identification of tones in conflicting condition dropped below chance level. Furthermore, when the target was presented in swapped tonal contexts (i.e., target produced in compatible contexts presented in conflicting contexts, and that produced in conflicting contexts presented in compatible contexts), listeners showed clear tendencies of compensating the influence of the tonal contexts when identifying the target, e.g., a rising tone that was flattened by a conflicting context in production can be perceived as a falling tone when it was presented in a new compatible context. This study thus indicates that when tones are greatly deviated from its canonical *f0* realization, listeners need to rely on the contextual tonal information in order to identify the target lexical tones. The question that remains interesting is how the target tones are processed in advance of access to the tonal contexts during speech recognition.

A second, hitherto independent, line of studies on contextual tonal perception has mainly focused on the issue of whether the sandhi-derived *f0* realization of lexical tones is perceptually distinguishable from another lexical tone (which has been proposed to be the targeted sandhi tone in most literature on tone sandhi). For instance, a few studies have been done on the perception of the Low tone sandhi in Beijing Mandarin. Although there is evidence that the sandhi-derived Low tone variant is processed differently from the lexical Rising tone in production encoding (Chen et al., 2011; Zhang et al., 2014; Nixon, Chen, & Schiller, 2015), and also articulated with slight acoustic differences from the Rising tone (e.g., Yuan & Chen, 2014 and references therein), perception studies have repeatedly shown that the Low+Low tonal sequence cannot be reliably distinguished from that of the Rising+Low sequence (e.g., Wang & Li, 1967; Speer et al., 1989). However, the issue of tonal perception in sandhi context is not limited to whether the tone sandhi patterns are neutralized in perception. Moreover, the end-state responses are likely to exhibit a different pattern from that of the real-time processing data (Spivey, 2007; also see discussion in Malins & Joanisse, 2010). Therefore even if listeners cannot reliably distinguish two tonal sequences in traditional meta-linguistic perceptual judgement tasks, it is still possible that the two tonal sequences are actually processed differently in spoken word recognition. So, the question that remains to be addressed in the literature is how lexical tones are processed during spoken word recognition, due to their allophonic sandhi variation.

The present study aims to fill the knowledge gaps on tonal perception by addressing the above two research questions. Via a word-recognition task within the Visual World Paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), this study sets out to investigate the processing of tonal variability in Tianjin Mandarin, which presents a good testing case as it exhibits interesting patterns of tonal variability in connected speech.

4.1.1 Tianjin Mandarin

Tianjin Mandarin is a dialect of Mandarin which is spoken in the urban areas of Tianjin city (about 100km from 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 (Zhang & Liu, 2011; Li & Chen, 2012, 2016). In connected speech, Tianjin Mandarin shows a range of complex tone sandhi patterns over disyllabic constituents. The received wisdom has been that these sandhi patterns involve the categorical change of one lexical tone to another. For example, when two T3s are combined, the first T3 is claimed to change into T2 (e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007). However, recent experimental data have shown that there is no complete tonal neutralization in Tianjin Mandarin, which means there is no categorical change of one lexical tone to another (Zhang & Liu, 2011; Li & Chen, 2012, 2016). Furthermore, two different types of tone sandhi have been observed in this language, one is Near-Merger Sandhi and the other No-Merger Sandhi, as illustrated in Figure 4.1.

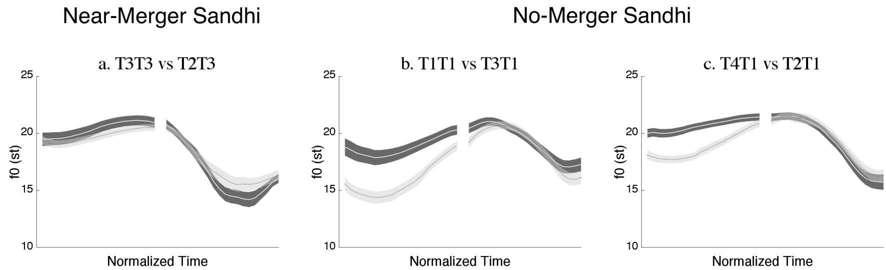

Figure 4.1 *f0 realization of three disyllabic sandhi sequences in Tianjin Mandarin (T3T3 in a, T1T1 in b and T4T1 in c) compared to their respective sandhi target sequences (T2T3 in a, T3T1 in b, T2T1 in c) as claimed in the literature. T3T3 is Near-Merger Sandhi; T1T1 and T4T1 are No-Merger Sandhi. Thick white lines indicate the mean f0 of the sandhi sequences (dark gray areas for ±1 standard error of mean); black lines indicate the mean f0 of the claimed target sequences (light gray areas for ±1 standard error of mean). Normalized time.*

In the Near-Merger Sandhi case, i.e., the T3T3 sequence in Figure 4.1*a*, the first T3 is realized with a high-rising *f0* contour, which makes the *f0* realization of this tonal sequence hardly distinguishable from that of the claimed target sequence (T2T3) in the literature. Figure 4.1*a* compares the *f0* contour of T3T3 to that of T2T3. The data were averaged across 72 tokens (six speakers, three items, two informational status, two repetitions) (Li & Chen, 2012, 2016). White lines with dark areas represent the *f0* realization of the sandhi pattern T3T3; and black lines with light areas represent the claimed target sequence T2T3. It can be seen from Figure 4.1*a* that, the *f0* realization of the two sequences are hardly distinguishable despite some subtle differences.

In the No-Merger Sandhi cases, the sandhi tones surface with altered *f0* realization, while maintaining its distinctiveness from the other lexical tonal contours. There are two No-Merger Sandhi cases in Tianjin Mandarin: T1T1 (Figure 4.1*b*) and T4T1 (Figure 4.1*c*).

T1T1 has been claimed to change into T3T1 in the literature (e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007). As shown in Figure 4.1*b*, the first T1 in T1T1 is realized with a slightly falling and rising *f0* tone, which is clearly unexpected from its canonical low-falling *f0* contour when produced in isolation. Although $T1_{sandhi}$ is realized similarly to T3 as in T3T1, its *f0* realization is significantly different from that of T3, suggesting no merging of $T1_{sandhi}$ with T3. Similarly, T4T1 is claimed to change into T2T1 (e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007). While the first $T4_{sandhi}$ in T4T1 is similar to the T2 *f0* contour, they are clearly different from each other (Figure 4.1*c*).

In short, the two types of tone sandhi differ in how closely the sandhi-derived *f0* contour resembles one of the lexical tones in the lexical tone system. In the Near-Merger case, the sandhi-derived tone is realized with an *f0* contour that is hardly distinguishable from its target lexical tone, while in the No-Merger cases, the sandhi-derived lexical tones show a similar *f0* contour to their corresponding target lexical tones while still maintaining their distinctiveness. The specific research question of interest here is whether and how the different types of tone sandhi variation affect tonal processing in online speech recognition?

4.1.2 Visual world paradigm and spoken word recognition

An auditory word-recognition task within the Visual World Paradigm (VWP) (Tanenhaus et al, 1995) was employed to tap into the processing of lexical tonal variation. A typical setup of the VWP in speech recognition studies is to present participants with four pictures or written words on a computer screen. The participants simultaneously hear an auditory stimulus, which corresponds to one of the four possibilities (i.e., the target) while their eye movement is tracked (see a review in Huettig, Rommers, & Meyer, 2011). Typically, a target is presented with a competitor and two distractors. The participants were asked to identify the words they have just heard and click on the target with a mouse.

Based on the assumption that eye movements are closely time-locked to the spoken-word processing, this paradigm makes it possible to tap into the time-course of auditory word recognition (e.g., Tanenhaus et al., 1995; Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Magnuson, & Tanenhaus, 2001; Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Beddor, McGowan, Boland, Coetzee, & Brasher, 2013; see also reviews in Tanenhaus, Magnuson, Dahan, & Chambers, 2000 and Huettig et al., 2011). Furthermore, the task involved in this paradigm is more natural speech perception experience than just asking the participants to make meta-linguistic judgements, which is typically used in traditional studies on tonal perception.

To make inferences about how the target auditory stimuli are processed, previous studies have mainly looked at how fast listeners start to fixate upon the target words (or competitors) as well as how long they look at the targets (or competitors). For example, in Tanenhaus et al. (1995), participants were asked to follow the instructions, e.g., "*pick up the candy*", and move real objects around. When the target object (e.g., *candy*) was presented together with another object (e.g., *candle*) whose name has phonological overlapping with the target, it took the participants longer time (i.e., 230ms) to initiate an eye movement to the target object *candy* than when a phonological similar object was absent (i.e., 145ms). This demonstrates that the speech recognition process actually starts before participants hearing the end of the word. In Allopenna et al. (1998), participants were presented with displays of four images on the computer screen, one for target corresponding to the auditory stimuli (e.g., *beaker*), one for "onset competitor" which has onset overlapping with the target (e.g., *beetle*, which shares the onset with *beaker*), as well as two for distractors that are phonologically unrelated with the target (e.g., *dolphin*, *carriage*). The results show that when participants heard the instruction "*pick up the beaker*", the proportion of looks to both *beaker* and *beetle* started to increase initially. However, as the speech unfolded, the proportion of looks to *beetle* started to decrease while only the picture corresponding to target gained more looks. In another experimental condition, the target words were presented with a "rhyme competitor" (e.g., *speaker*, which shares the rhyme with the target *beaker*). Compared to that of the "onset competitor", the initial increase of looks to "rhyme competitor" *speaker* occurred at a much later time point since it only overlapped with the target at a later part of the word. This indicates that, upon hearing a spoken word, listeners continually evaluate the unfolding speech with fine-grained sensitivity, and gradually activate certain lexical candidates that compete for recognition.

Recently, VWP has been successfully employed in perception studies of lexical tones produced in isolation (Malins & Joanisse, 2010; Shen et al., 2013). For example, in Shen et al. (2013), participants were presented with images corresponding to segmentally identical T2 and T3 monosyllabic words (e.g., bi^2 'nose' vs. bi^3 'pen') together with two other phonologically unrelated distractors (e.g., che^1 'car', dao^1 'knife'). They were played with T2 or T3 stimuli and asked to click on the corresponding images. Stimuli were normalized with 500ms duration. Both T2 and T3 were resynthesized so as to have identical *f0* contours from the onset until 200ms into the tones; from 200ms on, T2 was resynthesized with high *f0* offsets and T3 low offsets. Results show that like segments, the recognition of lexical tones starts before the end of the stimuli and the processing of lexical tones is in a similar incremental fashion as that observed for segments. For example, upon hearing a T3 stimulus, participants generally needed only around 350-450ms to initiate an eye movement to the image corresponding to the target. Given that it usually takes about 200ms to plan and execute an eye movement (Matin, Shao, & Boff, 1993), it can be inferred that participants have already recognized the tone as early as around 150-250ms after the stimulus onset, which is much earlier than the end of the stimulus (500ms after stimulus onset). Furthermore, as the time unfolds, the proportion of looks to T3 continues to increase from 30% up to 90%, while that for other possibilities drops to the minimum at the same time.

This present study extended the VWP into studies of lexical tone perception in connected speech, where disyllabic collocations in Tianjin Mandarin with different types of tonal variability were used as auditory stimuli (i.e., Near-Merger Sandhi, No-Merger Sandhi, No Sandhi). The first goal of this study is to establish whether lexical tones produced in contexts are comparable with those produced in isolation in terms of the eye movement patterns during auditory speech recognition; second, it is also of great interest to investigate how different types of tonal variability affect online tonal processing. Note that the present study is not particularly interested in whether certain sandhi sequence is ambiguous with its target sequence as claimed in the literature with segmental overlap for both syllables, e.g., *zhi³ fa³* 'fingering' vs. *zhi² fa³* 'executing justice'. Rather, the main focus is only on the perception of the first tone of different tonal variability, where there should be no ambiguity in the visual world context, e.g., *zhi³ fa³* 'fingering' vs. *zhi⁴ hou⁴* 'lag'.

To address the first issue, two competitors for each target were included: a baseline competitor which does not share segments or tones with the target, and a segmental competitor which has overlapping segments but unrelated tones with the target (following the name in Malins & Joanisse, 2010 for the ease of comparison). Malins and Joanisse (2010) show that when targets and competitors are segmentally identical but different in tones (e.g., *chuang²* 'bed' vs. *chuang¹* 'window'), there is greater competition between targets and competitors than in the baseline condition with no segmental overlap. Similarly, greater competition effect for segmental competitors than for baseline competitors should also be expected in the present study.

To achieve the second goal, three types of tonal variability were included for the target words, Near-Merger Sandhi, No-Merger Sandhi, and No Sandhi. It can be predicted that 1) for Near-Merger Sandhi, the first tone is processed with the most effort due to the altered *f0* realization and the potential ambiguity with another lexical tone in the tonal system; 2) No-Merger Sandhi should be less difficult to process than Near-Merger Sandhi, since its realization is altered but it is not supposed to be ambiguous with another tone in the tonal system; and 3) in the No Sandhi condition, the tonal realization of the first tone can be expected from the canonical realization, and thus no tonal ambiguity is involved; therefore, it should be processed with the least effort.

4.2 Method

4.2.1 Participants

Thirty-four native speakers of Tianjin Mandarin were recruited to participate in this experiment. All subjects were born in late 1980s or early 1990s (9 males, 25 females; Mean=22) 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. One subject was wearing cosmetic contact lens, resulting in potentially different eye movement data. Her data were excluded from further analyses. Data from another two subjects who were shortsighted without wearing glasses for correction were also excluded. The remaining 31 subjects had normal or corrected-to-normal vision. All participants provided written informed consent.

4.2.2 Stimuli

The target stimuli consisted of 36 highly lexicalized disyllabic collocations with two sandhi patterns in Tianjin Mandarin: near-merger sandhi (i.e., T3T3) and no-merger sandhi (i.e., T1T1, T4T1). In addition, 18 stimuli with only tonal coarticulation and no sandhi changes were included for further comparison (i.e., T4T4, T3T2, T3T4) (see Table 4.1). Target stimuli of different tonal variability types were matched in bigram mutual information which stands for the likelihood of two syllables co-occuring within a lexical item according to Da (2004). The mean bigram co-occuring frequency of all target stimuli was 5.0 (SD=3.7), which corresponds to strong collocation strength, and the three groups of stimuli were not significantly different from each other ($F(2)=0.26$, $p=0.773$).

Table 4.1 *Experimental design and sample stimuli.*

Near-Merger Sandhi			
Target	Baseline Competitor	Distractor 1	Distractor 2
雨水 (yu³ shui³) *rainwater*	拖鞋 (tuo¹ xie²) *slippers*	抽血 (chou¹ xie³) *to draw blood*	犀牛 (xi¹ niu²) *rhinoceros*
	Segmental Competitor	Distractor 1	Distractor 2
	预购 (yu⁴ gou⁴) *purchase in advance*	分号 (fen¹ hao⁴) *semicolon*	报警 (bao⁴ jing³) *to report to the police*
No-Merger Sandhi			
Target	Baseline Competitor	Distractor 1	Distractor 2
浴缸 (yu4 gang1) *bathtub*	冷气 (leng³ qi⁴) *cold air*	消防 (xiao¹ fang²) *fire-fighting*	激动 (ji¹ dong⁴) *excited*
	Segmental Competitor	Distractor 1	Distractor 2
	语录 (yu³ lu⁴) *quotation*	沙拉 (sha¹ la¹) *salad*	冬天 (dong¹ tian¹) *winter*
No Sandhi			
Target	Baseline Competitor	Distractor 1	Distractor 2
继续 (ji⁴ xu⁴) *to continue*	缆车 (lan³ che¹) *cable car*	雨衣 (yu³ yi¹) *raincoat*	手镯 (shou³ zhuo²) *bracelet*
	Segmental Competitor	Distractor 1	Distractor 2
	挤压 (ji³ ya¹) *to press*	椭圆 (tuo³ yuan²) *oval*	海港 (hai³ gang³) *harbor*

For each target, a corresponding baseline competitor and a segmental competitor were chosen (a within-item design). The baseline competitors did not share segment or tone with the targets (S-, T-). The segmental competitors shared the segment of the first syllable with the targets, but with an unrelated tone (S+, T-), which was neither the underlying lexical tone nor the lexical tone which was claimed in the literature as the targeted sandhi tone. The second syllables of the targets and competitors were different in terms of both tone and segment across all conditions. Take the target *zhi³ fa³* ('fingering') as an example: its baseline competitor was *wai⁴ tao⁴* ('coat') while the segmental competitor was "*zhi⁴ hou⁴*" ('lag'). Each target-competitor pair also had two distractors within the Visual-World Paradigm. The distractors did not share any tone or segments with the target or the competitor.

The targets and competitors were further controlled to be closely matched in terms of lexical frequency based on Cai and Brysbaert (2010) and orthographic complexity as we presented Chinese characters instead of pictures (following Shen et al., 2013), so that there was no significant difference between the targets and competitors for lexical frequency ($F(1)=0.087$, $p=0.768$) or visual complexity ($F(1)=0.156$, $p=0.694$). No participant reported difficulty of recognizing any character.

The auditory stimuli were pre-recorded by a male speaker of Tianjin Mandarin who was born in the 1980s. All stimuli were produced with the same loudness and speaking rate. The mean duration of the first syllable of the stimuli was 357ms (SD=71ms), and was not significantly different across the three conditions ($F(2)=0.47$, $p=0.627$); the mean duration of the entire disyllabic stimuli was 704ms (SD=98ms).

4.2.3 Procedure

Eye movements were recorded with an Eyelink 1000 system with a 35mm lens running at 500Hz. Visual stimuli were presented on a 21-inch ViewSonic G220f monitor (resolution: 1024*768 pixels; frame rate: 120Hz). Participants were seated comfortably with a chin rest and a forehand rest set at a distance of 69cm from the screen. All recordings and calibrations were done monocularly based on the left eyes and viewing was binocular (except for one participant whose right eye was tracked for tracking failure of the left eye). Gaze position was calibrated with a 13-point grid. Prior to the presentation of each trial, there was a drift check.

The experiment began with a block of seven training trials to familiarize the participants with the experimental setup and task. Participants were tested individually. Each trial consisted of a fixation-cross screen (500ms) with only a fixation cross at the very center of the screen *prior to* the stimuli screen. The participants were asked to look at the fixation cross until it disappeared. Then the auditory stimulus was played through a headphone simultaneously with the presentation of the visual stimuli screen. The task for the participants was to click on the lexical item they had heard with a mouse within 10 seconds (although no participant spent more than 10 seconds to make a decision).

In each trial, participants saw four disyllabic collocations on the screen, which consisted of a target corresponding to the auditory stimulus, a competitor, and two distractors. To maximally guarantee the non-overlapping of the parafoveal view when reading each collocation (e.g., Mielle et al., 2009), size and location of the visual stimuli were calculated accordingly to the screen (2.5cm * 5cm in height and width, corresponding to 5° visual angle; approximately at the centers of the four quadrants of the screen; for further details, we refer readers to Appendix I). The order of the four visual stimuli on the screen was counter-balanced. The trial order was pseudo-randomized so that auditory stimuli of the same tonal variability condition were not presented in consecutive trials.

4.2.4 Eye movement data analysis

Given that the two characters of a disyllabic collocation on the screen basically covered the foveal area (see Figure I in Appendix I), only eye gaze data within the boundary immediately around the collocation were included for the analyses of respective stimulus item (i.e., target, competitor and distractors). The eye movement data were recorded at 2ms intervals. The proportion of looks at each time point (every 2ms) was calculated. Trials in which subjects clicked on items other than the target (<1%) were excluded from the analyses. For better illustration of the eye movement patterns, data from trials containing blinks (17%) were also excluded.

The eye movement data were reported from the onset of the auditory stimulus to 1400ms post stimulus onset. The upper limit was chosen at the point where the proportion of looks to target had reached the maximum (following Malins & Joanisse, 2010).

The growth curve analysis was used for statistics (following Mirman, 2014) with the package *lme4* (Bates et al., 2014) in R (R Core Team, 2014). Given the large dataset obtained for this study, the statistical analysis was limited within certain time windows (i.e., part of the curves), which showed interesting patterns, instead of fitting the entire curves.

The base model was first constructed with only TIME (up to the fourth order components) in the fixed-factor structure and SUBJECT as the random factor on all time terms (following Mirman, 2010). Additional fixed factors were then assessed in a stepwise fashion. Only factors that significantly improved the model fits were added into the model. Parameter-specific *p*-values were estimated using the normal approximation (treating the *t*-value as a *z*-value).

4.3 Results

4.3.1 Baseline comparison

Figure 4.2 illustrates the mean proportion of looks to target averaged across three types of tonal variability, when the competitor is a segmental competitor (i.e., with segmental overlap) vs. when there is a baseline competitor (i.e., without segmental overlap). X-axis stands for the time since target onset and y-axis for the proportion of looks to target.

As can be seen from Figure 4.2, when there is segment overlap between the targets and the competitors (shown with solid line as SEGMENTAL), the proportion of looks to target shows a different pattern from that in the baseline competitor condition (shown with dotted line as BASELINE) mainly in two aspects. First, the proportion of looks to target in the segmental condition starts to increase at a much later time point (around 500ms) than that in the baseline condition (around 200ms). Second, the segmental condition shows lower overall proportion of looks to target than in the baseline condition, especially salient in the time window of 200-800ms, which approximately corresponds to about 600ms into the auditory stimuli (the first syllable is 357ms on the average). After 800ms, the differences between the two conditions are less dramatic.

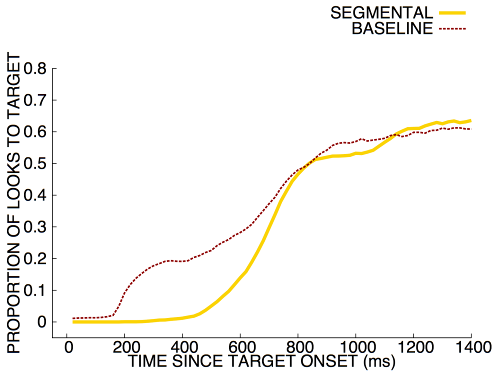

Figure 4.2 *The proportion of looks to target over time when target and competitor have segment sharing for the first syllable (SEGMENTAL) compared to the baseline control condition (BASELINE) averaged across three tonal variability types.*

A growth curve analysis was run for the comparison between segmental vs. baseline conditions within the 200-800ms time window. The full model included the interaction of the fourth-order time terms, COMPETITOR TYPE (two levels: SEGMENTAL & BASELINE) and TONAL VARIABILITY TYPE (three levels: NEAR-MERGER SANDHI, NO-MERGER SANDHI and NO SANDHI) as the fixed factor and SUBJECT as a random factor on all time terms. We also included FREQUENCY, NUMBER OF STROKES, BIGRAM MUTUAL INFORMATION and STRUCTURE of the target stimuli as additional fixed factors in the full model. The results confirmed a significant main effect of COMPETITOR TYPE over the time terms up to the third-order within the 200-800ms time window (intercept: $\beta=0.15$, $t=33.68$, $p<0.001$; slope: $\beta=-0.18$, $t=-7.39$, $p<0.001$; quadratic: $\beta=-0.13$, $t=-5.34$, $p<0.001$; cubic: $\beta=0.09$, $t=3.86$, $p<0.001$). This indicates that the proportions of looks to the target in two conditions were significantly different in the overall mean, the direction of curve change, as well as the steepness of the curve change.

4.3.2 Different tonal variability types

To further investigate the eye movement patterns in three types of tonal variability, we included analyses for looks to both targets and competitors in this section.

4.3.2.1 Looks to the target

Figure 4.3 shows the proportion of looks to target of three different tonal variability types when target and competitor shared segments. X-axis stands for the time since target onset, y-axis for the proportion of looks to target.

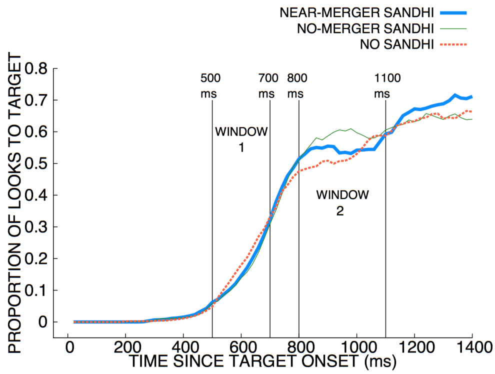

Figure 4.3 *The proportion of looks to target when targets and competitors have segmental overlap in three tonal variability types, aggregated across participants and items (thick solid line for Near-Merger Sandhi, thin solid line for No-Merger Sandhi, dotted line for No Sandhi).*

We can see from Figure 4.3 that, due to different types of tonal variability in the targets, there are subtle differences in the proportion of looks to the target. In general, the proportion of looks to target in all tonal variability types remains at the bottom (less than 0.1) until around 500ms, when the proportion starts to rise. This indicates that the participants were not looking at the targets until 500ms across all tonal variability types. However, after 500ms, the proportion of looks to target in different tonal variability types began to diverge, especially in two time windows: 500-700ms and 800-1100ms.

In the 500-700ms time window (WINDOW 1, Figure 4.3), the No-Sandhi condition shows a slightly faster increase in the proportion of looks to target than in the other two sandhi conditions. A growth curve analysis was run for the comparison among three tonal variability types within this window. The full model included the interaction of the fourth-order time terms and TONAL VARIABILITY TYPE (three levels: NEAR-MERGER SANDHI, NO-MERGER SANDHI and NO SANDHI) as the fixed factor and SUBJECT as the random factor on all time terms. Additional fixed factors included FREQUENCY, NUMBER OF STROKE, STRUCTURE and BIGRAM MUTUAL INFORMATION of the target, as well as NUMBER OF STROKE of the competitor. Although the difference is subtle, the growth curve analysis results suggested that the proportion of looks to the target in No-Sandhi condition had a significantly faster increasing rate than the other two conditions, in terms of the quadratic components (quadratic for Near-Merger Sandhi vs. No Sandhi: $\beta=0.05$, $t=1.99$, $p<0.05$; No-Merger Sandhi vs. No Sandhi: $\beta=0.05$, $t=2.04$, $p=0.04$). Two sandhi conditions were not significantly different from each other in any aspect within this time window.

The second interesting time window is from 800ms to 1100ms (WINDOW 2, Figure 4.3). Within this window, our full model of growth curve analysis included the interaction of the fourth-order time terms and TONAL VARIABILITY TYPE (three levels: NEAR-MERGER SANDHI, NO-MERGER SANDHI and NO SANDHI) as the fixed factor and SUBJECT as the random factor on all time terms. Additional fixed factors included

FREQUENCY, STRUCTURE and BIGRAM MUTUAL INFORMATION of the target stimuli, as well as STRUCTURE of the competitor. In this window, there are noticeable differences in the three conditions, among which two observations can be made.

First, No-Sandhi condition differs from the two sandhi conditions in terms of both the overall mean and the slope of proportion of looks at the target. There is clearly less proportion of looks at the target in the No-Sandhi condition, as suggested by the significantly results in the intercepts for Near-Merger Sandhi vs. No-Sandhi condition (intercept: $\beta=0.05$, $t=5.75$, $p<0.001$) and No-Merger Sandhi vs. No-Sandhi (intercept: $\beta=0.08$, $t=8.50$, $p<0.001$). In addition, the No-Sandhi condition had a significantly larger rising slope of the proportion of looks to the target than both Near-Merger Sandhi (slope for Near-Merger Sandhi vs. No-Sandhi: $\beta=-0.10$, $t=-3.12$, $p<0.01$) and No-Merger Sandhi (slope for No-Merger Sandhi vs. No-Sandhi: $\beta=-0.10$, $t=-3.29$, $p<0.001$). The slopes of two sandhi conditions did not significantly differ from each other.

Second, the two sandhi conditions showed significantly different proportions of looks to the target in terms of the overall mean. There are overall less looks to the target in Near-Merger Sandhi than that in the No-Merger Sandhi condition, as reflected in the significant results in the intercept between Near-Merger Sandhi vs. No-Merger Sandhi ($\beta=0.03$, $t=3.16$, $p<0.001$).

4.3.2.2 Looks to the competitor

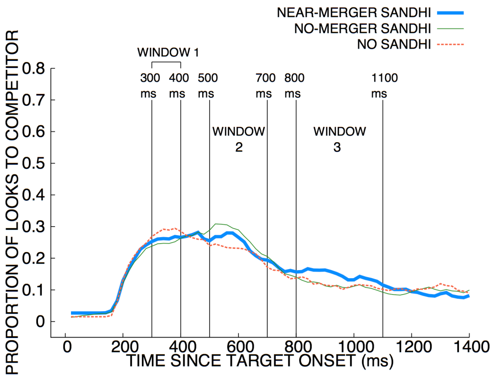

Figure 4.4 *The proportion of looks to competitor when target and competitor have segment overlap in three tonal variability types, aggregated across participants and items (thick solid line for Near-Merger sandhi, thin solid line for No-Merger sandhi, dotted line for No Sandhi).*

If not at the target, where else could the listeners be looking? We further analyzed the eye movement patterns over competitor. Figure 4.4 illustrates the proportion of looks to competitor in three tone sandhi types. X-axis stands for the time since target onset, y-axis for the proportion of looks to competitor. The time frames in WINDOWS 2 and 3 in Figure 4.4 correspond to that of WINDOW 1 and WINDOW 2 in Figure 4.3, respectively.

In Figure 4.4, the proportion of looks to competitor in all tonal variability types remains at the bottom, starting to rise and diverges after 200ms. Three time windows are particularly interesting: 300-400ms (WINDOW 1, Figure 4.4), 500-700ms (WINDOW 2, Figure 4.4) and 800-1100ms (WINDOW 3, Figure 4.4).

In the window of 300-400ms (WINDOW 1, Figure 4.4), the No-Sandhi condition shows relatively higher proportion of looks to the competitor than two sandhi conditions with a slight magnitude of difference. Taking a look at the whole curve for No-Sandhi condition in this figure, we could also see that the peak value for the No-Sandhi condition appears within the 300-400ms window. Growth curve analyses confirmed a significant difference in the overall mean between the Near-Merger Sandhi vs. No-Sandhi conditions (intercept: $\beta=-0.07$, $t=-5.68$, $p<0.001$) and between the No-Merger Sandhi vs. No-Sandhi conditions (intercept: $\beta=-0.06$, $t=-5.04$, $p<0.001$). Two sandhi conditions were not significantly different from each other.

In the time window from 500ms to 700ms (WINDOW 2, Figure 4.4), the proportion of looks to the competitor in the No-Merger Sandhi condition has a relatively higher mean than that in the other two conditions. This was confirmed by the significant difference in the overall mean between No-Merger Sandhi vs. No-Sandhi conditions (intercept: $\beta=0.04$, $t=4.57$, $p<0.001$) and No-Merger Sandhi vs. Near-Merger Sandhi conditions (intercept: $\beta=0.03$, $t=3.13$, $p<0.01$).

In the window from 800ms to 1100ms (WINDOW 3, Figure 4.4), the Near-Merger Sandhi condition shows relatively more overall proportions of looks to the competitor than both No-Merger Sandhi and No-Sandhi conditions. This could be confirmed by the significant difference in the intercept of the No-Merger Sandhi condition (intercept: $\beta=-0.03$, $t=-5.27$, $p<0.001$) and No-Sandhi condition (intercept: $\beta=-0.03$, $t=-4.99$, $p<0.001$), while No-Merger Sandhi and No-Sandhi were not significantly different from each other.

4.4 Discussion & conclusion

While it is well established that there is extensive variability in the *f0* realization of lexical tones due to the neighboring tones, it is less understood in how native listeners perceive contextual tonal variations in spoken word recognition. The present study set out to shed some light on this issue with evidence from Tianjin Mandarin, which exhibits rich layers of contextual tonal variability.

We employed the Visual World Paradigm to investigate the time course of disyllabic tone perception with evidence from Tianjin Mandarin. We examined the participants' looks to both targets and competitors within the Visual World Paradigm when they heard disyllabic stimuli with different types of tonal variability (i.e., Near-Merger Sandhi, No-Merger Sandhi, No Sandhi). Our results yield significant perceptual differences among different types of tone variation, which affects online speech processing differently, as reflected in the different eye movement patterns.

To first ensure the participants were behaving in a comparable way with what have been reported in studies of lexical tones produced in isolation, we included the baseline comparison where the targets and competitors had neither segmental nor tonal overlap. We found that when targets and competitors share the same segment for the first syllable, there was a generally smaller proportion of looks to the target than that in the baseline condition. This indicates a stronger overall competition between targets and competitors when there is segmental overlap, which is comparable with what have been reported in a recent perception study on monosyllabic Mandarin tones (Malins & Joanisse, 2010). What differs between ours and the earlier study is that we found much earlier divergence of proportion of looks to target: only around 200ms post target onset in ours, but around 600ms in Malins and Joanisse (2010). This is probably due to the fact that we used printed words as the visual stimuli rather than images as used in Malins and Joanisse (2010). Printed words have been argued to be more sensitive to phonological manipulations than images, at least for alphabetic languages (e.g., Weber, Melinger, & Tapia, 2007; Huettig et al., 2011). Shen et al. (2013) compared image vs. character stimuli, but claimed no difference. However, their experimental setups for two types of stimuli were not completely comparable for comparison. Future studies are therefore needed to verify the potentially different effects of two types of visual stimuli on eye movement patterns in logographic languages during online speech recognition.

To look into how different kinds of contextual tonal variability affect tonal perception, we compared the proportions of looks to both targets and competitors in three different tonal variability conditions. As predicted, we have observed significant differences among the three different types of tonal variability.

First, we have observed a different eye movement pattern upon hearing target stimuli with No-Sandhi tonal coarticulation from those in two sandhi conditions, with evidence from the gaze patterns over both targets and competitors. In terms of the proportion of looks to the target, No-Sandhi condition showed quicker increase of looks to the target compared to the two sandhi conditions in the time frame from 500ms to 700ms (WINDOW 1, Figure 4.3). Later on, participants did not seem to look at the No-Sandhi targets any more than they did in the two sandhi conditions in the time window from 800ms to 1100ms (WINDOW 2, Figure 4.3). Data of proportion of looks to the competitor suggested that the participants have shifted their attention from both targets and competitors to somewhere else at this time point, as in the time window from 800ms to 1100ms, the proportion of looks to the competitor in No-Sandhi condition was also lower than sandhi conditions (WINDOW 3, Figure 4.4). This, together with the finding that there are significantly more looks to competitors in the No-Sandhi condition than the two sandhi conditions at an earlier time point (300-400ms) (WINDOW 1, Figure 4.4), confirmed the relatively less effort required in processing No-Sandhi coarticulation, as the tones are realized with only subtle phonetic deviation.

Second, compared to the No-Sandhi coarticulation condition, two sandhi conditions have presented more challenge to listeners in tonal recognition due to the more dramatic *f0* contour distortion. This was reflected in later initiation of fixation on the targets in both sandhi conditions (WINDOW 1, Figure 4.3). Furthermore, two sandhi conditions showed different eye movement patterns from each other for both targets and competitors.

Between two sandhi conditions, No-Merger Sandhi seemed to be relatively easier to identify, as confirmed by the overall more looks to the target than Near-Merger Sandhi in the time window 800ms to 1100ms (WINDOW 2, Figure 4.3). The analyses of competitors in Figure 4.4 further suggested that No-Merger Sandhi already showed more proportion of looks to the competitors than that of Near-Merger Sandhi at an earlier time point from 500ms to 700ms (WINDOW 2, Figure 4.4), while in the time window from 800ms to 1100ms (WINDOW 3, Figure 4.4), looks to competitors for No-Merger Sandhi significantly declined to the same low proportion as for the No-Sandhi condition. This suggested the successful recognition of No-Merger Sandhi stimuli at this time point, as participants have shifted their attention away from both target and competitor.

Compared to No-Merger Sandhi, Near-Merger Sandhi seems to be more difficult to process, reflected in significantly less looks at the target in the 800-1100 time window than No-Merger Sandhi. The significantly higher proportion of looks to the competitor for Near-Merger Sandhi in WINDOW 3 of Figure 4.4 further suggested that, within the time window from 800ms to 1100ms, the Near-Merger Sandhi stimuli have not been completely recognized as participants were still searching for the target.

Previous studies have shown clear effect of tonal contexts on lexical tone identification (e.g., Lin & Wang, 1985; Fox & Qi, 1990; Moore & Jongman, 1997; Francis et al., 2003), especially when tones largely deviate from its canonical realization, where listeners need to rely on the contexts to identify the tones (Xu, 1994). On the one hand, our study has lend further support to this contextual effect; on the other hand, by extending this body of literature, our results show, for the first time, that tonal variation within the first syllable has already exerted an effect on tonal recognition, given comparable contextual facilitation from the second syllable for tonal identification.

For the No-Sandhi condition, where there is only minor variation while tonal distinctiveness is maintained, listeners relied relatively less on the contexts in recognizing the tone of the first syllable as reflected in the relatively faster growth in the proportion of looks to the target. For two sandhi conditions, the *f0* realization of the lexical tones was altered to such a great extent that the tonal distinctiveness was no longer kept. It would be difficult to identify the tone only based on the sandhi-derived *f0* realization (also as suggested in Zhou & Marslen-Wilson, 1997). Information of the second syllable was thus needed to a greater extent to identify the lexical tone of the first syllable, which was reflected in the much later looks to target in both conditions.

This thus has provided clear perceptual evidence in support of the distinction between two contextual tonal variation processes, i.e., tone sandhi vs. tonal coarticulation. Although there is some consensus on different *f0* variation triggered by tone sandhi vs. tonal coarticulation, previous impressionistic studies claim no essential distinction between the two processes (e.g., Chen, 2000). Our data showed that the different degrees of *f0* deviation due to tone sandhi and tonal coarticulation indeed have different perceptual consequences in the process of online speech recognition. This thus suggests the necessity to tease apart the two processes with care when one attempts to analyze and account for the phenomenon of contextual tonal variability.

Our results further suggested the need to differentiate two types of tone sandhi in Tianjin Mandarin. These different types of tonal variability over the first syllable clearly have different consequences on tonal perception; because otherwise, different eye movement patterns for different tone sandhi types could not have been observed. For No-

Merger Sandhi where sandhi tones are rather distinct from the canonical tonal realization of any lexical tone within the lexicon, the difficulty in identification is mainly due to the recognition of a distorted *f0*. Near-Merger Sandhi, comparatively, is more difficult to identify, given that the sandhi-derived tones are indeed quite undistinguishable from another lexical tone in the system. The successful identification thus involves both recognition of a distorted *f0* and the competition between two lexical tone candidates, which therefore requires more effort than for No-Merger Sandhi cases. While traditional perceptual studies on contextual tone sandhi variation mainly rely on native listeners' meta-linguistic knowledge of whether one tone is changed into another due to tone sandhi (e.g., Wang & Li, 1967; Speer et al., 1989), our paper is the first to look into how lexical tones with different types of tone sandhi are processed online.

Taking together, we have demonstrated that despite the commonly accepted "categorical perception" of lexical tones (e.g., Chan, Chuang, & Wang, 1975; Burnham & Jones, 2002; Francis et al., 2003; Xu, Gandour, & Francis, 2006; Xi, Zhang, Shu, Zhang, & Li, 2010; Zhang, Xi, Wu, Shu, & Li, 2012; but see Abramson, 1977 for non-categorical perception in Thai), native listeners are quite sensitive to fine-grained *f0* details in tone perception, which is in line with Malins and Joanisse (2010) and Shen et al. (2013). We observed a graded time course difference in online perception of contextual tonal variants, depending on how much the variants deviate from the canonical *f0* realization and whether there is competition with another lexical tone within the tonal inventory. Specifically, No-Sandhi tonal coarticulation was the easiest to recognize as the proportion of participants' looks to the target increased with the fastest rate among three conditions; Near-Merger Sandhi was more difficult to process than No-Merger sandhi, reflected in the slower increase of looks to the target as well as in the overall less proportion of looks to target. The further implications of our findings on possible linguistic theory of tonal variability and sandhi alternation needs to be explored in the future.

CHAPTER 5 THE ONLINE PERCEPTION OF SANDHI TONES IN TIANJIN MANDARIN – EVIDENCE FROM EYE MOVEMENTS

5.1 Introduction

In tonal languages that utilize melodic pitch contour changes to convey lexical meanings such as Mandarin Chinese, the main acoustic correlate of lexical tones - fundamental frequency (*f0*) changes - varies extensively in connected speech. Part of the variation can be due to the global prosody of the entire utterance (e.g., Xu, 1999; Yuan, 2004; Chen, 2003; Chen & Gussenhoven, 2008; Chen, 2010; Scholz & Chen, 2014a, 2014b; also see reviews in Chen, 2012; Zsiga, 2012). Others are more localized where the *f0* realization of a lexical tone varies as a function of its neighboring tonal contexts. One such contextual tonal alternation process is known as "tone sandhi", commonly regarded as the phonologized change of lexical tones in certain tonal contexts. Tone sandhi usually introduces great changes in the *f0* contours of the sandhi-derived tones (see reviews in Chen, 2000; Zhang, 2010). A well-known case of tone sandhi is the "Third-Tone" sandhi in Standard Chinese. When two dipping tones (Tone 3, hereafter as T3) are combined, the first one surfaces with a rising *f0* contour (hereafter $T3_{sandhi}$), similar to that of the lexical rising tone (Tone 2, hereafter as T2) in the language.

A much-debated issue that is crucial for tonal alternation theories is whether tone sandhi such as the "Third-Tone" sandhi in Standard Chinese involves complete neutralization of otherwise distinctive lexical tones as a result of the tonal sandhi change. Previous perception studies have shown that native listeners cannot reliably distinguish the sandhi variant $T3_{sandhi}$ vs. the lexical T2 (e.g., Wang & Li, 1967; Speer et al., 1989; Chen, 2013; Chen et al., 2015). This leads to the received view of T3 sandhi change in Standard Chinese as the change of T3 to T2, in which $T3_{sandhi}$, as a result, is regarded to completely neutralize with the lexical T2 (e.g., Chen, 2000; Yip, 2002; Chen et al., 2015). This perception-based point of view, however, is challenged by evidence from tone production studies which demonstrate subtle but quite consistent *f0* differences between $T3_{sandhi}$ and the lexical T2 in both read speech (e.g., Zee, 1980; Peng, 1996) and corpus data (Yuan & Chen, 2014). Furthermore, $T3_{sandhi}$ and the lexical T2 have also been shown to be processed differently during speech encoding (Chen et al., 2011; Zhang et al., 2014). If complete neutralization was indeed involved in tone sandhi, such systematic production differences between the sandhi-derived tone (i.e., $T3_{sandhi}$) and the claimed output tone (i.e., T2) should not have been expected. However, it is still unknown so far whether the sandhi-derived tones are processed more similarly to its non-sandhi variants or to the claimed output tones.

In a study related to this issue, Zhou and Marslen-Willson (1997) investigated the representation of tonal sandhi variants in Standard Chinese via two auditory-auditory priming experiments, based on the assumption that T3 changes to T2 before another T3. Results of the first experiment show that, compared to the control condition where primes and targets have unrelated tones and segments, the processing of disyllabic $T3_{sandhi}T3$

targets were significantly facilitated by a T3Tx prime, where $T3_{sandhi}$ in the target and T3 in the prime share the toneme but differ in the surface *f0* contour (Toneme+, Contour-). However, the targets were significantly inhibited by a T2Tx prime which surfaces with a similar *f0* with $T3_{sandhi}$ in the targets but has a different toneme (Toneme-, Contour+). The first experiment thus suggests that $T3_{sandhi}$ should be represented similarly as the non-sandhi T3 rather than the lexical T2 as hypothesized by the authors (following the literature). In the following experiment with T2Tx as targets and T3Tx, $T3_{sandhi}$T3 and T2T3 as auditory primes, their results show that compared to the control condition, T3Tx (Toneme-, Contour-) showed a significantly more inhibitory effect than $T3_{sandhi}$T3 (Toneme-, Contour+) and T2T3 (Toneme+, Contour+), suggesting similarity between T2 with $T3_{sandhi}$ but not T3. Assuming that T3 changes to T2 in T3T3 sequences (i.e., $T3_{sandhi}$=T2), we find the results reported here rather puzzling. The relationship between $T3_{sandhi}$ and T2 (or T3) thus is worthy of further research efforts to be clarified.

It is to be noted that while the *f0* information of lexical tones has been shown to be processed in an incremental fashion as segments in online speech recognition (Malins & Joanisse, 2010; Shen et al., 2013; Li & Chen, 2015), most previous studies on tone sandhi perception have focused on whether native listeners are able to discriminate between the sandhi tones (e.g., $T3_{sandhi}$ in Standard Chinese) and the so-called output tones (e.g., lexical T2 in Standard Chinese) in traditional meta-linguistic tone discrimination paradigms (e.g., Wang & Li, 1967; Speer et al., 1989). Zhou and Marslen-Wilson (1997) used an online auditory-auditory priming lexical decision paradigm, but only end-state responses were recorded. On the one hand, data of only end-state responses are inadequate to reveal the dynamic time-course differences in spoken word recognition prior to the end-state judgement (see further discussion in Malins & Joanisse, 2010). On the other hand, the end-state responses are likely to exhibit a different pattern from that of the real-time processing data, which could potentially lead to completely different understandings of the same phenomenon (Spivey, 2007). Further online speech recognition studies are therefore in great need to shed more light on how listeners recognize sandhi tones in real time.

To tap into this issue, the present study investigates the time course of online recognition of sandhi tones in Tianjin Mandarin. As will be introduced below, Tianjin Mandarin serves as an interesting test case here for its rich layers of contextual tonal sandhi variation patterns over disyllabic tonal sequences (e.g., Li & Liu, 1985; Chen, 2000; Wee et al., 2005; Zhang & Liu, 2011; Li & Chen, 2016).

5.2 Tianjin Mandarin

Tianjin Mandarin is a dialect of Mandarin Chinese, which is mainly spoken in the urban areas of Tianjin Metropolis, China. There are four lexical tones in Tianjin Mandarin: Tone 1 (T1) is a low-falling tone, Tone 2 (T2) a high-rising tone, Tone 3 (T3) a dipping tone and Tone 4 (T4) a high-falling tone (Li & Chen, 2016). Figure 5.1 illustrates the *f0* realization of the four lexical tones uttered in isolation, based on 50 samples for each tone, produced by a male speaker in his 20s at the time of recording (born in 1983).

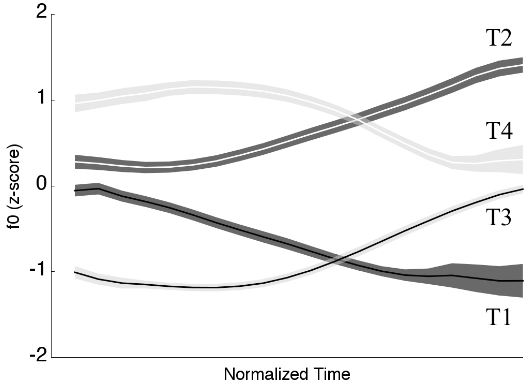

Figure 5.1 *Four lexical tones in Tianjin Mandarin produced in isolation (with normalized time). Lines stand for the mean. Gray areas stand for* ±*1 standard error of the mean.*

When two tones are combined, several of the disyllabic tonal sequences undergo tone sandhi changes from one tone to another in Tianjin Mandarin. For example, when two T3s are combined, the first one is realized with a high rising *f0*, resembling that of the lexical T2 with only subtle differences (Zhang & Liu, 2011; Li & Chen, 2016; but see e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007 which proposed a categorical change of lexical T3 to lexical T2). Furthermore, another type of tone sandhi has been reported for the tonal sequences of T1T1 and T4T1, where $T1_{sandhi}$ and $T4_{sandhi}$ are significantly different from their respective targeted output tones as claimed in the literature (i.e., T3 and T2, respectively) (Li & Chen, 2016; but see e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007 which categorize these tonal variants also as the categorical changes of lexical tones, i.e. T1 and T4, to T3 and T2, respectively). Given the different degrees of resemblance to another lexical tone within the tonal inventory (as illustrated in Figure 5.2 below), T3T3 has been classified as Near-Merger Sandhi while T1T1 and T4T1 as No-Merger Sandhi (Li & Chen, 2016).

Figure 5.2 illustrates the *f0* contours of the three tone sandhi sequences compared to that of their respective targeted output sequences as claimed in the literature. The data were averaged across 72 tokens (six speakers, three items, two informational status, two repetitions). White lines with dark areas represent the *f0* realization of the tone sandhi sequences; and black lines with light areas represent the claimed target output sequences. In the Near-Merger Sandhi case (i.e., the T3T3 sequence in Figure 5.2*a*), the first T3 is realized with a high-rising *f0* contour, which greatly resembles that of the first T2 in T2T3. The *f0* realization of T3T3 tonal sequence, therefore, shows great similarity to that of the claimed targeted output sequence (T2T3) in the literature despite that no complete neutralization is involved (Zhang & Liu, 2011; Li & Chen, 2016). In the No-Merger Sandhi cases (i.e., T1T1 and T4T1), both $T1_{sandhi}$ and $T4_{sandhi}$ surface with a much altered *f0* contour, while still maintaining tonal distinctiveness from their respective targeted output tonal contours claimed in the literature. Specifically, T1T1 (Figure 5.2*b*) has been described to change into T3T1 in the literature (e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007). As shown in Figure 5.2*b*, although the first T1 in T1T1 is realized similarly to T3 as in T3T1, its *f0* realization is clearly different from that of T3, suggesting no merging of T1$_{sandhi}$ with T3. For T4T1 (Figure 5.2*c*), similarly, although the first T4 in T4T1 is similar to the claimed targeted output - T2 (e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007), its *f0* realization is clearly different from that of lexical T2.

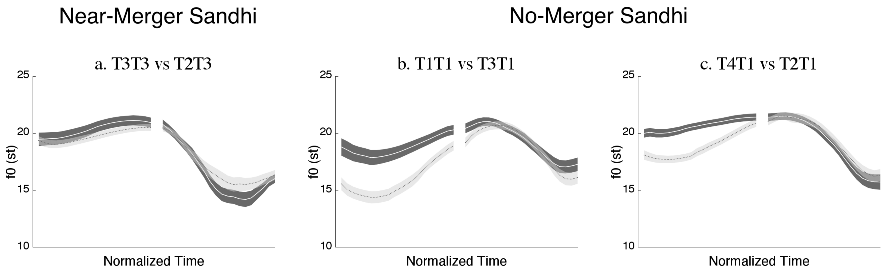

Figure 5.2 *f0 realization of three disyllabic sandhi sequences in Tianjin Mandarin (T3T3 in a, T1T1 in b and T4T1 in c) compared to their respective sandhi target sequences (T2T3 in a, T3T1 in b, T2T1 in c) as claimed in the literature. T3T3 is Near-Merger Sandhi; T1T1 and T4T1 are No-Merger Sandhi. Thick white lines indicate the mean f0 of the sandhi sequences (dark gray areas for ±1 standard error of mean); black lines indicate the mean f0 of the claimed target sequences (light gray areas for ±1 standard error of mean). Normalized time.*

The differences between the two sandhi types have been found to significantly influence the recognition of the lexical tones as in our recent eye-tracking study (Li & Chen, 2015 and Chapter 4). The eye movement data show that listeners have more difficulty in processing Near-Merger sandhi tone than No-Merger sandhi tones, due to the fact that for Near-Merger sandhi, the sandhi-derived tone is more similar to another lexical tone within the lexical tonal inventory in terms of the *f0* realization, while the No-Merger sandhi tones remain distinctive from the lexical tones with the most similar *f0* contours (see below for further details).

This study taps further into the nature of tone sandhi recognition. To be specific, we ask whether the online processing of sandhi tones is in a similar way to that of the non-sandhi variants (Toneme+, Contour-) or another lexical tone which has been claimed as the output target of the tone sandhi rules due to *f0* contour similarity (Toneme-, Contour+), and whether different sandhi types might differ in how the sandhi tones are recognized. These should shed further light on how tonal variants are represented in the mental lexicon.

5.3 Visual world paradigm

To investigate the recognition of tone sandhi, we employed the "Visual World Paradigm" (Tanenhaus et al., 1995) with an auditory word-recognition task following Li and Chen (2015), in which the eye movements of listeners are tracked while they listen to auditory stimuli. As in many experiments with the Visual World Paradigm, participants are presented with multiple objects (either as images or in written words) on a computer screen. They are required to follow instructions to complete certain tasks with one of these objects (i.e., the target) upon hearing an auditory stimulus corresponding to the object. Their eye movements are recorded simultaneously for later analyses. A typical display of visual stimuli consists of a target, a competitor as well as two distractors. The participants are asked to identify the words they have just heard and click on it with a mouse (see Huettig et al., 2011 for a review).

This paradigm has enabled us to tap into the time course of auditory word recognition based on the assumption that eye movements are closely time-locked to the spoken-word processing in an incremental fashion (e.g., Tanenhaus et al., 1995; Allopenna et al., 1998; Dahan et al. 2001a, 2001b; McMurray, Tanenhaus & Aslin, 2002; Beddor et al., 2013; see also reviews in Tanenhaus et al., 2000; Huettig et al., 2011). Furthermore, tasks involved in this paradigm are more natural than meta-linguistic judgement tasks in terms of perception experience (Spivey, 2007).

An important dependent variable in the visual world paradigm is the proportion of looks to each visual object over time - the likelihood of certain object being looked at during the time window. Two indices are found to effectively reflect how auditory stimuli are processed.

One is concerned with how fast listeners fixate on the target object, which is related to how early the target starts to be activated. For example, in Tanenhaus et al. (1995), it has been found that when the target object (e.g., *candy*) was presented together with another object whose name has phonological overlapping with the target (e.g., *candle*), participants took longer time (i.e., 230ms) to initiate an eye movement to the target object *candy* than when no phonologically similar object was presented (i.e., 145ms). This shows that when there exists a phonologically overlapping competitor, the activation of the target is delayed due to stronger competition effect in target recognition compared to when there is no such competitor.

The other important index is how likely listeners are looking at the targets (or competitors), which is believed to reflect the level of activation of the targets (or competitors). In Allopenna et al. (1998), participants were presented with four images on the computer screen, one corresponding to the auditory stimuli (target, e.g., *beaker*), one for "onset competitor" which has onset overlapping with the target (e.g., *beetle*, which shares the onset with *beaker*), and the rest two for distractors that are phonologically unrelated with the target (e.g., *dolphin*, *carriage*). When participants heard the instruction "*Pick up the beaker*", the proportion of looks to both *beaker* and *beetle* started to increase initially, indicating the start of activation for both words. However, as the speech unfolded, *beetle* was gradually deactivated and consequently, the proportion of looks to it started to decrease; while the target picture gained more looks given the increased activation of the target word.

This paradigm has been successfully used in recent studies on lexical tone perception, which demonstrate that listeners are sensitive to fine-grained *f0* details (Malins & Joanisse, 2010; Shen et al., 2013; Li & Chen, 2015). For example, Li and Chen (2015) investigated the effect of contextual tonal variation on the perception of lexical tones of Tianjin Mandarin in connected speech. Participants heard disyllabic auditory target stimuli, which undergo three types of tonal variation: Near-Merger Sandhi, No-Merger Sandhi and No

Sandhi coarticulation. Upon hearing each auditory stimulus, participants saw four disyllabic collocations on the computer screen: a Target which corresponds to the auditory stimulus (e.g., *zhi*^{3} *fa*^{3} 'fingering'), a Critical Competitor of which the first syllable shares the segments but has an unrelated tone with the target (e.g., *zhi*^{4} *hou*^{4} 'lag'), and two phonologically unrelated distractors (e.g., *san*^{1} *jiao*^{3} 'triangle'; *gua*^{4} *hao*^{4} 'register'). Listeners were asked to click, with a mouse, on the target amongst the four possibilities upon hearing each auditory stimulus. Results suggest that No-Sandhi coarticulation was the easiest to recognize, as in an earlier time window, participants' fixation on the targets increased at the fastest rate among the three conditions. Between the two sandhi conditions, Near-Merger Sandhi was more difficult to process than No-Merger Sandhi, which was reflected in the overall less proportion of looks to target in the Near-Merger Sandhi condition within a later time window.

In the present study, we set out to investigate the recognition time course of different target stimuli when the targets are presented with different types of competitors. Specifically, target stimuli in the present study included two types of tone sandhi, i.e., Near-Merger Sandhi (i.e., T3T3) and No-Merger Sandhi (i.e., T1T1, T4T1), as well as those with No Sandhi coarticulation (i.e., T4T4, T3T2, T3T4) as a control. Three types of disyllabic competitors were included for these targets which all have segmental overlap with the target for the first syllable: 1) Toneme Competitor whose first tone shares the underlying toneme with that of the target but with distinctively different *f0* realization from the target, i.e., Toneme+, Contour-; 2) Contour Competitor whose first tone has a different toneme but surfaces with a similar *f0* contour with that of the target, i.e., Toneme-, Contour+; 3) Segmental Competitor whose first tone does not share toneme or contour with the target, i.e., Toneme-, Contour-. Taking the target *zhi*^{3} *fa*^{3} ('fingering') for example, the Toneme Competitor is *zhi*^{3} *xiang*^{1} ('carton') whose first tone shares the T3 toneme with that of the target, the Contour Competitor is *zhi*^{2} *cheng*^{1} ('professional title') in which the *f0* contour of the first tone is similar to that of the target, and the Segmental Competitor is *zhi*^{4} *hou*^{4} ('lag') whose first tone is unrelated to that of the target. We are interested in how the proportion of looks to the target is affected by these three different competitor types, and how the effect of different competitors might interact with different target types.

5.3 Method

5.3.1 Participants

All participants in Chapter 4 participated in this experiment. The data from one subject who wore cosmetic contact lens were excluded. Data from another two subjects who were shortsighted without wearing glasses for correction were also excluded. All the remaining 31 subjects had normal or corrected-to-normal vision. All participants provided written informed consent.

5.3.2 Stimuli

The target stimuli in this experiment were the same as those in Chapter 4, including 18 near-merger sandhi targets (i.e., T3T3), 18 no-merger sandhi targets (i.e., T1T1, T4T1), and 18 no sandhi target (i.e., T4T4, T3T2, T3T4)[5].

For each target, a corresponding toneme competitor, a contour competitor and a segmental competitor were chosen (a within-item design). All competitors share segmental information with the targets at the first syllable, with manipulation only in tones. The first syllable of the toneme competitors shared the underlying toneme with that of the targets. The first tone of the contour competitors is the claimed output tone of that in the targets according to the sandhi rules claimed in the literature, whose *f0* contour resembles that of the targets. The segmental competitor does not share the underlying toneme or the surface *f0* realization with that of the targets, which are therefore unrelated to that of the targets. The second syllables of the targets and competitors are different in terms of both tone and segment across all conditions. Each target-competitor pair also has two distractors within the Visual World Paradigm. These distractors do not share any tone (toneme or surface contour) or segments with the target or the competitor. Table 5.1 illustrates the experimental design and sample stimuli.

The targets and competitors were further controlled to be closely matched in terms of lexical frequency based on Cai and Brysbaert (2010) as well as orthographic complexity (following Shen et al., 2013) as we presented Chinese characters instead of pictures. There was no significant difference between the targets and competitors for lexical frequency ($F(3)=0.108$, $p=0.955$) or visual complexity across conditions ($F(3)=1.526$, $p=0.209$). No participant reported difficulty of recognizing any character. All auditory stimuli were the same as those in Chapter 4.

Table 5.1 *Experimental design and sample stimuli.*

Near-Merger Sandhi			
Target	Toneme Competitor	Distractor 1	Distractor 2
	纸箱 (zhi^3 xiang^1) *carton*	电话 (dian^4 hua^4) *telephone*	斑马 (ban^1 ma^3) *zebra*
	Contour Competitor	Distractor 1	Distractor 2
指法 (zhi^3 fa^3) *fingering*	职称 (zhi^2 cheng^1) *professional title*	操场 (cao^1 chang^3) *playground*	机场 (ji^1 chang^3) *airport*
	Segmental Competitor	Distractor 1	Distractor 2
	滞后 (zhi^4 hou^4) *lag*	三角 (san^1 jiao^3) *triangle*	挂号 (gua^4 hao^4) *to register*

[5] No Sandhi stimuli included those claimed to undergo certain tone sandhi in impressionistic studies (e.g., Li & Liu, 1985; Chen, 2000; Wee et al. 2005) but not confirmed by empirical data in Chapter 3 and in Li & Chen (2016).

	No-Merger Sandhi		
Target	Toneme Competitor	Distractor 1	Distractor 2
	气球	樱花	香肠
	(qi^4 qiu^2)	($ying^1$ hua^1)	($xiang^1$ $chang^2$)
	balloon	*cherry blossom*	*sausage*
	Contour Competitor	Distractor 1	Distractor 2
器官	棋牌	钢琴	接力
(qi^4 $guan^1$)	(qi^2 pai^2)	($gang^1$ qin^2)	(jie^1 li^4)
organ	chess and cards	*piano*	*rally*
	Segmental Competitor	Distractor 1	Distractor 2
	乞丐	餐巾	空气
	(qi^3 gai^4)	(can^1 jin^1)	($kong^1$ qi^4)
	beggar	*napkin*	*air*
	No Sandhi		
Target	Toneme Competitor	Distractor 1	Distractor 2
	稳重	办公	记号
	(wen^3 $zhong^4$)	(ban^4 $gong^1$)	(ji^4 hao^4)
	steady	*to work*	*marker*
	Contour Competitor	Distractor 1	Distractor 2
吻合	温度	裤袜	外币
(wen^3 he^2)	(wen^1 du^4)	(ku^4 wa^4)	(wai^4 bi^4)
to match	*tempareture*	*leggings*	*foreign currency*
	Segmental Competitor	Distractor 1	Distractor 2
	问好	住址	袋鼠
	(wen^4 hao^3)	(zhu^4 zhi^3)	(dai^4 shu^3)
	to greet	*home address*	*kangaroo*

5.3.3 Procedure and eye movement data analysis

We followed the same experimental procedure and eye movement data analysis as that in Chapter 4. The eye movement data were also reported from the onset of the auditory stimulus to 1400ms post stimulus onset following Chapter 4. We employed the growth curve analyses procedure with the package *lme4* (Bates et al., 2014) in R (R Core Team, 2014) as in Chapter 4.

The critical independent factors included TARGET TYPE and COMPETITOR TYPE. We also included the control factors TARGET FREQUENCY, TARGET STROKE NO., TARGET STRUCTURE, TARGET BIGRAM INFO., COMPETITOR FREQUENCY, COMPETITOR STROKE NO., as well as COMPETITOR STRUCTURE. Our main interest was to see the possible effects of TARGET TYPE and COMPETITOR TYPE after having regressed out possible mainly stimulus-intrinsic factors as listed above. Multiple comparisons with Bonferroni adjustment were conducted with the function *glht* in package *multcomp* (Hothorn et al., 2008) whenever necessary.

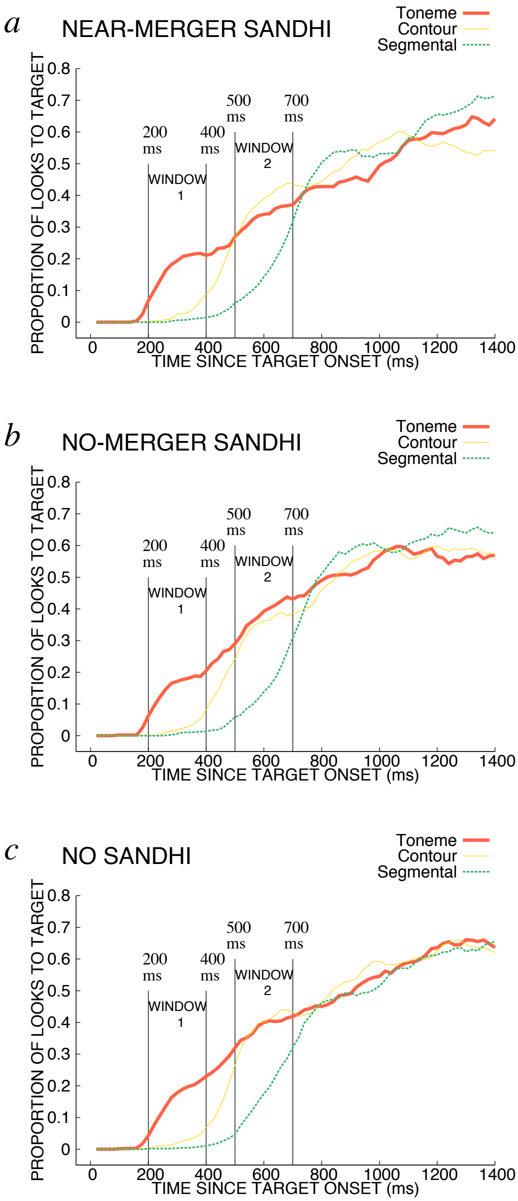

Figure 5.3 *Mean proportion of looks to target when targets of different types were presented with different competitors. Targets with Near-Merger Sandhi were plotted in a, No-Merger Sandhi targets in b, No Sandhi targets in c. Thick solid lines stand for the condition when presented with a Toneme Competitor, thin solid lines when presented with a Contour Competitor, thin dashed lines when presented with a Segmental Competitor. Aggregated across participants and items.*

5.4 Results

To investigate how sandhi tones are recognized, we included analyses for looks to both targets and competitors in this section.

5.4.1 Looks to the target

Figure 5.3 shows the mean proportion of looks to the target when each type of target was presented with different competitors (illustrated with different line types). X-axis stands for the time since auditory target onset, y-axis for the proportion of looks to target/competitor. Targets with Near-Merger Sandhi were plotted in *a*, No-Merger Sandhi targets in *b*, No Sandhi targets in *c*.

We can see from Figure 5.3 that for all target types, there are great differences in the proportion of looks to the target due to different competitors, especially from 200ms to 700ms after auditory stimuli onset. After 700ms, the differences between the three competitor conditions are much more reduced. We further zoomed into two particular time windows which show interesting eye movement patterns in each target type. The first window is from 200-400ms after auditory target onset. Given the 200ms delay to plan and execute an eye movement (Matin et al., 1993), 200-400ms relatively corresponds to 200ms into the auditory stimuli which maximally ensures no information of the second syllable is available during this time window, as the minimal duration of the first syllable is 220ms in our dataset. The second window is from 500ms to 700ms, following Li and Chen (2015) and data reported in Chapter 4 in order to make a direct comparison.

From visual inspection of the 200-400ms time window (WINDOW 1, Figures 5.3*a,b,c*), the proportion of looks to the target in the three competitor conditions seems to start increasing at different time points. Participants took the least time to initiate an eye movement to the target when there is a Toneme Competitor, i.e., around 200ms after the auditory stimulus onset. The proportion of looks to the target with the presentation of the other two types of competitors started to increase at much later time points, i.e., around 300ms for the Contour Competitor condition and after 400ms for the Baseline Competitor condition. This has resulted in a much greater overall mean (around 20%) and a much faster increase in the proportion of looks to the target in Toneme Competitor condition than in the other two competitor conditions within this time window. The Contour Competitor condition and Segmental Competitor condition show only slight differences in the proportion of looks to the target, both less than 10%. Similar patterns could be consistently observed across three target types.

A growth curve analysis was run for the comparison of the proportion of looks to the target among the three competitor conditions within the 200-400ms window across target types. The best-fit model contained the interaction of TIME (up to fourth-order) and COMPETITOR TYPE (three levels: TONEME COMPETITOR, CONTOUR COMPETITOR, SEGMENTAL COMPETITOR) in the fixed factor structure. TARGET TYPE did not significantly improve the model fit ($\chi^2(10)=12.10$, $p=0.28$, n.s.), which suggested that different sandhi types in the target did not affect the proportion of looks patterns within the 200-400ms time window. Among all the controlled factors, only COMPETITOR STRUCTURE showed significant improvement of the model ($\chi^2(2)=27.62$, $p<0.001$), and thus was added to the model. SUBJECT was considered as a random factor over all time terms. Table 5.2 shows the estimation of the effect of competitor types.

Table 5.2 *Pairwise comparison of the effect of competitor types on the proportion of looks to target within the time window 200-400ms.*

	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=-0.14 z=-51.75 p<0.001	β=-0.17 z=-64.25 p<0.001	β=-0.03 z=-11.87 p<0.001
Slope	β=-0.03 z=-3.69 p<0.01	β=-0.11 z=-12.41 p<0.001	β=-0.08 z=-8.66 p<0.001
Quadratic	β=0.07 z=8.20 p<0.001	β=0.04 z=4.76 p<0.001	β=-0.03 z=-3.42 p<0.01
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.

The statistic results confirmed the observations made from WINDOW 1 of Figure 5.3. As shown in Table 5.2, the proportion of looks to target in the Toneme Competitor condition had a significantly higher overall mean than that of the other two conditions (see significant results in the intercept of Contour vs. Toneme and Segmental vs. Toneme in Tables 5.2*a,b*). The Contour Competitor condition had a slightly but significantly higher overall mean than that of Segmental Competitor condition (see significant results in the intercept of Baseline vs. Contour in Table 5.2*c*). Furthermore, the Toneme Competitor condition showed a significantly faster proportion increase than the Contour Competitor condition, which was further faster than the Segmental Competitor condition, as reflected by the significant differences in the slope and quadratic terms in Tables 5.2*a,b,c*).

The second interesting time window is from 500ms to 700ms (WINDOW 2, Figures 5.3*a,b,c*). In this window, the proportion of looks to target in Toneme and Contour Competitor conditions keep increasing at a similar rate (both rising from 30% to around 40%). The Segmental Competitor condition, however, just starts to increase but at a relatively faster increasing rate (from less than 10% to around 30%).

Within this window, both COMPETITOR TYPE ($\chi^2(10)=2047.1$, $p<0.001$) and TARGET TYPE ($\chi^2(10)=30.25$, $p<0.001$) significantly improved the model fits. They also showed significant interactions ($\chi^2(20)=50.69$, $p<0.001$). To investigate the effect of competitor in each target type, we further subset the data to run separate analyses within each target type. Estimates of the effect of competitors are listed in Table 5.3.

Two observations can be made from Table 5.3. First, in all three target types, both the

Toneme Competitor condition and the Contour Competitor condition were significantly different from the Segmental Competitor condition in terms of significantly higher overall mean but slower increase in the proportion of looks to the target. This was reflected by the significant results in the intercept, slope as well as the quadratic of Segmental vs. Toneme and Segmental vs. Contour in Tables 5.3I*b,c*, II*b,c*, III*b,c*. Meanwhile, the differences between the Toneme and Contour Competitor conditions were relatively smaller. Only very subtle overall mean differences were observed (see significant results in the intercept only in Tables 5.3I*a*, II*a*, and non-significant results in Table 5.3III*a*).

Second, it is notable that we observed a significantly larger overall mean of the proportion of looks to target in the Contour Competitor condition than in Toneme Competitor condition for Near-Merger Sandhi targets. This could be seen from the significant results in the Contour vs. Toneme of Table 5.3Ia.

Table 5.3 *Pairwise comparison of the effect of competitor types on the proportion of looks to target within the time window 500-700ms in different target types.*

	I. Near-Merger Sandhi		
	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=0.04 z=5.07 p<0.001	β=-0.17, z=-18.73 p<0.001	β=-0.21 z=-23.35 p<0.001
Slope	n.s.	β=0.21 z=7.47 p<0.001	β=0.17 z=5.87 p<0.001
Quadratic	n.s.	n.s.	β=0.11 z=3.71 p<0.01
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.
	II. No-Merger Sandhi		
	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=-0.06 z=-6.14 p<0.001	β=-0.21 z=-24.79 p<0.001	β=-0.16 z=-17.02 p<0.001
Slope	n.s.	β=0.16 z=5.85 p<0.001	β=0.20 z=7.05 p<0.001
Quadratic	n.s.	β=0.09 z=3.19 p<0.05	β=0.10 z=3.51 p<0.01
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.

III. No Sandhi			
	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	n.s.	β=-0.20 z=-21.86 p<0.001	β=-0.21 z=-22.95 p<0.001
Slope	n.s	β=0.21 z=7.17 p<0.001	β=0.18 z=6.25 p<0.001
Quadratic	n.s.	n.s.	n.s.
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.

5.4.2 Looks to the competitor

To observe where participants were looking if the target is not looked at, this section further analyzes the eye movement patterns over competitors. Looks to competitor were analyzed in a similar fashion to that for looks to target. Figure 5.4 shows the mean proportion of looks to the competitor when each type of target was presented with different competitors (illustrated with different line types). X-axis stands for the time since auditory target onset, y-axis for the proportion of looks to target/competitor. Targets with Near-Merger Sandhi were plotted in *a*, No-Merger Sandhi targets in *b*, No Sandhi targets in *c*. Two observations can be made.

First, it can be seen from WINDOW 1 in Figures 5.4a,b,c that, participants directed more looks to the Contour Competitor and the Segmental Competitor (both above 20%) than to the Toneme Competitor (less than 20%) at the beginning of the auditory stimuli (200-400ms). Meanwhile, the Contour Competitor and the Segmental Competitor are relatively similar to each other during this time window. This pattern was consistently observed across three target types, which was confirmed by the statistic results that TARGET TYPE did significantly improve the model fit of the data. As shown in Table 5.4, results of the best-fit model confirmed this with significant results in the proportion of looks between the Contour Competitor condition vs. the Toneme Competitor condition (see significant results in the intercept in Figure 5.4*a*) and Segmental Competitor vs. Toneme Competitor condition (see significant results in the intercept and slope in Figure 5.4*b*). Contour Competitor condition and Segmental Competitor condition did not differ significantly from each other.

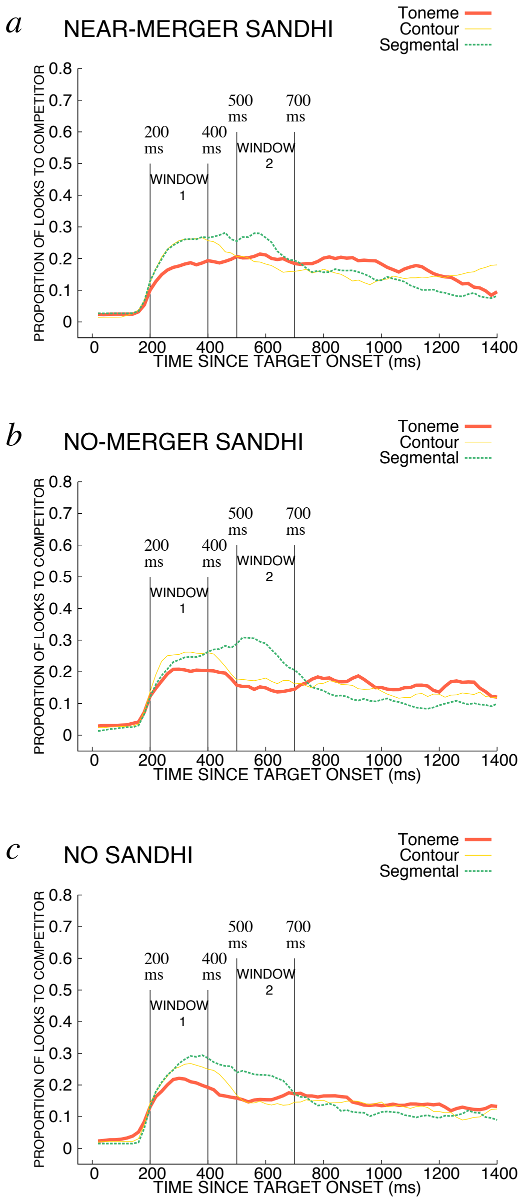

Figure 5.4 *Mean proportion of looks to competitor when targets of different types were presented with different competitors. Targets with Near-Merger Sandhi were plotted in a, No-Merger Sandhi targets in b, No Sandhi targets in c. Thick solid lines stand for the condition when presented with a Toneme Competitor, thin solid lines when presented with a Contour Competitor, thin dashed lines when presented with a Segmental Competitor. Aggregated across participants and items.*

Table 5.4 *Pairwise comparison of the effect of competitor types on the proportion of looks to competitor within the time window 200-400ms.*

	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=0.06 z=11.92 p<0.001	β=0.05 z=11.30 p<0.001	n.s.
Slope	n.s.	β=0.06 z=3.94 p<0.01	n.s.
Quadratic	n.s.	n.s.	n.s.
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.

Second, within the 500-700ms window (WINDOW 2, Figures 5.4*a,b,c*), the participants looked at the Segmental Competitor (above 20%) more than at the Toneme and Contour Competitors (both less than 20%). This was reflected by the significant results in both the intercept and the slope of Segmental vs Toneme and Segmental vs. Contour in Figures 5.5Ib,c, IIb,c, IIIb,c. The Toneme Competitor condition and the Contour Competitor condition further differ from each other according to different target types. For both No-Merger Sandhi and No Sandhi targets, participants' looks to the Toneme and Contour Competitors clearly drop compared to that within the previous time window (WINDOW 1, Figures 5.4b,c). While for Near-Merger Sandhi targets, the proportion of looks to the competitor in the Toneme Competitor condition remains as that in the previous time window (WINDOW 1, Figure 5.4a). This was confirmed by the significantly higher overall mean in the looks to the Toneme Competitor than Contour Competitor (see significant results in the intercept of Contour vs. Toneme in Table 5.5I*a*).

Table 5.5 *Pairwise comparison of the effect of competitor types on the proportion of looks to competitor within the time window 500-700ms in different target types.*

	I. Near-Merger Sandhi		
	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=-0.02 z=3.33 p<0.05	β=0.03 z=3.77 p<0.01	β=0.06 z=6.83 p<0.001
Slope	n.s.	β=-0.08 z=-3.21 p<0.05	n.s.
Quadratic	n.s.	n.s.	n.s.
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.

II. No-Merger Sandhi			
	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=0.03 z=3.63 p<0.01	β=0.12 z=15.58 p<0.001	β=0.09 z=11.18 p<0.001
Slope	n.s.	β=-0.13 z=-5.30 p<0.001	β=-0.12 z=-4.99 p<0.001
Quadratic	n.s.	n.s.	n.s.
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.
III. No Sandhi			
	a. Contour vs. Toneme	*b.* Segmental vs. Toneme	*c.* Segmental vs. Contour
Intercept	β=-0.02 z=-3.423 p<0.01	β=0.05 z=7.16 p<0.001	β=0.08 z=10.54 p<0.001
Slope	n.s	β=-0.11 z=-4.57 p<0.001	β=-0.07 z=-3.06 p<0.05
Quadratic	n.s.	n.s.	n.s.
Cubic	n.s.	n.s.	n.s.
Quartic	n.s.	n.s.	n.s.

5.5 Discussion & conclusion

This study taps into the online recognition of tonal sandhi variants in Tianjin Mandarin via the Visual World Paradigm. We examined the participants' looks to the target and competitor when they were presented with three types of competitors upon hearing two types of tone sandhi (i.e., Near-Merger Sandhi and No-Merger Sandhi) as well as stimuli with tonal coarticulation and no sandhi involved (i.e., No Sandhi) as a baseline for comparison. For each target stimulus, we included a Toneme Competitor which has toneme overlap with the target, a Contour Competitor whose *f0* contour resembles that of the target but of a different toneme, as well as a Segmental Competitor with an unrelated tone to that of the target. Results show that in all target types, participants were sensitive to the types of competitors together presented with the targets, which was reflected in the different eye movement patterns when participants were presented with different types of competitors.

First, regardless of the target types, there is a Toneme Competitor benefit across all target types within an earlier time window (200-400ms). During this stage, participants directed significantly more looks to the target (WINDOW 1, Figure 5.3) and less looks to the competitor (WINDOW 1, Figure 5.4) in the Toneme Competitor condition. This was reflected by significantly higher overall mean and significantly faster increase of the proportion of looks to target as well as the significantly lower overall mean of the proportion of looks to competitor in the Toneme Competitor condition. The other two competitor conditions (i.e., Contour Competitor condition and Contour Competitor condition) exhibit relatively similar eye movement patterns (reflected in the gaze pattern for both target and competitor). This has thus suggested that a Contour Competitor (Toneme-, Contour+) plays a similar role in the processing of the target to that of the Segmental Competitor (Toneme-, Contour-) at the first 200ms of the target stimuli, both of which show stronger activation than the Toneme Competitor (Toneme+, Contour-).

Second, in the later time window (500-700ms), the activation of targets in both Toneme and Contour Competitor conditions continues to increase, while the activation of target in the Segmental Competitor condition just starts to increase (WINDOW 2, Figure 5.3). This was confirmed by the significantly lower but faster increase in proportion of looks to the target in the Segmental Competitor condition than the other two competitor conditions. The data of looks to competitor also show that the participants directed more looks to the Segmental Competitor more than Toneme and Contour Competitors, suggesting a stronger competition effect between targets and Segmental Competitors.

It has been repeatedly observed in previous studies on spoken word recognition within the Visual World Paradigm that, segmental overlap between the target and the competitor gives rise to competition effect in listeners' recognition of the spoken stimuli (see e.g., Tanenhaus et al., 1995; Allopenna et al., 1998; Dahan et al., 2001a, 2001b; McMurray et al., 2002; Beddor et al., 2013). Based on recent evidence that segments and lexical tones play a similar role in constraining word recognition (Malins & Joanisse, 2011), we predicted that tonal overlap, or at least partial tonal overlap, should induce further competition effect compared to the condition with no tonal overlap. Therefore, we would have expected more competition between target and competitor in the Toneme Competitor condition and the Contour Competitor condition than in the Segmental Competitor condition. Instead, what we have observed was that within the early window (i.e., 200-400ms), there was a delayed activation of the target in the Contour Competitor and the Segmental Competitor conditions, suggesting a different processing trajectory of the target words in the face of their Toneme Competitor (i.e., earlier and greater activation) from that against a Segmental or Contour competitor. Interestingly, regardless of how far the surface *f0* realization of the target tone deviates from the surface *f0* realization of its Toneme Competitor (i.e., across the three target types), the pattern of early activation of target with Toneme Competitor remains. This suggests that at an earlier stage of processing, listeners are processing the tonal information, despite their contextual variation, as their underlying lexical tones. Thus even for targets with Near-Merger tone sandhi, their Contour Competitors, despite the similarity in tonal pitch contours, exert a similar competition effect just as the Segmental Competitors (i.e., segmental sharing only).

A similar delayed activation of the correct target has been observed in Malins and Joanisse (2010), where participants did not look at the target at all during the initial 200ms into the first syllable when there is a Segmental Competitor. Moreover, they also observed that compared to the baseline condition, participants initially directed more looks to the target when target and competitor have tonal overlap in addition to some segmental overlap with the targets. This is then in line with our earlier activation of target with

Toneme Competitor (although we had lexical tonal overlap as well as complete segmental overalp).

At a later time window of processing (500-700ms), however, we observed a different pattern of looks as a function of both target types and competitor types. Specifically, two observations are to be noted. First, across target types, there was a continued delay in the activation of the targets with Segmental Competitor. The activation trajectory for the target in the Contour Competitor condition shows a more comparable pattern to that in the Toneme Competitor condtion, which is a reversed pattern from the earlier window. Despite their similarity, there were subtle yet consistent differences across the target types. For the Near-Merger targets, where the *f0* contours of the target and the Contour Competitor resemble to such an extent that they are almost ambiguous, the activation of the target in the Contour Competitor condition reached a higher level than that in the Toneme Competitor condition. In the No-Merger Sandhi and No Sandhi target conditions, where the *f0* contours of the target and the competitor are clearly distinctive from each other, the Contour Competitor enhanced the activation of the target, compared to the Segmental Competitor, but to a much smaller extent than the Toneme Competitor.

Overall, our data show that the Contour Competitors (Toneme-, Contour+) play a comparable role to the Segmental Competitors (Toneme-, Contour-) during the initial activation of the target. Therefore, despite the claims in the literature that for Near-Merger tone sandhi, one lexical tone undergoes sandhi change and is realized as another lexical tone, our data suggest that upon hearing the target word, which has a similar *f0* contour as the Contour Competitor, it is the underlying lexical tone of the target word that is concurrently activated together with the segments. This explains the different activation patterns of the target word when there is a Toneme Competitor versus when there is a Contour or Segmental Competitor. This finding then lends support to the proposal in the literature that tonal variation due to tone sandhi is better viewed as allophonic variation of the underlying lexical tones and are stored together with the lexical toneme as allotones in the mental lexicon (Chen et al., 2011; Nixon et al., 2015). This is also in line with recent results of an ERP study on lexical tonal processing in Standard Chinese (X. Li & Y. Chen, in press). With an oddball paradigm, this study investigated the effect of allophonic variation on the mental representation and neural processing of lexical tones produced in isolation. All stimuli share segments /ma/ but with different lexical tones: high-level T1, rising T2, and dipping T3. The ERP results show that among the four oddball conditions (T1/T3, T3/T1, T2/T3, T3/T2; standard/deviant), the T1-T3 pair elicited symmetrical mismatch negativity effects (MMN), while the T2-T3 pair only showed asymmetrical MMN effects where there were significantly greater and earlier MMN effects in the T2/T3 condition than that in the reversed T3/T2 condition. This thus suggests the co-activation of representations of both non-sandhi T3 and $T3_{sandhi}$ even when listeners only hear a non-sandhi T3 in isolation.

Despite that tonal contour similarity (between the target and the competitor) did not seem to exert an effect on the activation of target tone within the earlier time window, it is important to note that there was a contour similarity effect at a later processing stage. Within the time window of 500-700ms, the tonal contour has been completed or is near completion. By then, contour similarity of the Contour Competitor seemed to trigger similar patterns of target activation as in the Toneme Competitor condition, evident across the target types. This is in sharp contrast to the delayed activation of target with Segmental Competitor. While more research is needed to understand the process, one possibility to explain this differences is that the activation of the target is facilitated either with underlying lexical tonal sharing or surface contour similarity. While in the Segmental Competitor condition, the delayed activation of the targets may be due to the lack of such facilitation effect. Furthermore, listeners seemed to be sensitive to the subtle differences in the contour similarity. When the target and competitor show great contour overlap as in the Near-Merger Sandhi case, the contour overlap seems to be more likely to facilitate the target activation, compared to when the target and the Contour Competitor share less contour similarity (i.e., in the No-Merger Sandhi and No Sandhi target conditions).

Our results are thus different from conclusions made in traditional tone sandhi perception studies which usually employ meta-linguistic judgement tasks. These studies have mainly investigated the perception of sandhi tones in Standard Chinese which only has one Near-Merger sandhi (i.e., the change of a 3rd Low tone before another Low; known as Third-Tone/Low-Tone Sandhi), and have argued for complete neutralization of the Sandhi Low tone with the lexical Rising tone due to the lack of reliable perception between the two tones (e.g., Wang & Li, 1967; Speer et al., 1989; Chen, 2013; Chen et al., 2015). Given the multiple layers of tone sandhi in Tianjin Mandarin, we show, with eye tracking data which reveal the dynamic processing of spoken words, that tone sandhi does not have to be the categorical change from one lexical tone to another. Regardless of how the sandhi-derived tones resemble another lexical tone within the tonal system, native listeners consistently process the claimed output tones (in the literature) as an unrelated tone from an early processing stage and was able to fixate upon the correct target word despite great similarity between the sandhi variant and the lexical tone competitor which is claimed in the literature to be the sandhi output tone. In other words, within the Visual World Paradigm, which is much more sensitive to the dynamic activation processes given a speech input, we have shown that listeners are able to differentiate a sandhi tonal variant from its contour competitor lexical tone well above chance level. This thus echoes well with recent experimental data showing subtle but consistent acoustic differences (see e.g., Zee, 1980; Peng, 1996; Yuan & Chen, 2014; Zhang & Liu, 2011; Li & Chen, 2016) as well as different tonal encoding processes in online speech production between sandhi-derived tonal variants and the lexical tones claimed to be their sandhi output tones in the literature (Chen et al., 2011; Nixon et al., 2015; Zhang et al., 2014). This thereby argues firmly against the view that tone sandhi in these Mandarin dialects involves the categorical change of one lexical tone into another.

CHAPTER 6 PROSODICALLY CONDITIONED NEUTRAL TONE REALIZATION IN TIANJIN MANDARIN[6]

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: $z = \frac{f0_x - f0_{mean}}{f0_{SD}}$ (z: z-score; $f0_x$: the observed *f0* value in Hz; $f0_{mean}$: the mean *f0* value of the speaker in Hz; $f0_{SD}$: 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.

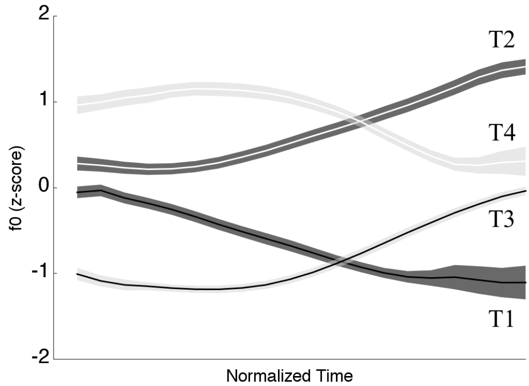

Figure 6.1 *Lexical tones produced in isolation. Lines stand for the mean. Gray areas stand for ±1 standard error of mean. Tone 1 (T1) is illustrated with black line and dark gray area; Tone 2 (T2) with white line and dark gray area; Tone 3 (T3) with black line and light gray area; Tone 4 (T4) with white 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 *wo³de* 'mine'), the second syllable of reduplicative forms (e.g., the second *ma* in *ma¹ma* 'mother'), or the second syllable of disyllabic lexical items (e.g., *jin* in tian¹jin *'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 *f0* 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 *f0* 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 *f0* realization as reported in the literature. If the rising *f0* 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 *f0* 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 *f0* realization. Otherwise, the so-called "special rising target" would be called into question.

2) *How is neutral tone f0 realization conditioned by different prosodic boundaries?*

By examining TM neutral tone realization, we also aimed to investigate how in general, the *f0* realization of neutral tone is conditioned by different prosodic boundaries. To this end, we were interested in how the *f0* 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 *f0* 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^1 shuo^1*… ("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^1 zhu^4 le na^4 zhi^1 lao^{3}shu^3.

T1 N|T1 T1

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 jing^1 ying^2 le yi^4 jia^1 can^1 ting^1.

T3 N||T1 T2

He said sisiter 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 *f0* 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^2 de mao^1** zhua^1 zhu^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^1 zhu^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^1 zhu^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 ($\chi^2(9)=10183$, $p<0.001$).

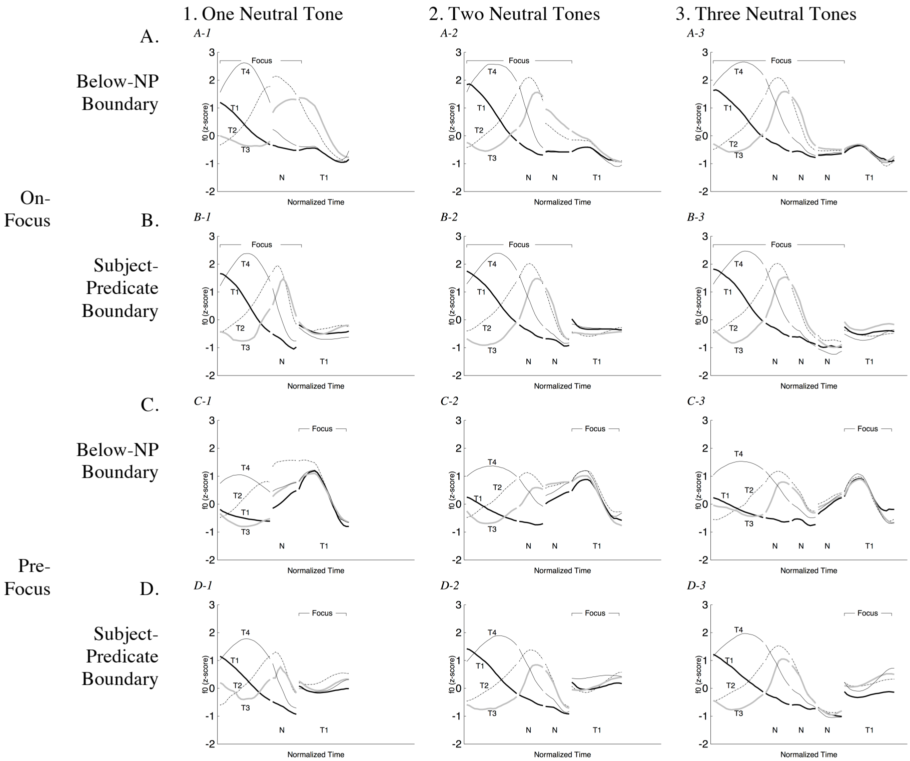

Figure 6.2 *Mean f0 contours of neutral tone syllables in different tonal contexts with different prosodic boundary inserted between neutral tone and the following full tone under different focus conditions. Normalized time.*

The domain of *f0* raising is revealed by comparing *f0* 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 *f0* was indeed due to the so-called rising neutral tonal target, we would have observed consecutive rising neutral tone *f0* contours over all neutral tone syllables or continuous rising *f0* 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 *f0* rising contours. For those which does show rising *f0* 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 *f0* contour of neutral tone was unlikely due to the underlying *f0* 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 (1st: $\chi^2(9)=16874$, $p<0.001$; 2nd: $\chi^2(9)=4980.4$, $p<0.001$).

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

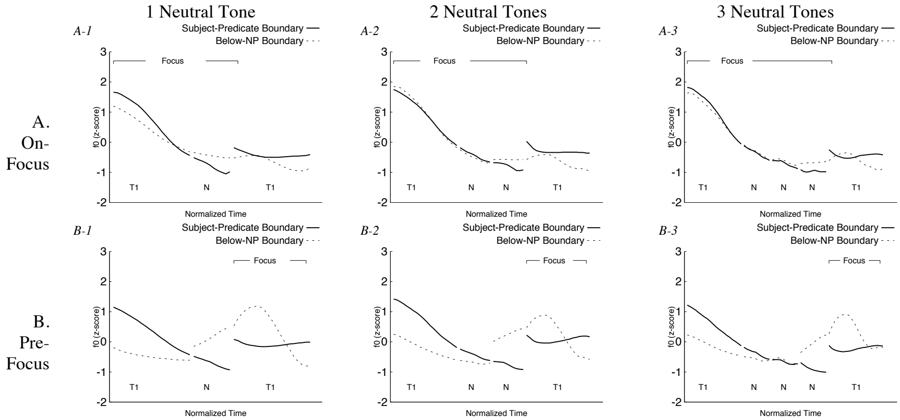

Figure 6.3 *Neutral tone realization in the context T1N(N)(N)T1 with different prosodic boundaries between neutral tone and the following T1 in different focus conditions. Row A is for On-Focus condition, Row B for Pre-Focus condition. Normalized time.*

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.*

			a. On-Focus			
	1st Neutral Tone		2nd Neutral Tone		3rd Neutral Tone	
1N	*Intercept*	$\beta=0.25$ $t=9.94$ $p<0.001$				
	Slope	$\beta=0.26$ $t=3.32$ $p<0.001$				
	Quadratic	n.s.				
2N	-		*Intercept*	$\beta=0.18$ $t=12.14$ $p<0.001$		
			Slope	$\beta=0.22$ $t=4.54$ $p<0.001$		
			Quadratic	n.s.		
3N	-		-		*Intercept*	$\beta=0.17$ $t=4.96$ $p<0.001$
					Slope	$\beta=0.23$ $t=2.12$ $p<0.05$
					Quadratic	n.s.
			b. Pre-Focus			
	1st Neutral Tone		2nd Neutral Tone		3rd Neutral Tone	
1N	*Intercept*	$\beta=0.87$ $t=38.15$ $p<0.001$				
	Slope	$\beta=1.14$ $t=15.79$ $p<0.001$				
	Quadratic	n.s.				
2N	-		*Intercept*	$\beta=0.99$ $t=57.70$ $p<0.001$		
			Slope	$\beta=0.78$ $t=14.48$ $p<0.001$		
			Quadratic	n.s.		

3N	-	-	*Intercept*	β=0.99 t=62.40 p<0.001
			Slope	β=0.64 t=12.71 p<0.001
			Quadratic	β=-0.16 t=-3.22 p<0.01

Two observations can be made from Figure 6.3. First, neutral tone immediately preceding a Below-NP Boundary (dotted lines) shows a salient raised *f0* contour, in contrast to the continued *f0* 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 1st Neutral Tone of 1N in Tables 6.1*a*, *b*, the 2nd Neutral Tone of 2N in Tables 6.1*a*, *b*, the 3rd Neutral Tone of 3N in Tables 6.1*a*, *b*).

Second, the magnitude of *f0* raising was related to the focus conditions of the neutral tone(s). The raising of the neutral tone *f0* 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 *f0* 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 *f0* 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 *f0* 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 *f0*, 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 *f0* realization of neutral tone is conditioned by different prosodic boundaries, a question which has not been addressed in the literature.

6.4.1 Nature of rising neutral tone *f0*

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 *f0* realization after different lexical tones when there are multiple neutral tone syllables. Others accounted for the neutral tone *f0* realization as the transition (or interpolation) between its preceding and following lexical tonal targets (e.g., Shih, 1987; van Santen et al., 1998). Based on our data, this also seems implausible. As the number of neutral tone syllables increased, the *f0* realization of the neutral tone sequence should not have first reached for a mid-low target and then rose, if the *f0* 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 *f0*.

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 *f0* 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 *f0* 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 *f0* 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 *f0* 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 *f0* 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 *f0* raising effect. In Tianjin Mandarin, the low-register falling T1 triggers an anticipatory *f0* raising effect (Li & Chen, 2016), which is absent in Standard Chinese. In addition, we observed an interesting effect of prosodic boundaries on neutral tone *f0* 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.

CHAPTER 7 CONCLUSION

This dissertation investigated the production and perception of contextual tonal variation based on data from Tianjin Mandarin. While most of the previous studies on speech variability have focused on variation at the segmental level, this dissertation complements the existing literature with data on lexical tonal variation, a suprasegmental feature of the sound system. With evidence from Tianjin Mandarin, a dialect of the Northern Mandarin family, this dissertation provides an in-depth view of how the acoustic realization of lexical tones may vary due to both the local contextual factors (Chapter 3) and the global prosodic factors (Chapter 6), and how the different types of tonal variability further affect the online perception of the lexical tones (Chapters 4 and 5).

After an introduction to the sound system of Tianjin Mandarin (**Chapter 2**), **Chapter 3** examined the acoustic realization of two different types of local tonal variation. Specifically, a direct comparison was made of the *f0* variabilities introduced by two local tonal variation processes: tone sandhi and anticipatory tonal coarticulation, which was designed to address the question of how lexical tones are realized with varying *f0* contours in disyllabic domains without pre-exclusion of any tonal sequence claimed to involve tone sandhi processes. To further understand the nature of tone sandhi (especially the direction and possible cyclic application of tone sandhi) in Tianjin Mandarin, trisyllabic constituents were also investigated.

With the technique of Growth Curve Analysis (GCA) (Mirman, 2014), this chapter observed a range of variation in the *f0* realization of the four lexical tones due to different following tones. Of particular interest were two different anticipatory raising effects due to the low-falling T1 and the low-dipping T3, where T1 consistently gives rise to a greater *f0* rise over the preceding tones and T3 significantly raises the overall *f0* mean of the preceding tones. In the sequences of T3T3, T1T1 and T4T4, the raising effects have brought about major changes in the *f0* contour of the preceding tones so that their surface *f0* may not be predictable from the characteristics of the canonical tonal *f0* contours. Variation over the three sequences was therefore attributed to tone sandhi. Other cases are better accounted for as tonal coarticulation, as the anticipatory raising effects have not affected the *f0* distinctiveness of the preceding tones. Their characteristic *f0* contours can still be expected to accord with the canonical tonal shapes. Over trisyllabic sequences, however, the three confirmed disyllabic tone sandhi sequences were not consistently observed, counter predictions based on claims in the literature (e.g., Li & Liu, 1985; Chen, 2000; Hyman, 2007). Specifically, tone sandhi applications over T1T1 and T4T1 sequences are only possible when these disyllabic sequences are aligned to the right edge of a trisyllabic constituent; while the T3T3 sequence triggers consistent application of sandhi changes regardless of its position within a trisyllabic utterance. Furthermore, no tonal neutralization was observed over any of the sandhi processes. Rather, we argue for two different types of tone sandhi in terms of the degrees of *f0* resemblance between the sandhi-derived tonal contour and the tonal contour of another lexical tone within the system. To be specific, T3T3 is a near-merger sandhi case, as its sandhi output $T3_{sandhi}T3$

shows an almost merged *f0* contour with the lexical tonal sequence of T2T3. T1T1 and T4T1 can be classified as no-merger sandhi sequences, as their sandhi realizations have some resemblance to but remain distinctive from the contours of other lexical tones in the language.

This chapter thus presents the first empirical study to directly compare these two types of tonal variation within a well-controlled acoustic experiment. It is to be recognized that the boundary between tone sandhi and tonal coarticulation is not straight-forward, and tone sandhi and tonal coarticulation do show some commonalities as well as differences. Nevertheless, given the rich layers of tonal variation due to different tonal variation processes, our results show that it is important to first objectively observe the patterns of *f0* variation due to these sources, without which, any discussion upon the differences between these two variation types is unfounded. In addition, most of the previous impressionistic studies have considered tonal variation as the categorical change of tonal identity, which has created a difficult situation and unnecessary complications for theoretical accounts of tonal variability. By showing that tone sandhi does not involve the categorical change from one tone to another, this chapter has greatly simplified the claimed complex tone sandhi system in this dialect. Consequently, our results have cast serious doubts on the previous theoretical accounts of tonal alternation based on impressionistic data, and called for future theoretical works on tonal alternation based on the empirically more solid findings.

Chapter 4 further tapped into the perceptual consequences of the different tonal variability, namely, near-merger sandhi, no-merger sandhi and no sandhi tonal coarticulation. While the literature has suggested a clear contextual effect on listeners' identification of lexical tones, the current chapter asks how a deviated tone itself is processed by a listener prior to the access to the disambiguating tonal contexts during speech recognition. To address this question, this chapter employed a word-recognition task within the Visual World Paradigm (Tanenhaus et al., 1995).

Via examining the participants' looks to both targets and competitors within the paradigm when they heard disyllabic stimuli with different types of tonal variability (i.e., near-merger sandhi, no-merger sandhi, no sandhi tonal coarticulation), this chapter observed significant differences in the eye movement patterns as a function of different types of tonal variability. This suggests a clear effect of tonal variability on the perception of lexical tones prior to the facilitation of the neighboring tonal contexts. To be specific, tonal sequences with no sandhi tonal coarticulation was the easiest to recognize as the proportion of participants' looks to the target increased with the fastest rate among the three tonal variability types. Both sandhi variability types showed slower increase of looks to the target compared to no sandhi, between which near-merger sandhi was more difficult to process than no-merger sandhi, reflected in the overall less proportion of looks to target in the near-merger sandhi condition.

This chapter is the first study which shows that, prior to the access of the following tonal contexts, the deviated tone itself has already exerted effects on listeners' recognition of speech, lending further evidence for tonal perception that speech perception is an incremental process during which listeners constantly update their perceptual decisions of the incoming speech stimuli. From a methodological perspective, while traditional perception studies on tone sandhi mainly rely on native listeners' meta-linguistic knowledge of whether one tone is changed to another, this chapter is also the first to look into how lexical tones with different types of tone sandhi are processed online. Furthermore, this chapter, from a perceptual point of view, has suggested the necessity of differentiating tonal variability of different types that have been observed in Chapter 3.

Chapter 5 further investigates the nature of tone sandhi perception, as a follow-up study of Chapter 4. This chapter asks whether the online processing of sandhi tones is in a similar way to that of their non-sandhi variants of the same toneme or another lexical tone which has been claimed as the sandhi output due to their *f0* contour similarity. Via the same Visual World Paradigm as in Chapter 4, this chapter examined how listeners process tone sandhi targets when presented with different types of competitors, i.e., a Toneme Competitor which has toneme overlap but no *f0* contour overlap with the target, a Contour Competitor with *f0* contour overlap but no toneme overlap, as well as a Segmental Competitor with neither toneme nor *f0* contour overlap.

Results show that in all target types, participants were sensitive to the types of competitors together presented with the targets. First, Contour Competitors and Segmental Competitors play a comparable role during the initial activation of the target. Both show greater inhibition of the target activation. Therefore, upon hearing the target word, only the underlying toneme of the target word was concurrently activated together with the segments at an earlier stage. Second, at a later stage, contour similarity between the target and the Contour Competitor generally showed a comparable facilitation effect to that of the underlying toneme overlap in the Toneme Competitor condition. This facilitation effect due to *f0* contour overlap was further observed between the target and the Contour Competitor for near-merger sandhi targets.

In previous impressionistic studies, tone sandhi in many Chinese dialects has been commonly regarded as the tonal change from one lexical tone into another lexical tone within the tonal inventory, based on evidence from traditional perceptual studies using offline meta-linguistic tonal discrimination tasks. This chapter is the first study showing, with eye tracking data which reveal the dynamic processing of spoken words, that tone sandhi does not have to be the categorical change from one lexical tone to another. Rather, tonal sandhi variation is better viewed as allophonic variation of the underlying lexical tones and are stored together with the lexical toneme as allotones in the mental lexicon. Together with evidence form the acoustic data, this chapter, from a perception perspective, casts serious doubts on the claimed nature of tone sandhi as the complete tonal neutralization as proposed in the previous impressionistic literature on Tianjin tone sandhi.

Chapter 6 examined how the *f0* realization of neutral tone is affected by global prosodic factors such as prosodic structure. This chapter focused on the rising neutral tone realization before a low-falling tone and asked whether Tianjin Mandarin has a special rising neutral tone target as claimed, and how the *f0* realization of neutral tone is conditioned by different prosodic boundaries.

Via investigating the acoustic realization of the so-called "rising neutral tone" in different prosodic boundary conditions and produced with different preceding discourse contexts, this chapter first shows that in Tianjin Mandarin, the rising neutral tone *f0* contour before T1 does not need to be treated as a special tonal target. Rather, this chapter argues that the raising effect is due to the anticipatory raising effect brought about by the following T1, which is consistent with the general raising effect of T1 towards the preceding lexical full tones as observed in Chapter 3. By increasing the number of neutral tone syllables, a mid-low tonal neutral tone target emerges, which is of a similar nature to that of the neutral tone in Standard Chinese. Furthermore, the results of this chapter suggest an interesting effect of prosodic boundaries on neutral tone *f0* 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, the present chapter 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 lexical tones.

Previous studies have shown that, prosodic boundaries play an important role in conditioning the realization of segments and lexical full tones. This chapter provides the first set of data on how neutral tone realization is affected by different prosodic boundaries. Previous impressionistic studies propose that Tianjin Mandarin has a special rising neutral tone target when the neutral tone is followed by a lexical low-falling tone (Wang & Jiang, 1997; Wang, 2002; Lu & Wang, 2012), challenging our current understanding of neutral tone based on data from Standard Chinese (Chen & Xu, 2006). Results of this chapter suggest that this is not the case. While Tianjin Mandarin and Standard Chinese do differ in terms of the lexical full tone system, where the low-falling tone (T1) is present in Tianjin Mandarin but absent in Standard Chinese, the neutral tone realization in Tianjin Mandarin exhibits the similar mid-low neutral tone target as observed in Standard Chinese. It is the different lexical tone inventories between the two Mandarin varieties that give rise to different neutral tone realization.

In conclusion, with data from well-controlled experiments, this dissertation has demonstrated a rather complex structure of tonal variability in Tianjin Mandarin as well as the perceptual consequences of tonal variability in speech recognition. Several further implications can be made.

First, while most of the previous impressionistic studies have considered tonal variability as the categorical change of tonal identity, by considering multiple factors that might affect the realization of lexical tones in connected speech, this dissertation shows that tonal variation might not be as simple as the lexical tonal identity change. This has cast serious doubts on conclusions made on tonal variability in previous impressionistic studies. On the one hand, this dissertation demonstrates that careful experimental examination of tonal realization can reveal phenomena that have never been reported with impressionistic observations, e.g., the anticipatory raising of a following low-falling tone, which can shed light on theoretically important questions. On the other hand, our results show that some of the variability described in the literature as due to special tonal representations or tonal alternation processes could already be well explained by taking into account multiple factors that affect tonal realization. All of these new findings call for revisiting of the theoretical accounts on tonal alternation based on solid empirical evidence.

Second, in previous studies on tonal variability, much research effort has been made to understand how tones are varied in Standard Chinese. This dissertation demonstrates different yet interesting tonal variability patterns in Tianjin Mandarin. It is to be noted that Tianjin Mandarin is very similar to Standard Chinese not only in terms of the number of lexical tones, but also in terms of the phonemic and the phonotactic systems. Yet the two dialects show very different patterns of tonal variability and surface *f0* realizations. This study thus exemplifies that to gain better insight into the nature of tonal variability cross-linguistically, future studies on different dialects/languages are in great need.

Third, the present dissertation is based on data elicited from a generation that is born in the 1980s. It is possible that what have been reported in the impressionistic studies could be observed in speakers of an older generation. It is therefore important to collect data from older generations in the future to make cross-generation comparisons, results of which would certainly have great implications on language changes.

Last but not least, findings of the present dissertation will also serve as a great testing ground for further work in psycho/neuro-linguistic research to investigate the cognitive and neural mechanisms of lexical tones and speech variability in both production and perception.

REFERENCES

Abramson, A. S. (1977). Noncategorical perception of tone categories in Thai. *Journal of the Acoustical Society of America 61*(1): S66.

Abramson, A. S. (1979). The coarticulation of tones: An acoustic study of Thai. In V. P. T. L. Thongkum, P. Kullavanijaya, & M. R. K. Tingsabadh (Eds.), *Studies in Tai and Mon-Khmer phonetics and phonology: In honour of Eugenie Henderson* (pp. 1-16). Bangkok: Chulalongkorn University Press.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language 38*: 419-439.

Andruski, J. E. (2006). Tone clarity in mixed pitch/phonation-type tones. *Journal of Phonetics 34*: 388-404.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. Retrieved from: <https://cran.r-project.org/web/packages/lme4/index.html>.

Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Phonological structure and phonetic form - Papers in laboratory phonology III* (pp. 7-33). Cambridge: Cambridge University Press.

Beddor, P. S., McGowan, K. B., Boland, J. E., Coetzee, A. W., & Brasher, A. (2013). The time course of perception of coarticulation. *Journal of the Acoustical Society of America 133*(4): 2350-2366.

Boersma, P., & Weenink, D. (2011). *Praat: Doing phonetics by computer*. Retrieved from: <http://www.fon.hum.uva.nl/praat/>.

Brunelle, M. (2009a). Northern and southern Vietnamese tone coarticulation: A comparative case study. *Journal of the Southeast Asian Linguistics Society 1*: 49-62.

Brunelle, M. (2009b). Tone perception in northern and southern Vietnamese. *Journal of Phonetics 37*: 79-96.

Büring, D. (2009). Towards a typology of focus realization. In M. Zimmermann, & C. Féry (Eds.), *Information Structure. Theoretical, Typological, and Experimental Perspectives*. Oxford: Oxford University Press.

Burnham, D., & Jones, C. (2002). Categorical perception of lexical tone by tonal and non-tonal language speakers. In *Proceedings of the 9th Australian International Conference on Speech Science & Technology* (pp. 515-520). Melbourne: Australian Speech Science & Technology Association.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics 26*: 173-199.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE 5*.

Chan, S. W., Chuang, C.-K., & Wang, W. S.-Y. (1975). Cross-linguistic study of categorical perception for lexical tone. *Journal of the Acoustical Society of America 58*(1): S119.

Chang, Y.-C., & Hsieh, F.-F. (2012). Tonal coarticulation in Malaysian Hokkien: A typological anomaly? *The Linguistic Review 29*: 37-73.

Chao, Y.-R. (1920). A system of tone-letters. *Le Maître Phonétique 45*: 24–27.

Chao, Y.-R. (1948). *Mandarin primer*. Cambridge: Harvard University Press.

Chao, Y.-R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chen, A. (2013). *Universal biases in the perception of Mandarin tones, from infancy to adulthood* (PhD thesis). Universiteit Utrecht, the Netherlands.

Chen, A., Liu, L., & Kager, R. (2015). Cross-linguistic perception of Mandarin tone sandhi. *Language Science 48*: 62-69.

Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.

Chen, Y. (2003). *The phonetics and phonology of contrastive focus in Beijing Mandarin* (PhD thesis). State University of New York, Stony Brook, New York.

Chen, Y. (2004). *Focus and intonational phrase boundary in Standard Chinese*. Paper presented at International Symposium on Chinese Spoken Language Processing, Hong Kong, China.

Chen, Y. (2010). Post-focus f0 compression - Now you see it, now you don't. *Journal of Phonetics 38*: 517-525.

Chen, Y. (2012). Tonal variation. In A. C. Cohn, C. Fougeron, & M. K. Huffman_(Eds.), *The Oxford handbook of laboratory phonology* (pp. 103-114). Oxford: Oxford University Press.

Chen, Y. (2015). Qingsheng (Neutral tone). In R. Sybesma (Ed). *Encyclopedia of Chinese language and linguistics*. Brill. Retrieved from: <http://www.brill.com/products/online-resources/encyclopedia-chinese-language-and-linguistics#BIONOTE_1>.

Chen, Y., & Braun, B. (2006). Prosodic realization in information structure categories in Standard Chinese. In R. Hoffmann, & H. Mixdorff (Eds.) *Proceedings of the 3rd International Conference on Speech Prosody*. Dresden: TUD Press.

Chen, Y., & Gussenhoven C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics 36*: 724-746.

Chen, Y., & Gussenhoven, C. (in press). Shanghai Chinese. *Journal of the International Phonetic Association*.

Chen, Y., Shen, R., & Schiller, N. (2011). *Representation of allophonic tone sandhi variants*. Paper presented at Annual Conference on Architectures & Mechanism for Language Processing, Paris, France.

Chen, Y., & Xu, Y. (2006). Production of weak elements in speech - Evidence from f0 patterns of neutral tone in Standard Chinese. *Phonetica 63*: 47-75.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics 32*: 141-176.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics 29*: 115-190.

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics 33*: 121-157.

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech 29*(3): 201-220.

Da, J. (2004). A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. In P. Zhang, T. Xie, & J. Xu (Eds.) *The Studies on the Theory and Methodology of the Digitalized Chinese Teaching to Foreigners: Proceedings of the Fourth International Conference on New Technologies in Teaching and Learning Chinese* (pp. 501-511). Beijing: Tsinghua University Press.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001a). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology 42*: 317-367.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001b). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes 16*(5/6): 507-534.

Duanmu, S. (2000). *The Phonology of Standard Mandarin*. Oxford: Oxford University Press.

Egido, C., & Cooper, W. E. (1980). Blocking of alveolar flapping in speech production: The role of syntactic boundaries and deletion sites. *Journal of Phonetics 8*: 175-184.

Ernestus, M. (2012). Segmental within-speaker variation. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.) *The Oxford handbook of laboratory phonology* (pp. 93-102). Oxford: Oxford University Press.

Fery, C. (2007). Information structural notions and the fallacy of invariant correlates. *Interdisciplinary Studies on Information Structure 6*: 161-184.

Féry, C. (2013). Focus as prosodic alignment. *Natural Language & Linguistic Theory 31*: 683-734.

Féry, C., & Ishihara, S. (2009). How focus and givenness shapes prosody. In M. Zimmermann, & Caroline F. (Eds.) *Information structure. Theoretical, typological, and experimental perspectives*. Oxford: Oxford University Press.

Fougeron, C. (1999). Prosodically conditioned articulatory variations: A review. *U.C.L.A. Working Papers in Phonetics 97*: 1-80.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics 29*: 109-135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America 101*(6): 3728-3740.

Fox, R. A., & Qi, Y.-Y. (1990). Context effect in the perception of lexical tone. *Journal of Chinese Linguistics 18*(2): 261-284.

Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics 65*(7): 1029-1044.

Fromkin, V. A. (Ed.). (1978). *Tone: A linguistic survey*. London: Academic Press.

Frota, S. (2012). Prosodic structure, constituents and their implementation. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.) *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press.

Gandour, J., Ponglorpisit, S., Dechongkit, S., Khunadorn, F., Boongird, P., & Potisuk, S. (1993). Anticipatory tonal coarticulation in Thai noun compounds after unilateral brain damage. *Brain and Language 45*: 1-20.

Gandour, J., Potisuk, S., & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics 22*: 477-492.

Gandour, J., Potisuk, S., Ponglorpisit, S., Dechongkit, S., Khunadorn, F., & Boongird, P. (1996). Tonal coarticulation in Thai after unilateral brain damage. *Brain and Language 52*: 505-535.

Gandour, J. T. (1978). The perception of tone. In V. A. Fromkin (Ed.) *Tone: A linguistic survey* (pp. 41-76). London: Academic Press.

Gussenhoven, C. (2008). Notions and subnotions in information structure. *Acta Linguistica Hungarica 55*(3-4): 381-395.

Han, G. (Ed.) (1993). *Tianjin fangyan* [Tianjin Mandarin]. Beijing, Yanshan Chubanshe.

Han, M. S., & Kim, K.-O. (1974). Phonetic variation of Vietnamese tones in disyllabic utterances. *Journal of Phonetics 2*: 223-232.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal 50*(3): 346-363.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica 137*: 151-171.

Hung, T.-N. (1987). Tianjin tone sandhi: Towards a unified approach. *Journal of Chinese Linguistics 15*(2): 274-305.

Hyman, L. M. (2007). Universals of tone rules: 30 years later. In T. Riad, & C. Gussenhoven (Eds.) *Tones and tunes 1: Typology studies in word and sentence prosody* (pp. 1-34). Berlin: Mouton de Gruyter.

Kabagema-Bilan, E., López-Jiménez, B., & Truckenbrodt, H. (2011). Multiple focus in Mandarin Chinese. *Lingua 121*: 1890-1905.

Karlgren, B. (1915-1926). Etúdes sur la phonologie Chinoise [Studies of Chinese phonology]. In Archives d'Etúdes Orientales. Leiden & Stockholm: E. J. Brill.

Keating, P. A. (2003). *Phonetic encoding of prosodic structure.* Paper presented at the 6th International Seminar on Speech Production, Sydney, Australia.

Kochanski, G., Shih, C., & Jing, H. (2003). Hierarchical structure and word strength prediction of Mandarin prosody. *International Journal of Speech Technology 6*: 33-43.

Kuang, J., & Wang, H. (2006). T3 sandhi at the boundaries of different prosodic hierarchies. *Journal of Chinese Phonetics 1*: 125-131.

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics 35*(3): 301-320.

Ladd, D. R. (2008). *Intonational phonology.* Cambridge: Cambridge University Press.

Ladefoged, P. (1999). American English. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 41-44). International Phonetic Association.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages.* Oxford: Blackwell Publishers.

Lee, C.-Y. (2007). Does horse activate mother? Processing lexical tone in form priming. *Language and Speech 50*(1): 101-123.

Lee, W.-S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association 33*(1): 109-112.

Lee, W.-S., & Zee, E. (2014). Chinese phonetics. In C.-T. J. Huang, Y.-H. A. Li, & A. Simpson (Eds.) *The handbook of Chinese linguistics* (pp. 369-399). Wiley Blackwell.

Lee-Kim, S.-I. (2014). Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association 44*(03): 261-282.

Li, Q., & Chen, Y. (2011). *Neutral tone realization in Tianjin Mandarin*. Paper presented at 17th International Congress of Phonetic Sciences, Hong Kong, China.

Li, Q. & Chen, Y. (2012). *Trisyllabic tone sandhi in Tianjin Mandarin*. Paper presented at the Third International Symposium on Tonal Aspects of Languages, Nanjing, China.

Li, Q., & Chen, Y. (2015). Effect of contextual tonal variation on speech recognition: Evidence from eye movements. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, UK.

Li, Q., & Chen, Y. (2016). An acoustic study of contextual tonal variation in Tianjin Mandarin. *Journal of Phonetics 54*: 123-150.

Li, Q., & Chen, Y. (under review). Prosodically conditioned neutral tone realization in Tianjin Mandarin.

Li, X., & Chen, Y. (in press). Representation and processing of lexical tone and tonal variants: Evidence from the mismatch negativity. *PLoS ONE*.

Li, X., & Liu, S. (1985). Tianjin fangyan de liandu biandiao [Tone sandhi in Tianjin Mandarin]. *Zhongguo Yuwen 1*: 76-80.

Liang, L. (2008). *Hanyu zhonghediao de kuafangyan yanjiu* [A cross-linguistic study of neutral tone in Chinese]. Tianjin: Nankai Daxue Chubanshe.

Lin, H.-B., & Repp, B. H. (1989). Cues to the perception of Taiwanese tones. *Haskins Laboratories Status Report on Speech Research SR-99/100*: 137-147.

Lin, M., & Wang, W. S. (1985). Shengdiao ganzhi wenti [Tone perception]. *Zhongguo Yuyan Xuebao 2*: 59-69.

Lin, M., & Yan, J. (1980). Beijinghua qingsheng de shengxue xingzhi [The acoustic properties of neutral tone in Beijing Mandarin]. *Fangyan 3*: 166-178.

Lin, Y.-H. (2007). *Sounds of Chinese*. Cambridge: Cambridge University Press.

Lin, Y.-H. (2014). Segmental phonology. In C.-T. J. Huang, Y.-H. A. Li, & A. Simpson (Eds.) *The handbook of Chinese linguistics* (pp. 400-421). Wiley Blackwell.

Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica 62*: 70-87.

Liu, S., & Samuel, A. G. (2007). The role of Mandarin lexical tones in lexical access under different contextual conditions. *Language and Cognitive Processes 22*(4): 566-594.

Liu, Y., & Gao, J. (2003). Qu-qu liandu biandiao guize: Shehui yuyanxue bianxiang [T4-T4 tone sandhi rule: Sociolinguistic variation]. *Tianjin Shifan Daxue Xuebao (Shehui Kexue Ban) 5*: 65-69.

Lu, J. (1997). Tianjin fangyan zhong de yizhong xinde liandu biandiao [A new tone sandhi rule in Tianjin Mandarin]. *Tianjin Shida Xuebao 4*: 67-72.

Lu, J., & Wang, J. (2012). *Hanyu qingsheng de youxuanlun fenxi* [An optimality account of neutral tone in Chinese]. Tianjin: Tianjin Daxue Chubanshe.

Lu, Y. (1995). *Putonghua de qingsheng he erhua* [Neutral tone and Erhua in Standard Chinese]. Beijing: Shangwu Yinshuguan.

Ma, Q. (2005). "Tianjinhua liandu biandiao zhi mi" de youxuanlun jieshi [An optimality account of "the paradox of Tianjin tone sandhi"]. *Zhongguo Yuwen 6*: 561-568.

Ma, Q., & Jia, Y. (2006). Tianjinhua shangsheng de liangtiao "biandiao guize" bianxi [An analysis of two "sandhi rules" of T3 in Tianjin Mandarin]. *Tianjin Shifan Daxue Xuebao (Shehui Kexue Ban) 1*: 53-58.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language 62*: 407-420.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition 8*(1): 1-71.

Matin, E., Shao, K. C., & Boff, H. R. (1993). Saccadic overhead: Informational-processing time with and without saccades. *Perception & Psychophysics 53*(4): 372-380.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology 18*: 1-86.

McGuire, G. (2010). A Brief Primer on Experimental Designs for Speech Perception. Unpublished manuscript. Retrieved from: <http://people.ucsc.edu/~gmcguir1/experiment_designs.pdf>.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition 86*: B33-B42.

Mei, T.-L. (1977). Tones and tone sandhi in 16th century Mandarin. *Journal of Chinese Linguistics 5*: 237-260.

Miellet, S., O'Donnell, P., & Sereno, S. C. (2009). Parafoveal magnification. *Psychological Science 20*(6): 721-728.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton: CRC Press.

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America 102*(3): 1864-1877.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.

Nixon, J. S., Chen, Y., & Schiller, N. O. (2015). Multi-level processing of phonetic variants in speech production and visual word processing: Evidence from Mandarin lexical tones. *Language, Cognition and Neuroscience 30*(5): 491-505.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition 52*(3): 189-234.

Palmer, A. (1969). Thai tone variants and the language teachers. *Language Learning - A Journal of Applied Linguistics 19*(3-4): 287-299.

Pan, H.-H. (2007). The effects of prosodic boundaries on nasality in Taiwan Min. *Journal of the Acoustical Society of America 121*(6): 3755-3769.

Pan, H.-H., & Tai, Y.-H. (2006). Boundaries and tonal articulation in Taiwanese Min. In R. Hoffmann, & H. Mixdorff (Eds.) *Proceedings of the 3rd International Conference on Speech Prosody*. Dresden: TUD Press.

Peng, S.-H. (1996). *Phonetic implementation and perception of place coarticulation and tone sandhi* (PhD thesis). The Ohio State University, Ohio.

Peng, S.-H. (1997). Production and perceptioin of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics 25*: 371-400.

Potisuk, S., Gandour, J., & Harper, M. P. (1997). Contextual variation in trisyllabic sequences of Thai tones. *Phonetica 54*: 22-42.

Prom-on, S., Liu, F., & Xu, Y. (2012). Post-low bouncing in Mandarin Chinese: Acoustic analysis and computational modeling. *Journal of the Acoustical Society of America 132*(1): 421-432.

R Core Team (2014). *R: A language and environment for statistical computing*. Retrieved from: <https://www.r-project.org>.

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication 6*(4): 343-352.

Schirmer, A., Tang, S.-L., Penney, T. B., Gunter, T. C., Chen, H.-C. (2005). Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience 17*(1): 1-12.

Scholz, F. (2012). *Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese* (PhD thesis). Universiteit Leiden, the Netherlands.

Scholz, F., & Chen, Y. (2014a). The independent effects of prosodic structure and information status on tonal coarticulation: Evidence from Wenzhou Chinese. In J. Caspers, Y. Chen, W. Heeren, J. Pacilly, N. O. Schiller, & E. van Zanten (Eds.) *Above and beyond the segments: Experimental linguistics and phonetics* (pp. 275-287). John Benjamins Publishing Company.

Scholz, F., & Chen, Y. (2014b). Sentence planning and f0 scaling in Wenzhou Chinese. *Journal of Phonetics 47*: 81-91.

Selkirk, E. (1995). Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith (Ed.) *The handbook of phonological theory* (pp. 550-569). London: Blackwell.

Selkirk, E. (2001). The syntax-phonology interface. In N. J. Smelser, & P. B. Baltes (Eds.) *International encyclopedia of the social & behavioral sciences*. Elsevier.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research 25*(2): 193-247.

Shen, J., Deutsch, D., & Rayner, K. (2013). On-line perception of Mandarin tones 2 and 3: Evidence from eye movements. *Journal of the Acoustical Society of America 133*(5): 3016-3029.

Shen, X. S. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics 18*: 281-295.

Shen, X. S. (1991). Question intonation in natural speech: A study of Changsha Chinese. *Journal of the International Phonetic Association 21*: 19-28.

Shen, X. S. (1992). On tone sandhi and tonal coarticulation. *Acta Linguistica Hafniensia: International Journal of Linguistics 25*(1): 83-94.

Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech 34*(2): 145-156.

Shen, X. S., Lin, M., & Yan, J. (1993). f0 turning point as an f0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustical Society of America 4*(1): 2241-2243.

Shi, F. (1986). Tianjin fangyan shuangzizu shengdiao fenxi [An analysis of disyllabic tonal combination in Tianjin Mandarin]. *Yuyan Yanjiu 1*: 77-90.

Shih, C. (1986). *The prosodic domain of tone sandhi in Chinese* (PhD thesis). University of California, San Diego, California.

Shih, C. (1987). The Phonetics of the Chinese Tonal System. Unpublished manuscript. Retrieved from: <http://www.researchgate.net/publication/247944578_The_phonetics_of_the_Chinese_tonal_system>.

Shih, C. (1997). Mandarin third tone sandhi and prosodic structure. In J. Wang, & N. Smith (Eds.) *Studies in Chinese phonology* (pp. 81-123). Berlin: Mouton de Gruyter.

Shih, C., & Konchanski, G. P. (2000). *Chinese tone modeling with Stem-ML*. Paper presented at the International Conference on Spoken Language Processing, Beijing, China.

Speer, S. R., Shih, C., & Slowlaczek, M. L. (1989). Prosodic structure in language understanding: Evidence from tone sandhi in Mandarin. *Language and Speech 32*(4): 337-354.

Spivey, M. J. (2007). *The continuity of mind*. New York: Oxford University Press.

Tabain, M. (2003). Effects of prosodic boundary on /aC/ sequences: Acoustic results. *Journal of the Acoustical Society of America 113*(1): 516-531.

Tan, F. (1987). Tone sandhi in the Tianjin dialect. *Journal of Chinese Linguistics 15*(2): 228-246.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research 29*(6): 557-580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science 268*: 1632-1634.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry 30*: 219-255.

van Santen, J., Shih, C., & Möbius, B. (1997). Intonation. In R. Sproat (Ed.) *Multilingual text-to-speech synthesis: The Bell labs approach*. Springer.

Wang, B., & Xu, Y. (2011). Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics 39*: 595-611.

Wang, H. S. (2002). The prosodic effects on Taiwan Min tones. *Language and Linguistics 3*(4): 839-852.

Wang, J. (2002). Youxuanlun he tianjinhua de liandu biandiao ji qingsheng [Optimility theory and tone sandhi and neutral tone in Tianjin Mandarin]. *Zhongguo Yuwen 4*: 363-372.

Wang, J., & Jiang, H. (1997). Tianjinhua qingsheng de yuyin xingzhi he yinxi fenxi [The phonetic properties and phonological analysis of neutral tone in Tianjin Mandarin]. In *Qinzhu zhongguo shehui kexueyuan yuyan yanjiusuo jiansuo sishiwu zhounian xueshu lunwenji.*

Wang, W. S., & Li, K.-P. (1967). Tone 3 in Pekinese. *Journal of Speech and Hearing Research 10*: 629-636.

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

Weber, A., Melinger, A., & Tapia, L. L. (2007). *The mapping of phonetic information to lexical representations in Spanish: Evidence from eye movements.* Paper presented at 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Wee, L.-H., Yan, X., & Chen, M. Y. (2005). *Yinan yu luxiang – Lun Tianjin fangyan de liandu biandiao* [Problems and solutions – On tone sandhi in Tianjin Mandarin]. Beijing: Shangwu Yinshuguan.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America 91*(3): 1707-1717.

Wu, Z. (1985). Putonghua sanzizu biandiao guilü [Rules of trisyllabic tone variation in Standard Chinese]. *Zhongguo Yuyan Xuebao 2*.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience 170*: 223-231.

Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America 95*(4): 2240-2253.

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

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics 27*: 55-105.

Xu, Y. (2001). Sources of tonal variations in connected speech. *Journal of Chinese Linguistics Monograph Series 17*: 1-31.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America 120*(2): 1063-1074.

Yang, Y., & Wang, B. (2002). *Acoustic correlates of hierarchical prosodic boundary in Mandarin.* Paper presented at the 1st International Conference on Speech Prosody, Aix-en-Provence, France.

Yip, M. (1980). *The tonal phonology of Chinese* (PhD thesis). Massachusetts Institute of Technology, Massachusetts.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Yu, A. C. L. (2007). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology 24*: 187-214.

Yuan, J. (2004). *Perception of Mandarin intonation.* Paper presented at International Symposium on Chinese Spoken Language Processing, Hong Kong, China.

Yuan, J., & Chen, Y. (2014). 3rd tone sandhi in Standard Chinese: A corpus approach. *Journal of Chinese Linguistics 42*(1): 218-236.

Zee, E. (1980). A spectrographic investigation of Mandarin tone sandhi. *UCLA Working Papers 49*: 98-116.

Zhang, C., Xia, Q., & Peng, G. (2014). Mandarin third tone sandhi requires more effortful phonological encoding in speech production: Evidence from an ERP study. *Journal of Neurolinguistics 33*: 149-162.

Zhang, J. (2010). Issues in the analysis of Chinese tone. *Language and Linguistics Compass 4*(12): 1137-1153.

Zhang, J., & Liu, J. (2010). *Patterns of tone sandhi productivity in Tianjin Chinese*. Paper presented at 160th Meeting Acoustical Society of America, Cancun, Mexico.

Zhang, J., & Liu, J. (2011). Tone sandhi and tonal coarticulation in Tianjin Chinese. *Phonetica 68*: 161-191.

Zhang, L., Xi, J., Wu, H., Shu, H., & Li, P. (2012). Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition. *Neuroreport 23*(1): 35-39.

Zhang, N. (1997). The avoidance of the third tone sandhi in Mandarin Chinese. *Journal of East Asian Linguistics 6*: 293-338.

Zhang, Z.-S. (1987). The paradox of Tianjin: Another look. *Journal of Chinese Linguistics 15*(2): 228-246.

Zhou, X., & Marslen-Wilson, W. (1997). The abstractness of phonological representation in the Chinese mental lexicon. In H.-C. Chen (Ed.) *Cognitive processing of Chinese and related Asian languages* (pp. 3-26). Hong Kong: The Chinese University Press.

Zsiga, E. (2012). Contrastive tone and its implementation. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.) *The Oxford handbook of laboratory phonology* (pp. 196-207). Oxford: Oxford University Press.

Zsiga, E., & Nitisaroj, R. (2007). Tone features, tone perception, and peak alignment in Thai. *Language and Speech 50*(3): 343-383.

ENGLISH SUMMARY

As an important means of our daily communication, spoken language is notoriously rich in its variability due to various factors. Previous studies have paid much attention to how people produce and perceive different types of segmental variation typically in non-tonal languages. The majority of the world's languages, however, are tonal languages, in which pitch patterns over syllables distinguish lexical meanings. Much less, however, has been investigated and understood on how suprasegmental features such as lexical tones vary in their acoustic realizations in different contexts in speech production and how the variabilities further affect listeners' perception of the lexical tones and the recognition of speech. This dissertation provides a comprehensive study of both local and global tonal variability in Tianjin Mandarin with a series of well-controlled experiments. Tianjin Mandarin is well-known for its intriguing tonal variation patterns, and therefore presents as a good test case for a better understanding of the tonal variability phenomenon.

The dissertation consists of seven chapters. **Chapter 1** briefly introduces the main research issues to be discussed in this dissertation, followed by a description of the phonological system of Tianjin Mandarin in **Chapter 2.**

Chapter 3 sets out to investigate the *f0* variability induced by two tonal variation processes, i.e., tone sandhi and tonal coarticulation, which are known to greatly vary the *f0* realization of lexical tones in a local domain. Previous impressionistic studies have proposed complex patterns of tonal sandhi patterns over disyllabic constituents, as well as the conflicting directionality of sandhi alternations and possible iterative application of disyllabic sandhi rules over trisyllabic constituents. Via examining *f0* variation of lexical tones as a function of different following tones, this chapter observed an interesting anticipatory raising effect of the lexical low-falling tone (T1) which has never been mentioned in the literature. Furthermore, this chapter confirmed three tonal sequences with tone sandhi changes: 1) two low-falling tones (T1T1), 2) a high-falling tone followed by a low-falling tone (T4T1), and 3) two low-dipping tones (T3T3). These disyllabic tone sandhi patterns, however, were not consistently observed within trisyllabic sequences, as claimed in the literature. Specifically, while T3T3 showed sandhi application regardless of its position within a trisyllabic sequence, T1T1 and T4T1 sequences showed sandhi changes only when they were right aligned. Last but not least, no tonal neutralization was observed over any of the sandhi processes. Rather, this chapter argues that T3T3 is a near-merger sandhi case, while T1T1 and T4T1 can be classified as no-merger sandhi sequences.

Chapters 4 and 5 investigate the perceptual consequences of the local tonal variability that has been discussed in Chapter 3. **Chapter 4**, via the visual world paradigm, examines the effect of contextual tonal variation on speech recognition in Tianjin Mandarin, where three types of contextual tonal variation have been observed: Near-Merger Sandhi, No-Merger Sandhi, and No-Sandhi tonal coarticulation. Listeners were asked to identify the target speech amongst an array of four possibilities upon hearing a disyllabic collocation, while their eye movements were tracked. Results suggest that native listeners are sensitive to fine-grained phonetic details in contextual variation of lexical tones. No-Sandhi

Coarticulation was the easiest to recognize as participants fixated on the targets the earliest among the three conditions. Between the two sandhi variation types, Near-Merger Sandhi was more difficult to process than No-Merger Sandhi, reflected in the overall less proportion of looks to target. To follow up, **Chapter 5** further examines the nature of tone sandhi perception. Previous perception studies with traditional meta-linguistic tonal judgement tasks claimed great perceptual ambiguity between sandhi-derived tones and another lexical tone in the tonal system, suggesting a complete lexical tonal neutralization due to tone sandhi. This chapter, via the same Visual World Paradigm as in Chapter 4, shows that the sandhi-derived tones are in fact perceived more similarly to their non-sandhi variants of the same toneme rather than another lexical tone that is claimed to be the output tone of sandhi change according to rules reported in previous impressionistic studies. This chapter, therefore, from a perceptual perspective, confirmed that tone sandhi does not involve the categorical change of one lexical tone to another. Rather, sandhi derived tones are better viewed as just a variant of the lexical tone that undergoes tone sandhi change.

Based upon findings for the local source of tonal variability as discussed in Chapters 3-5, **Chapter 6** sets out to further understand how global factors might influence the tonal *f0* realization. This chapter examined the effect of prosodic boundaries on the neutral tone realization in Tianjin Mandarin. Neutral tone refers to the prosodically weak element which shows much reduced acoustic properties and great variability in the *f0* realization. The neutral tone in Tianjin Mandarin exhibits interesting context-dependent *f0* realization, where the neutral tone exhibits consistent rising *f0* realization preceding a low-falling lexical tone. The results showed that the rising neutral tone *f0* contour before T1 does not need to be treated as a special tonal target as claimed in the literature, but rather should be due to the raising effect brought about by the following T1 as first proposed in Chapter 3. It is shown that the neutral tone in Tianjin Mandarin has a mid-low neutral tone target similar to that in Standard Chinese. In addition, by manipulating the prosodic boundary between the neutral tone and the following T1, an interesting effect of prosodic boundary on the neutral tone *f0* realization was observed. 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, this chapter shows an interaction of prosodic boundary and focus, which further argues for the independent implementation of the two mechanisms.

Chapter 7 revisits the research questions and concludes the main findings in this dissertation. This chapter also discusses the limitations and provides suggestions for future research.

SAMENVATTING IN HET NEDERLANDS

Als een belangrijk middel voor onze dagelijkse communicatie is gesproken taal bijzonder rijk aan variabiliteit dankzij verscheidene factoren. Eerdere studies hebben veel aandacht besteed aan de wijze waarop mensen verschillende soorten segmentele variatie produceren en waarnemen in typisch niet-toontalen. De meerderheid van de talen van de wereld wordt echter gevormd door toontalen, waarin tonale patronen over lettergrepen lexicale betekenissen onderscheiden. Er is echter veel minder onderzocht en bekend hoe suprasegmentele kenmerken zoals lexicale tonen variëren in akoestische realisaties in verschillende contexten van spraakproductie en hoe deze variabiliteiten verder invloed hebben op de perceptie van lexicale tonen en de herkenning van spraak door luisteraars. Dit proefschrift levert een diepgaande studie van zowel lokale als globale toonvariatie in het Tianjin-Mandarijn middels een serie zorgvuldig gecontroleerde experimenten. Het Tianjin-Mandarijn is berucht om zijn intrigerende patronen van tonale variatie en leent zich daarom als een goede testcase om een beter begrip te bereiken van het fenomeen van tonale variatie.

Het proefschrift bestaat uit zeven hoofdstukken. **Hoofdstuk 1** biedt een korte inleiding tot de belangrijkste onderzoekskwesties die in het proefschrift besproken worden, gevolgd door een beschrijving van het fonologische systeem van het Tianjin-Mandarijn in **Hoofdstuk 2**.

Hoofdstuk 3 onderzoekt de variabiliteit in *f0* als gevolg van twee processen van tonale variatie, nl. toonsandhi en tonale coarticulatie, waarvan beide bekend is dat ze grote invloed hebben op de *f0*-realisatie van lexicale tonen binnen een lokaal domein. Eerdere impressionistische studies hebben complexe patronen van toonsandhi voorgesteld over tweelettergrepige constituenten en mogelijke iteratieve toepassing van tweelettergrepige sandhiregels op drielettergrepige constituenten. Door *f0*-variatie van lexicale tonen te onderzoeken als een functie van de verschillende volgende tonen, observeert dit hoofdstuk een interessant anticipatorisch verhogingseffect op de lexicale laag-vallende toon (T1) dat niet eerder is opgemerkt in de literatuur. Daarnaast bevestigt dit hoofdstuk drie sequenties van tonen waarin veranderingen optreden door toonsandhi: 1) twee laag-vallende tonen (T1T1), 2) een hoog-vallende toon gevolgd door een laag-vallende toon (T4T1), en 3) twee laag-dippende tonen (T3T3). Deze tweelettergrepige toonsandhipatronen werden echter niet consistent gevonden binnen drielettergrepige sequenties, zoals beweerd in de literatuur. Specifiek vertoonde T3T3 toepassing van sandhi onafhankelijk van zijn positie binnen een drielettergrepige sequentie, maar vertoonden T1T1- en T4T1-sequenties alleen sandhi-aanpassingen aan het rechter uiteinde van de drielettergrepige sequentie. Als laatste werd er geen tonale neutralisatie geobserveerd over enig sandhiproces. In plaats daarvan stelt dit hoofdstuk dat T3T3 een geval is van *near-merger*-sandhi, terwijl T1T1 en T4T1 geclassificeerd worden als *no-merger*-sandhisequenties.

Hoofdstuk 4 en 5 onderzoeken de perceptuele consequenties van de lokale toonvariabiliteit die wordt besproken in hoofdstuk 3. **Hoofdstuk 3** onderzoekt, middels het *visual world paradigm*, het effect van contextuele tonale variatie op de herkenning van spraak in het Tianjin-Mandarijn, waar drie soorten contextuele tonale variatie geobserveerd zijn: *Near-Merger*-sandhi, *No-Merger*-sandhi, en *No-Sandhi* tonale coarticulatie. Luisteraars werden gevraagd om, na het horen van een tweelettergrepige collocatie, de bedoelde spraak te identificeren uit een verzameling van vier mogelijkheden, terwijl hun oogbewegingen gevolgd werden. De resultaten suggereren dat *native* luisteraars gevoelig zijn voor fijnmazige fonetische details in de contextuele variatie van lexicale tonen. *No-Sandhi*-coarticulatie werd het gemakkelijkst herkend, aangezien proefpersonen zich daar het snelst fixeerden op de doelen onder de drie condities. Tussen de twee typen sandhi-variatie was *Near-Merger*-sandhi moeilijker te verwerken dan *No-Merger*-sandhi, wat gereflecteerd werd in de algehele lagere proportie van fixaties op het doel. **Hoofdstuk 5** volgde dit op met verder onderzoek naar de aard van de perceptie van toonsandhi. Eerdere perceptiestudies met traditionele meta-linguïstische toonbeoordelingstaken rapporteren grote perceptuele ambiguïteit tussen tonen die afgeleid zijn van sandhi en andere tonen die aanwezig zijn in het tonale systeem, hetgeen een complete neutralisatie van lexicale tonen suggereert als gevolg van toonsandhi. Dit hoofdstuk toont aan, via het zelfde *visual world paradigm* als in hoofdstuk 4, dat de tonen afgeleid van sandhi echter als dichter bij de non-sandhi-variant van het zelfde toneem waargenomen worden, dan als een andere lexicale toon die beweerd wordt de outputtoon van sandhiverandering te zijn volgens de regels die eerdere impressionistische studies rapporteren. Dit hoofdstuk toont daarom aan, vanuit een perceptueel perspectief, dat het bij toonsandhi niet om de categorische verandering gaat van de ene lexicale toon naar de andere. In plaats daarvan laten door sandhi afgeleide tonen zich beter beschrijven als eenvoudigweg een variant van de lexicale toon die verandering door toonsandhi ondergaat.

Gebaseerd op de bevindingen voor de lokale oorzaak van toonvariabiliteit zoals besproken in hoofdstuk 3-5 probeert **Hoofdstuk 6** verder te begrijpen hoe globale factoren invloed kunnen hebben op de realisatie van de tonale *f0*. Dit hoofdstuk onderzoekt het effect van prosodische grenzen op de realisatie van de neutrale toon in het Tianjin-Mandarijn. *Neutrale toon* verwijst naar het prosodisch zwakke element dat sterk gereduceerde akoestische eigenschappen en grote variabiliteit in *f0*-realisatie vertoont. In het Tianjin-Mandarijn vertoont de neutrale toon interessante contextgebonden *f0*-realisatie, inhoudende dat neutrale toon consequent een stijgende *f0*-realisatie vertoont voor een laag-vallende lexicale toon. De resultaten laten zien dat de stijgende *f0*-contour van neutrale toon voor T1 niet gezien hoeft te worden als een speciaal tonaal doel zoals beweerd in de literatuur, maar verklaard dient te worden uit het stijgende effect veroorzaakt door de volgende T1 zoals eerst voorgesteld in hoofdstuk 3. Aangetoond wordt dat de neutrale toon in het Tianjin-Mandarijn een middenlaag neutraal toondoel heeft dat lijkt op dat in het Standaardchinees. Daarnaast wordt er een interessant effect geobserveerd van prosodische begrenzing op de realisatie van de *f0* van de neutrale toon, door de prosodische grens tussen een de neutrale toon en de volgende T1 te manipuleren. Specifiek kan het stijgende effect van de volgende T1 geblokkeerd raken door een hooggelegen prosodische grens (d.w.z. een Subject-Predicaat-grens) gevolgd door een laaggelegen grens (d.w.z. een grens onder de NP). Als laatste legt dit hoofdstuk een interactie bloot van de prosodische grens en focus, hetgeen verder pleit voor een onafhankelijke implementatie van de beide mechanismen.

Hoofdstuk 7 gaat terug naar de onderzoeksvragen en besluit de hoofdbevindingen in het proefschrift. Dit hoofdstuk bespreekt ook de beperkingen van het onderzoek en draagt suggesties aan voor verder onderzoek.

APPENDICE

Appendix I

The size and location of the visual stimuli in the Visual World Paradigm were calculated as follows, as illustrated in Figure I. As shown in Figure I*a*, the diameter (D) of the 5° visual circle (parafoveal visual region, as illustrated with circles in Figure I) on the screen was calculated according to the formula: $D = 2 \times d \times \tan\frac{\theta}{2}$. Therefore, given the distance from the eyes to the monitor (*d*=69cm), the diameter of each circle should be 6.03cm.

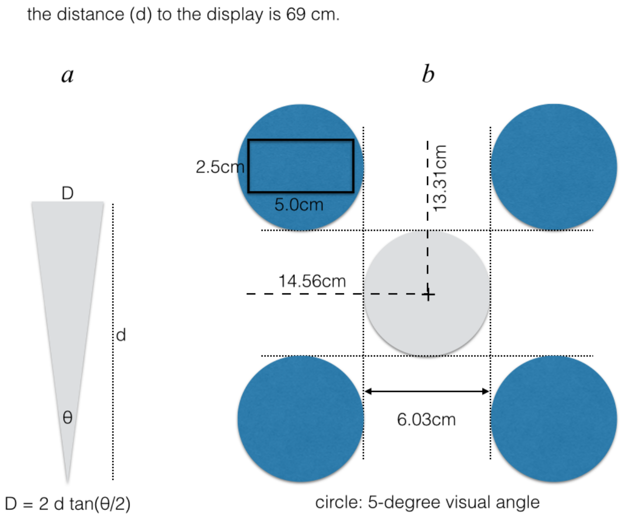

Figure I *Illustration of the size and location of the disyllabic words.*

As shown in Figure I, the fixation cross was located at the very center of the screen (as illustrated with a black cross in Figure I). The screen was then divided into nine regions according to the central circle with the fixation cross as the center. The circles for the four words were located at the four corners respectively.

Each disyllabic word was typed in Song, and resized into a 2.5cm * 5cm rectangle (as illustrated with the solid black rectangle in Figure I), which was maximally fit into the circle; the farther side edge of the words were 14.56cm to the left or right of the screen central vertical lines, and 13.31cm to the upper or lower of the screen central horizontal lines.

Appendix II

The full lists of disyllabic and trisyllabic words used in Chapter 3 are given below.

II-1. Disyllabic word list.

Tonal Sequences	Morpho-syntactic Structures		
	Compound	Modifier+Head	Verb+Object
T1T1	爹妈 die^{1} ma^{1} *parents*	家猫 jia^{1} mao^{1} *domestic cat*	弯腰 wan^{1} yao^{1} *to bend down*
T1T2	规模 gui^{1} mo^{2} *scale*	歌迷 ge^{1} mi^{2} *fans*	猜谜 cai^{1} mi^{2} *to guess a riddle*
T1T3	枢纽 shu^{1} niu^{3} *hub*	糙米 cao^{1} mi^{3} *brown rice*	推理 tui^{1} li^{3} *to reason*
T1T4	归纳 gui^{1} na^{4} *to conclude*	闺蜜 gui1 mi^{4} *girlfriends*	揭密 jie^{1} mi^{4} *to reveal a secret*
T2T1	拿捏 na^{2} nie^{1} *to ponder*	白猫 bai^{2} mao^{1} *white cat*	行医 xing^{2} yi^{1} *to practice medicine*
T2T2	皮毛 pi^{2} mao^{2} *fur*	活埋 huo^{2} mai^{2} *to bury alive*	回眸 hui^{2} mou^{2} *to look back*
T2T3	涂抹 tu^{2} mo^{3} *to smear*	鸵鸟 tuo^{2} niao^{3} *ostrich*	淘米 tao^{2} mi^{3} *to wash rice*
T2T4	稠密 chou^{2} mi^{4} *dense*	胡闹 hu^{2} nao^{4} *to make trouble*	夺目 duo^{2} mu^{4} *dazzling*
T3T1	扭捏 niu^{3} nie^{1} *coy*	奶妈 nai^{3} ma^{1} *nanny*	爽约 shuang^{3} yue^{1} *to fail to keep an appointment*
T3T2	楷模 kai^{3} mo^{2} *model*	铁牛 tie^{3} niu^{2} *tractor*	起锚 qi^{3} mao^{2} *to raise anchor*
T3T3	子女 zi^{3} nü^{3} *child*	舞女 wu^{3} nü^{3} *dancing girl*	洗脑 xi^{3} nao^{3} *to brainwash*
T3T4	美妙 mei^{3} miao^{4} *wonderful*	果木 guo^{3} mu^{4} *fruit tree*	把脉 ba^{3} mai^{4} *to take one's pulse*

T4T1	触摸 chu⁴ mo¹ *to touch*	后妈 hou⁴ ma¹ *step mother*	喂猫 wei⁴ mao¹ *to feed the cat*
T4T2	计谋 ji⁴ mou² *stratagem*	亚麻 ya⁴ ma² *linen*	蓄谋 xu⁴ mou² *to premeditate*
T4T3	购买 gou⁴ mai³ *to buy*	密码 mi⁴ ma³ *password*	喂奶 wei⁴ nai³ *to breastfeed*
T4T4	咒骂 zhou⁴ ma⁴ *to curse*	内幕 nei⁴ mu⁴ *inside story*	盗墓 dao⁴ mu⁴ *to rob a grave*

II-2. Trisyllabic list 1: σσσ.

Tonal Sequences	Word 1	Word 2
T1T1T1	乌拉圭 wu¹ la¹ gui¹ *Uruguay*	撒哈拉 sa¹ ha¹ la¹ *Sahara*
T1T1T2	西班牙 xi¹ ban¹ ya² *Spain*	阿拉伯 a¹ la¹ bo² *Arab*
T1T1T3	希拉里 xi¹ la¹ li³ *Hilary*	巴哈马 ba¹ ha¹ ma³ *Bahamas*
T1T1T4	伊拉克 yi¹ la¹ ke⁴ *Iraq*	科威特 ke¹ wei¹ te⁴ *Kuwait*
T1T2T1	昆德拉 kun¹ de² la¹ *Kundera*	科伦坡 ke¹ lun² po¹ *Colombo*
T1T2T2	科摩罗 ke¹ mo² luo² *Comoros*	巴格达 ba¹ ge² da² *Baghdad*
T1T2T3	巴拿马 ba¹ na² ma³ *Panama*	多明我 duo¹ ming² wo³ *Domingo*
T1T2T4	加拿大 jia¹ na² da⁴ *Canada*	匈牙利 xiong¹ ya² li⁴ *Hungary*
T1T3T1	安卡拉 an¹ ka³ la¹ *Ankara*	詹姆斯 zhan¹ mu³ si¹ *James*

T1T3T2	苏里南 su1 li3 nan2 *Surinam*	多里达 duo1 li3 da2 *Derrida*
T1T3T3	托马里 tuo1 ma3 li3 *Tomari*	扎卡里 zha1 ka3 li3 *Zachry*
T1T3T4	多普勒 duo1 pu3 le4 *Doppler*	冈比亚 gang1 bi3 ya4 *Gambia*
T1T4T1	菲律宾 fei1 lv4 bin1 *Philippine*	阿克拉 a1 ke4 la1 *Accra*
T1T4T2	喀麦隆 ka1 mai4 long2 *Cameroon*	乌克兰 wu1 ke4 lan2 *Ukraine*
T1T4T3	哈勒姆 ha1 le4 mu3 *Haarlem*	安道尔 an1 dao4 er3 *Andorra*
T1T4T4	伽利略 jia1 li4 lüe4 *Galilei*	巴洛克 ba1 luo4 ke4 *Baroque*
T2T1T1	尼加拉 ni2 jia1 la1 *Niagara*	达拉斯 da2 la1 si1 *Dallas*
T2T1T2	柏拉图 bo2 la1 tu2 *Plato*	德黑兰 de2 hei1 lan2 *Tehran*
T2T1T3	博拉美 bo2 la1 mei3 *Polamer*	卢森堡 lu2 sen1 bao3 *Luxembourg*
T2T1T4	黎巴嫩 li2 ba1 nen4 *Lebanon*	明斯克 ming2 si1 ke4 *Minsk*
T2T2T1	吉隆坡 ji2 long2 po1 *Kuala Lumpur*	俄罗斯 e2 luo2 si1 *Russia*
T2T2T2	婆罗门 po2 luo2 men2 *Brahman*	迪德罗 di2 de2 luo2 *Diderot*
T2T2T3	德雷姆 de2 lei2 mu3 *Dremu*	尼泊尔 ni2 bo2 er3 *Nepal*
T2T2T4	尤迪特 you2 di2 te4 *Udeet*	格兰特 ge2 lan2 te4 *Grant*

T2T3T1	牙买加 ya^2 mai^3 jia^1 *Jamaica*	莱比锡 lai^2 bi^3 xi^1 *Leipzig*
T2T3T2	梅里达 mei^2 li^3 da^2 *Merida*	图瓦卢 tu^2 wa^3 lu^2 *Tuvalu*
T2T3T3	劳卡鲁 lao^2 ka^3 lu^3 *Lokarus*	伯里瓦 bo^2 li^3 wa^3 *Bolivar*
T2T3T4	伯努利 bo^2 nu^3 li^4 *Bernoulli*	维也纳 wei^2 ye^3 na^4 *Vienna*
T2T4T1	摩纳哥 mo^2 na^4 ge^1 *Monaco*	伯利兹 bo^2 li^4 zi^1 *Belize*
T2T4T2	罗纳德 luo^2 na^4 de^2 *Ronald*	卢旺达 lu^2 wang^4 da^2 *Rwanda*
T2T4T3	伯克里 bo^2 ke^4 li^3 *Berkeley*	尼日尔 ni^2 ri^4 er^3 *Niger*
T2T4T4	维克特 wei^2 ke^4 te^4 *Victor*	华盛顿 hua^2 sheng^4 dun^4 *Washington*
T3T1T1	塔波拉 ta^3 bo^1 la^1 *Tabora*	美因兹 mei^3 yin^1 zi^1 *Mainz*
T3T1T2	比基尼 bi^3 ji^1 ni^2 *Bikini*	马拉维 ma^3 la^1 wei^2 *Malawi*
T3T1T3	果哥里 guo^3 ge^1 li^3 *Gogol*	里斯本 li^3 si^1 ben^3 *Lisbon*
T3T1T4	法拉利 fa^3 la^1 li^4 *Ferrari*	索菲亚 suo^3 fei^1 ya^4 *Sofia*
T3T2T1	海明威 hai^3 ming^2 wei^1 *Hemingway*	马尼拉 ma^3 ni^2 la^1 *Manila*
T3T2T2	卡琳达 ka^3 lin^2 da^2 *Kalinda*	所罗门 suo^3 luo^2 men^2 *Solomon*
T3T2T3	马德里 ma^3 de^2 li^3 *Madrid*	纽伦堡 niu^3 lun^2 bao^3 *Nuremberg*

T3T2T4	马其顿 ma^3 qi^2 dun^4 *Macedonia*	塔吉克 ta^3 ji^2 ke^4 *Tajik*
T3T3T1	卡米拉 ka^3 mi^3 la^1 *Camilla*	马耳他 ma^3 er^3 ta^1 *Malta*
T3T3T2	马里兰 ma^3 li^3 lan^2 *Maryland*	土耳其 tu^3 er^3 qi^2 *Turkey*
T3T3T3	索马里 suo^3 ma^3 li^3 *Somalia*	卡塔尔 ka^3 ta^3 er^3 *Qatar*
T3T3T4	柬埔寨 jian^3 pu^3 zhai^4 *Cambodia*	雅典娜 ya^3 dian^3 na^4 *Athena*
T3T4T1	卡耐基 ka^3 nai^4 ji^1 *Carnegie*	瓦杜兹 wa^3 du^4 zi^1 *Vaduz*
T3T4T2	普契尼 pu^3 qi^4 ni^2 *Puccini*	比利时 bi^3 li^4 shi^2 *Belgium*
T3T4T3	马六甲 ma^3 liu^4 jia^3 *Malacca*	马绍尔 ma^3 shao^4 er^3 *Marshall*
T3T4T4	玛丽娜 ma^3 li^4 na^4 *Marina*	以色列 yi^3 se^4 lie^4 *Isreal*
T4T1T1	孟加拉 meng^4 jia^1 la^1 *Bengal*	墨西哥 mo^4 xi^1 ge^1 *Mexico*
T4T1T2	布拉格 bu^4 la^1 ge^2 *Prague*	圣基茨 sheng^4 ji^1 ci^2 *Saint Kitts*
T4T1T3	爱丁堡 ai^4 ding^1 bao^3 *Edinburgh*	奥巴马 ao^4 ba^1 ma^3 *Obama*
T4T1T4	麦当娜 mai^1 dang^1 na^4 *Madonna*	奥斯陆 ao^4 si^1 lu^4 *Oslo*
T4T2T1	洛仑兹 luo^4 lun^2 zi^1 *Lorentz*	慕尼黑 mu^4 ni^2 hei^1 *Munich*
T4T2T2	不来梅 bu^4 lai^2 mei^2 *Bremen*	布隆迪 bu^4 long^2 di^2 *Burundi*

T4T2T3	洛伦佐 luo^4 lun^2 zuo^3 *Lorenzo*	立陶宛 li^4 tao^2 wan^3 *Lithuania*
T4T2T4	奥兰治 ao^4 lan^2 zhi^4 *Orange*	克林顿 ke^4 lin^2 dun^4 *Clinton*
T4T3T1	帕美拉 pa^4 mei^3 la^1 *Pamela*	萨瓦多 sa^4 wa^3 duo^1 *Salvador*
T4T3T2	布吕尼 bu^4 lü^3 ni^2 *Bruni*	爱尔兰 ai^4 er^3 lan^2 *Ireland*
T4T3T3	亚美里 ya^4 mei^3 li^3 *Amelie*	曼海姆 man^4 hai^3 mu^3 *Mannheim*
T4T3T4	布鲁诺 bu^4 lu^3 nuo^4 *Bruno*	利比亚 li^4 bi^3 ya^4 *Libya*
T4T4T1	洛佩兹 luo^4 pei^4 zi^1 *Lopez*	鹿特丹 lu^4 te^4 dan^1 *Rotterdam*
T4T4T2	拜占庭 bai^4 zhan^4 ting^2 *Byzantium*	密克罗 mi^4 ke^4 luo^2 *Mikro*
T4T4T3	瑞贝卡 rui^4 bei^4 ka^3 *Rebecca*	洛丽塔 luo^4 li^4 ta^3 *Lolita*
T4T4T4	奥地利 ao^4 di^4 li^4 *Austria*	意大利 yi^4 da^4 li^4 *Italy*

II-3. Trisyllabic list 2: σσ+σ.

Tonal Sequences	Word 1	Word 2
T1T1T1	拖拉机 tuo^1 la^1 ji^1 *tractor*	出租车 chu^1 zu^1 che^1 *taxi*
T1T1T2	公安局 gong^1 an^1 ju^2 *police station*	工商局 gong^1 shang^1 ju^2 *Trade and Industry Bureau*
T1T1T3	拉丁舞 la^1 ding^1 wu^3 *Latin dance*	飞机场 fei^1 ji^1 chang^3 *airport*

T1T1T4	基督教 ji^{1} du^{1} jiao^{4} *Christianity*	夫妻店 fu^{1} qi^{1} dian^{4} *family-run shop*
T1T2T1	金鱼缸 jin^{1} yu^{2} gang^{1} *goldfish tank*	天狼星 tian^{1} lang^{2} xing^{1} *Sirius*
T1T2T2	乌龙球 wu^{1} long^{2} qiu^{2} *own goal*	安全局 an^{1} quan^{2} ju^{2} *Security Bureau*
T1T2T3	葱油饼 cong^{1} you^{2} bing^{3} *onion pancake*	冰球手 bing^{1} qiu^{2} shou^{3} *hockey player*
T1T2T4	资格证 zi^{1} ge^{2} zheng^{4} *certificate*	侏罗纪 zhu^{1} luo^{2} ji^{4} *Jurassic*
T1T3T1	资本家 zi^{1} ben^{3} jia^{1} *capitalist*	风景区 feng^{1} jing^{3} qu^{1} *scenic district*
T1T3T2	宗主国 zong^{1} zhu^{3} guo^{2} *suzerain*	资产权 zi^{1} chan^{3} quan^{2} *property rights*
T1T3T3	斑点狗 ban^{1} dian^{3} gou^{3} *Dalmatian dog*	深水港 shen^{1} shui^{3} gang^{3} *deepwater harbor*
T1T3T4	天主教 tian^{1} zhu^{3} jiao^{4} *Catholicism*	斑马线 ban^{1} ma^{3} xian^{4} *zebra line*
T1T4T1	音乐家 yin^{1} yue^{4} jia^{1} *musician*	生物钟 sheng^{1} wu^{4} zhong^{1} *biochronometer*
T1T4T2	机器人 ji^{1} qi^{4} ren^{2} *robot*	工作狂 gong^{1} zuo^{4} kuang^{2} *workaholic*
T1T4T3	登记表 deng^{1} ji^{4} biao^{3} *registration form*	工艺品 gong^{1} yi^{4} pin^{3} *handicraft*
T1T4T4	关帝庙 guan^{1} di^{4} miao^{4} *Temple of General Kuan*	生物链 sheng^{1} wu^{4} lian^{4} *food chain*
T2T1T1	楼梯间 lou^{2} ti^{1} jian^{1} *stairwell*	红灯区 hong^{2} deng^{1} qu^{1} *red light district*
T2T1T2	鱼肝油 yu^{2} gan^{1} you^{2} *fish oil*	传播学 chuan^{2} bo^{1} xue^{2} *communication*

T2T1T3	服装厂 fu2 zhuang1 chang3 *garment factory*	流星雨 liu2 xing1 yu3 *meteor shower*
T2T1T4	牙周病 ya2 zhou1 bing4 *periodontal disease*	神经病 shen2 jing1 bing4 *psycho*
T2T2T1	南极洲 nan2 ji2 zhou1 *Antarctica*	流行歌 liu2 xing2 ge1 *pop song*
T2T2T2	蝴蝶结 hu2 die2 jie2 *bow knot*	直肠癌 zhi2 chang2 ai2 *rectum cancer*
T2T2T3	篮球场 lan2 qiu2 chang3 *basketball court*	流行语 liu2 xing2 yu3 *popular word*
T2T2T4	能源部 neng2 yuan2 bu4 *Ministry of Energy*	神农架 shen2 nong2 jia4 *(a place in Hubei Province)*
T2T3T1	平板车 ping2 ban3 che1 *trolley*	平顶山 ping2 ding3 shan1 *(a city in Henan Province)*
T2T3T2	龙井茶 long2 jing3 cha2 *Longjing tea*	牛仔裙 niu2 zai3 qun2 *jeans skirt*
T2T3T3	长短腿 chang2 duan3 tui3 *leg length discrepancy*	华表奖 hua2 biao3 jiang3 *Marble Pillar Award*
T2T3T4	营养液 ying2 yang3 ye4 *nutrient*	寒武纪 han2 wu3 ji4 *Cambrian*
T2T4T1	杂技班 za2 ji4 ban1 *acrobatic team*	杂货摊 za2 huo4 tan1 *grocery stall*
T2T4T2	国庆节 guo2 qing4 jie2 *National Day*	杂技团 za2 ji4 tuan2 *acrobatic society*
T2T4T3	游乐场 you2 le4 chang3 *playground*	玩具厂 wan2 ju4 chang3 *toy factory*
T2T4T4	疾病册 ji2 bing4 ce4 *list of diseases*	糖尿病 tang2 niao4 bing4 *diabetes*
T3T1T1	闪光灯 shan3 guang1 deng1 *flash light*	镁光灯 mei3 guang1 deng1 *limelight*

T3T1T2	火车头 huo3 che1 tou2 *locomotive*	母亲河 mu3 qin1 he2 *mother river*
T3T1T3	奖金卡 jiang3 jin1 ka3 *bonus card*	广东省 guang3 dong1 sheng3 *Guangdong Province*
T3T1T4	火车票 huo3 che1 piao4 *train ticket*	火车站 huo3 che1 zhan4 *train station*
T3T2T1	主人翁 zhu3 ren2 weng1 *hero*	里程碑 li3 cheng2 bei1 *milestone*
T3T2T2	狗头铡 gou3 tou2 zha2 *kobold guillotine*	北极熊 bei3 ji2 xiong2 *polar bear*
T3T2T3	打折卡 da3 zhe2 ka3 *coupon*	死亡谷 si3 wang2 gu3 *death valley*
T3T2T4	友情价 you3 qing2 jia4 *friendly price*	恐龙蛋 kong3 long2 dan4 *dinosaur egg*
T3T3T1	古董车 gu3 dong3 che1 *antique car*	宝马车 bao3 ma3 che1 *BMW car*
T3T3T2	火把节 huo3 ba3 jie2 *torch festival*	理想国 li3 xiang3 guo2 *Utopia*
T3T3T3	演讲稿 yan3 jiang3 gao3 *speech text*	总统府 zong3 tong3 fu3 *president office*
T3T3T4	雅宝路 ya3 bao3 lu4 *Yabao Road*	老虎凳 lao3 hu3 deng4 *tiger bench*
T3T4T1	可乐机 ke3 le4 ji1 *coke machine*	百叶窗 bai3 ye4 chuang1 *venetian blind*
T3T4T2	保健球 bao3 jian4 qiu2 *bowling*	乳腺癌 ru3 xian4 ai2 *breast cancer*
T3T4T3	统计表 tong3 ji4 biao3 *statistic form*	笔记本 bi3 ji4 ben3 *notebook*
T3T4T4	感叹句 gan3 tan4 ju4 *exclamatory sentence*	打印店 da3 yin4 dian4 *print shop*

T4T1T1	外交官 wai^4 jiao^1 guan^1 *diplomat*	地方官 di^4 fang^1 guan^1 *local official*
T4T1T2	太空船 tai^4 kong^1 chuan^2 *spaceship*	电机房 dian^4 ji^1 fang^2 *engine room*
T4T1T3	贵宾卡 gui^4 bin^1 ka^3 *VIP card*	便签纸 bian^4 qian^1 zhi^3 *sticker*
T4T1T4	外交部 wai^4 jiao^1 bu^4 *Ministry of Foreign Affairs*	太空站 tai^4 kong^1 zhan^4 *space station*
T4T2T1	地球村 di^4 qiu^2 cun^1 *earth village*	共鸣腔 gong^4 ming^2 qiang^1 *resonance cavity*
T4T2T2	太平洋 tai^4 ping^2 yang^2 *Pacific Ocean*	橡皮泥 xing^4 pi^2 ni^2 *plasticine*
T4T2T3	互联网 hu^4 lian^2 wang^3 *Internet*	棒球手 bang^4 qiu^2 shou^3 *baseball player*
T4T2T4	外来妹 wai^4 lai^2 mei^4 *girls from other places*	空格键 kong^4 ge^2 jian^4 *space bar*
T4T3T1	电子钟 dian^4 zi^3 zhong^1 *electronic clock*	样板间 yang^4 ban^3 jian^1 *sample room*
T4T3T2	盗版盘 dao^4 ban^3 pan^2 *pirated CD*	电影节 dian^4 ying^3 jie^2 *film festival*
T4T3T3	电子表 dian^4 zi^3 biao^3 *electronic watch*	电影展 dian^4 ying^3 zhan^3 *movie exhibition*
T4T3T4	电影票 dian^4 ying^3 piao^4 *movie ticket*	电影院 dian^4 ying^3 yuan^4 *cinema*
T4T4T1	纪念碑 ji^4 nian^4 bei^1 *monument*	计算机 ji^4 suan^4 ji^1 *computer*
T4T4T2	纪念堂 ji^4 nian^4 tang^2 *memorial hall*	电信局 dian^4 xin^4 ju^2 *tele-comm office*
T4T4T3	信用卡 xin^4 yong^4 ka^3 *credit card*	记忆体 ji^4 yi^4 ti^3 *RAM*

T4T4T4	数据库 shu⁴ ju⁴ ku⁴ *database*	计算器 ji⁴ suan⁴ qi⁴ *calculator*

II-4. Trisyllabic list 3: σ+σσ.

Tonal Sequences	Word 1	Word 2
T1T1T1	高工资 gao¹ gong¹ zi¹ *high salary*	新飞机 xin¹ fei¹ ji¹ *new aircraft*
T1T1T2	新中国 xin¹ zhong¹ guo² *new China*	多功能 duo¹ gong¹ neng² *multi-function*
T1T1T3	宽肩膀 kuan¹ jian¹ bang³ *wide shoulder*	光污染 guang¹ wu¹ ran³ *light pollution*
T1T1T4	轻音乐 qing¹ yin¹ yue⁴ *light music*	新花样 xin¹ hua¹ yang⁴ *new trick*
T1T2T1	干蹄筋 gan¹ ti² jin¹ *dry tendon*	双职工 shuang¹ zhi² gong¹ *working parents*
T1T2T2	东柏林 dong¹ bo² lin² *East Berlin*	西柏林 xi¹ bo² lin² *west Berlin*
T1T2T3	新油桶 xin¹ you² tong³ *new oil tank*	鲜牛奶 xian¹ niu² nai³ *fresh milk*
T1T2T4	新邮件 xin¹ you² jian⁴ *new mail*	新词汇 xin¹ ci² hui⁴ *new vocabulary*
T1T3T1	湿纸巾 shi¹ zhi³ jin¹ *wet tissue*	清辅音 qing¹ fu³ yin¹ *voiceless consonant*
T1T3T2	猪口条 zhu¹ kou³ tiao² *pig tongue*	金耳环 jin¹ er³ huan² *gold earring*
T1T3T3	新总统 xin¹ zong³ tong³ *new president*	新产品 xin¹ chan³ pin³ *new product*
T1T3T4	亲姐妹 qin¹ jie³ mei⁴ *blood sister*	新产业 xin¹ chan³ ye⁴ *new industry*

T1T4T1	湿面巾 shi^{1} mian^{4} jin^{1} *wet facial tissue*	新校区 xin^{1} xiao^{4} qu^{1} *new campus*
T1T4T2	猪大肠 zhu^{1} da^{4} chang^{2} *pig intestine*	新纪元 xin^{1} ji^{4} yuan^{2} *new era*
T1T4T3	新电影 xin^{1} dian^{4} ying^{3} *new movie*	花大姐 hua^{1} da^{4} jie^{3} *ladybug*
T1T4T4	新概念 xin^{1} gai^{4} nian^{4} *new concept*	中世纪 zhong^{1} shi^{4} ji^{4} *middle age*
T2T1T1	团中央 tuan^{2} zhong^{1} yang^{1} *league center*	毛蜘蛛 mao^{2} zhi^{1} zhu^{1} *hairy spider*
T2T1T2	红蜻蜓 hong^{2} qing^{1} ting^{2} *red dragonfly*	瓷花瓶 ci^{2} hua^{1} ping^{2} *ceramic vase*
T2T1T3	红糙米 hong^{2} cao^{1} mi^{3} *red rice*	零增长 ling^{2} zeng^{1} zhang^{3} *zero increase*
T2T1T4	咸鸭蛋 xian^{2} ya^{1} dan^{4} *salty duck egg*	皮车座 pi^{2} che^{1} zuo^{4} *feather seat*
T2T2T1	蓝瓷砖 lan^{2} ci^{2} zhuan^{1} *blue ceramic tile*	男同胞 nan^{2} tong^{2} bao^{1} *fellow men*
T2T2T2	前王朝 qian^{2} wang^{2} chao^{2} *previous dynasty*	明长城 ming^{2} chang^{2} cheng^{2} *Ming Great Wall*
T2T2T3	毒苹果 du^{2} ping^{2} guo^{3} *toxic apple*	男旗手 nan^{2} qi^{2} shou^{3} *male flag-bearer*
T2T2T4	咸白菜 xian^{2} bai^{2} cai^{4} *salty cabbage*	雌杨树 ci^{2} yang^{2} shu^{4} *female poplar*
T2T3T1	纯酒精 chun^{2} jiu^{3} jing^{1} *pure alcohol*	浊辅音 zhuo^{2} fu^{3} yin^{1} *voiced consonant*
T2T3T2	红嘴唇 hong^{2} zui^{3} chun^{2} *red lip*	蓝宝石 lan^{2} bao^{3} shi^{2} *sapphire*
T2T3T3	前总统 qian^{2} zong^{3} tong^{3} *former president*	男保姆 nan^{2} bao^{3} mu^{3} *male nanny*

T2T3T4	蓝孔雀 lan2 kong3 que4 *blue peacock*	雌孔雀 ci2 kong3 que4 *female peacock*
T2T4T1	男卫兵 nan2 wei4 bing1 *male guard*	白衬衫 bai2 chen4 shan1 *white shirt*
T2T4T2	红气球 hong2 qi4 qiu2 *red balloon*	南半球 nan2 ban4 qiu2 *south hemisphere*
T2T4T3	源代码 yuan2 dai4 ma3 *source code*	红墨水 hong2 mo4 shui3 *red ink*
T2T4T4	全自动 quan2 zi4 dong4 *automatic*	全世界 quan2 shi4 jie4 *entire world*
T3T1T1	假观音 jia3 guan1 yin1 *fake Buddha*	党中央 dang3 zhong1 yang1 *party center*
T3T1T2	小家庭 xiao3 jia1 ting2 *small family*	紫蜻蜓 zi3 qing1 ting2 *purple dragonfly*
T3T1T3	伪君子 wei3 jun1 zi3 *hypocrite*	水污染 shui3 wu1 ran3 *water pollution*
T3T1T4	冷空气 leng3 kong1 qi4 *cold air*	总司令 zong3 si1 ling4 *general commander*
T3T2T1	假投资 jia3 tou2 zi1 *fake investment*	假结婚 jia3 jie2 hun1 *fake marriage*
T3T2T2	小提琴 xiao3 ti2 qin2 *violin*	老同学 lao3 tong2 xue2 *old classmate*
T3T2T3	镁离子 mei3 li2 zi3 *magnesium ion*	钾离子 jia3 li2 zi3 *potassium ion*
T3T2T4	软白菜 ruan3 bai2 cai4 *soft cabbage*	软实力 ruan3 shi2 li4 *soft power*
T3T3T1	女主播 nü3 zhu3 bo1 *anchorwoman*	死火山 si3 huo3 shan1 *dead volcano*
T3T3T2	老党员 lao3 dang3 yuan2 *old party member*	小老头 xiao3 lao3 tou2 *little old man*

T3T3T3	女总统 nü^3 zong^3 tong^3 *female president*	女总理 nü^3 zong^3 li^3 *female prime minister*
T3T3T4	紫孔雀 zi^3 kong^3 que^4 *purple peacock*	小眼镜 xiao^3 yang^3 jing^4 *small glasses*
T3T4T1	小汽车 xiao^3 qi^4 che^1 *small car*	水立方 shui^3 li^4 fang^1 *Water Cube*
T3T4T2	伪政权 wei^3 zheng^4 quan^2 *puppet regime*	女队员 nü^3 dui^4 yuan^2 *female member*
T3T4T3	老电影 lao^3 dian^4 ying^3 *old movie*	伪政府 wei^3 zheng^4 fu^3 *puppet government*
T3T4T4	假面具 jia^3 mian^4 ju^4 *guise*	小报告 xiao3 bao4 gao4 *secret report*
T4T1T1	硬钢盔 ying^4 gang^1 kui^1 *hard helmet*	市中心 shi^4 zhong^1 xin^1 *city center*
T4T1T2	旧中国 jiu^4 zhong^1 guo^2 *old China*	大家庭 da^4 jia^1 ting^2 *big family*
T4T1T3	负增长 fu^4 zeng^1 zhang^3 *negative increase*	重金属 zhong^4 jin^1 shu^3 *heavy metal*
T4T1T4	旧家具 jiu^4 jia^1 ju^4 *old furniture*	半封建 ban^4 feng^1 jian^4 *semi-feudal*
T4T2T1	半元音 ban^4 yuan^2 yin^1 *semi-vowel*	大熊猫 da^4 xiong^2 mao^1 *giant panda*
T4T2T2	臭流氓 chou^4 liu^2 mang^2 *hooligan*	大结局 da^4 jie^2 ju^2 *finale*
T4T2T3	热牛奶 re^4 niu^2 nai^3 *hot milk*	负离子 fu^4 li^2 zi^3 *anion*
T4T2T4	辣白菜 la^4 bai^2 cai^4 *spicy cabbage*	大屏幕 da^4 ping^2 mu^4 *large screen*
T4T3T1	厚脚跟 hou^4 jiao^3 gen^1 *thick heel*	副主编 fu^4 zhu^3 bian^1 *vice editor-in-chief*

T4T3T2	厚脸皮 lou^{4} lian^{3} pi^{2} *cheeky*	代总裁 dai^{4} zong^{3} cai^{2} *acting CEO*
T4T3T3	臭老九 chou^{4} lao^{3} jiu^{3} *intellectual*	大拇指 da^{4} mu^{3} zhi^{3} *thumb*
T4T3T4	旧广告 jiu^{4} guang^{3} gao^{4} *old commercial*	电暖器 dian^{4} nuan^{3} qi^{4} *heater*
T4T4T1	亚健康 ya^{4} jian^{4} kang^{1} *sub-health*	电饭锅 dian^{4} fan^{4} guo^{1} *rice cooker*
T4T4T2	热气球 re^{4} qi^{4} qiu^{2} *hot air balloon*	热炕头 re^{4} kang^{4} tou^{2} *heated bed*
T4T4T3	废报纸 fei^{4} bao^{4} zhi^{3} *waste newspaper*	木地板 mu^{4} di^{4} ban^{3} *wooden floor*
T4T4T4	慢动作 man^{4} dong^{4} zuo^{4} *slow motion*	半自动 ban^{4} zi^{4} dong^{4} *semi-automatic*

Appendix III

The full list of stimuli in Chapter 4 is given below.

III-1. Near-Merger Sandhi targets and respective segmental competitors.

		Stimuli	Log Freq.	N. of Strokes	Bigram Mutual Info.
1	Target	指法 (zhi3 fa3) *fingering*	0.9	17	-0.72
	Segmental Competitor	滞后 (zhi4 hou4) *lag*	0.95	18	-
2	Target	反响 (fan3 xiang3) *repercussion*	1.83	13	3.72
	Segmental Competitor	范例 (fan4 li4) *example*	1.23	16	-
3	Target	赌本 (du3 ben3) *money for gambling*	0.48	17	0.61
	Segmental Competitor	杜撰 (du4 zhuan4) *to fabricate*	1.43	22	-
4	Target	赌友 (du3 you3) *gambler friend*	0.48	16	4.37
	Segmental Competitor	度假 (du4 jia4) *to take a vacation*	2.85	20	-
5	Target	俯首 (fu3 shou3) *to be obedient*	1	19	8.12
	Segmental Competitor	附录 (fu4 lu4) *appendix*	0.9	15	-
6	Target	抚养 (fu3 yang3) *to bring up*	2.59	16	7.85
	Segmental Competitor	负债 (fu4 zhai4) *debt*	1.58	16	-
7	Target	雨水 (yu3 shui3) *rainwater*	1.86	12	5.07
	Segmental Competitor	预购 (yu4 gou4) *purchase in advance*	0.3	18	-
8	Target	主管 (zhu3 guan3) *person in charge*	2.69	19	4.09
	Segmental Competitor	注释 (zhu4 shi4) *annotation*	1.49	20	-

9	Target	虎口 (hu³ kou³) *jaws of death*	1.46	11	4.03
	Segmental Competitor	互助 (hu⁴ zhu⁴) *to help each other*	1.8	11	-
10	Target	火把 (huo³ ba³) *torch*	1.91	11	3.14
	Segmental Competitor	货架 (huo⁴ jia⁴) *goods shelf*	1.63	17	-
11	Target	咫尺 (zhi³ chi³) *very short distance*	1.15	13	14.10
	Segmental Competitor	制定 (zhi⁴ ding⁴) *to formulate*	2.51	16	-
12	Target	眼睑 (yan³ jian³) *eyelids*	1.45	23	8.18
	Segmental Competitor	厌恶 (yan⁴ wu⁴) *to dislike*	2.61	16	-
13	Target	使馆 (shi³ guan³) *embassy*	1.57	19	4.32
	Segmental Competitor	事件 (shi⁴ jian⁴) *event*	3.44	14	-
14	Target	土改 (tu³ gai³) *land reform*	0	10	5.63
	Segmental Competitor	兔肉 (tu⁴ rou⁴) *rabbit meat*	1.04	14	-
15	Target	统管 (tong³ guan³) *overall administration*	0.3	23	0.06
	Segmental Competitor	痛恨 (tong⁴ hen⁴) *to hate*	2.32	21	-
16	Target	仰角 (yang³ jiao³) *angle of elevation*	0.3	13	1.84
	Segmental Competitor	样式 (yang⁴ shi⁴) *pattern*	1.81	16	-
17	Target	指导 (zhi³ dao³) *guide*	2.68	15	7.43
	Segmental Competitor	制造 (zhi⁴ zao⁴) *to manufacture*	3.2	18	-
18	Target	起码 (qi³ ma³) *at least*	2.99	18	6.61
	Segmental Competitor	气愤 (qi⁴ fen⁴) *furious*	2.06	16	-

III-2. No-Merger Sandhi targets and respective segmental competitors.

		Stimuli	Log Freq.	N. of Strokes	Bigram Mutual Info.
1	Target	僵尸 (jiang^1 shi^1) *mummy*	2.63	18	9.44
	Segmental Competitor	降落 (jiang^4 luo^4) *to land*	2.84	20	-
2	Target	尖刀 (jian^1 dao^1) *sharp knife*	0.9	8	7.25
	Segmental Competitor	渐变 (jian^4 bian^4) *to change graudually*	0.3	19	-
3	Target	低压 (di^1 ya^1) *low pressure*	1.23	13	2.24
	Segmental Competitor	地势 (di^4 shi^4) *topography*	1.3	14	-
4	Target	肩章 (jian^1 zhang^1) *shoulder badge*	1.15	19	4.72
	Segmental Competitor	健将 (jian^4 jiang^4) *athelete*	1.32	19	-
5	Target	交心 (jiao^1 xin^1) *to lay one's heart bare*	1.4	10	-0.99
	Segmental Competitor	教具 (jiao^4 ju^4) *teaching aids*	0.78	19	-
6	Target	收心 (shou^1 xin^1) *to get serious*	0	10	-0.28
	Segmental Competitor	授课 (shou^4 ke^4) *to teach*	1.32	21	-
7	Target	优生 (you^1 sheng^1) *healthy birthing*	0.3	11	0.88
	Segmental Competitor	诱变 (you^4 bian^4) *mutagenesis*	0	17	-
8	Target	收编 (shou^1 bian^1) *to incorporate armies*	0.3	18	3.78
	Segmental Competitor	受贿 (shou^4 hui^4) *to take bribes*	1.7	18	-
9	Target	搭车 (da^1 che^1) *to get a lift*	1.11	16	3.46
	Segmental Competitor	大麦 (da^4 mai^4) *barley*	1.32	10	-

10	Target	世间 (shi4 jian1) *world*	2.12	12	4.2
	Segmental Competitor	史册 (shi3 ce4) *historical records*	1.78	10	-
11	Target	汽车 (qi4 che1) *car*	3.12	11	9.38
	Segmental Competitor	起义 (qi3 yi4) *revolt*	1.91	13	-
12	Target	器官 (qi4 guan1) *organ*	2.48	24	6.27
	Segmental Competitor	乞丐 (qi3 gai4) *beggar*	2.12	7	-
13	Target	浴缸 (yu4 gang1) *bathtub*	2.44	19	11.22
	Segmental Competitor	语录 (yu3 lu4) *quotation*	1.15	17	-
14	Target	市区 (shi4 qu1) *urban area*	2.4	9	5.59
	Segmental Competitor	使坏 (shi3 huai4) *to play dirty*	1.43	15	-
15	Target	误工 (wu4 gong1) *absent from work*	0.3	12	3.4
	Segmental Competitor	舞剧 (wu3 ju4) *dance drama*	0.78	24	-
16	Target	异心 (yi4 xin1) *disloyalty*	0.48	10	-1.19
	Segmental Competitor	椅背 (yi3 bei4) *back of chair*	1.11	21	-
17	Target	滞销 (zhi4 xiao1) *unsaleable*	0.3	24	9.08
	Segmental Competitor	纸币 (zhi3 bi4) *paper money*	1.61	11	-
18	Target	面纱 (mian4 sha1) *veil*	1.77	16	3.46
	Segmental Competitor	免税 (mian3 shui4) *duty-free*	1.79	19	-

III-3. No Sandhi targets and respective segmental competitors.

Stimuli	Log Freq.	N. of Strokes	Bigram Mutual Info.

1	Target	继续 (ji4 xu4) *to continue*	3.67	21	11.76
	Segmental Competitor	挤压 (ji3 ya1) *to press*	1.82	15	-
2	Target	意愿 (yi4 yuan4) *wish*	2.45	27	2.3
	Segmental Competitor	乙肝 (yi3 gan1) *hepatitis B*	0.6	8	-
3	Target	事例 (shi4 li4) *example*	1.11	16	2.14
	Segmental Competitor	史诗 (shi3 shi1) *epic*	1.72	13	-
4	Target	教义 (jiao4 yi4) *religious regulations*	1.81	14	3.2
	Segmental Competitor	脚尖 (jiao3 jian1) *tiptoe*	1.93	17	-
5	Target	去世 (qu4 shi4) *to pass away*	3.02	10	2.55
	Segmental Competitor	取消 (qu3 xiao1) *to cancle*	3.28	18	-
6	Target	上市 (shang4 shi4) *to go on market*	1.97	8	0.48
	Segmental Competitor	赏光 (shang3 guang1) *to honour with presence*	1.45	18	-
7	Target	检查 (jian3 cha2) *to check*	3.37	20	11.24
	Segmental Competitor	见解 (jian4 jie3) *opinion*	2.01	17	-
8	Target	简直 (jian3 zhi2) *simply*	3.41	21	8.86
	Segmental Competitor	舰长 (jian4 zhang3) *captain*	2.2	14	-
9	Target	主题 (zhu3 ti2) *topic*	2.77	17	4.5
	Segmental Competitor	住址 (zhu4 zhi3) *home address*	2.33	14	-
10	Target	吻合 (wen3 he2) *to match*	2.55	13	5.6
	Segmental Competitor	问好 (wen4 hao3) *to greet*	2.64	12	-
11	Target	妥协 (tuo3 xie2) *to compromise*	2.39	13	11.25
	Segmental Competitor	拓展 (tuo4 zhan3) *to develop*	1.72	18	-

12	Target	指明 (zhi^3 ming^2) *to point out*	1.93	17	0.46
	Segmental Competitor	质朴 (zhi4 pu3) *simple*	1	14	-
13	Target	史册 (shi^3 ce^4) *historical records*	1.73	10	6.21
	Segmental Competitor	试纸 (shi4 zhi3) *regeant paper*	1.08	15	-
14	Target	舞会 (wu^3 hui^4) *Prom*	2.76	20	3.98
	Segmental Competitor	物产 (wu4 chan3) *resource*	0.7	14	-
15	Target	享誉 (xiang^3 yu^4) *to enjoy fame*	1.04	21	6.36
	Segmental Competitor	巷口 (xiang4 kou3) *end of alley*	0.6	12	-
16	Target	手锯 (shou^3 ju^4) *hand saw*	0.48	17	1.3
	Segmental Competitor	瘦小 (shou4 xiao3) *thin and small*	1.63	17	-
17	Target	暑假 (shu^3 jia^4) *summer vacation*	2.25	23	11.06
	Segmental Competitor	术语 (shu4 yu3) *termiology*	2.24	14	-
18	Target	简历 (jian^3 li^4) *resumé*	2.22	17	4.6
	Segmental Competitor	健美 (jian4 mei3) *fitness*	1.62	19	-

Appendix IV

The full list of stimuli in Chapter 5 is given below.

IV-1. Near-Merger Sandhi targets (T3T3) and respective competitors.

		Stimuli	Log Freq.	N. of Stroke	Bigram Mutual Info.
1	Target	指法 (zhi^3 fa^3) *fingering*	0.9	17	-0.72
	Toneme Competitor	纸箱 (zhi^3 xiang^1) *carton*	1.3	22	-
	Contour Competitor	职称 (zhi^2 cheng^1) *professional title*	0.48	21	-
	Segmental Competitor	滞后 (zhi^4 hou^4) *lag*	0.95	18	-
2	Target	反响 (fan^3 xiang^3) *repercussion*	1.83	13	3.72
	Toneme Competitor	反击 (fan^3 ji^1) *counter-attack*	2.54	9	-
	Contour Competitor	凡间 (fan^2 jian^1) *the mortal world*	1.2	10	-
	Segmental Competitor	范例 (fan^4 li^4) *example*	1.23	16	-
3	Target	赌本 (du^3 ben^3) *money for gambling*	0.48	17	0.61
	Toneme Competitor	堵车 (du^3 che^1) *traffic jam*	2.09	15	-
	Contour Competitor	读音 (du^2 yin^1) *pronunciation*	1.28	19	-
	Segmental Competitor	杜撰 (du^4 zhuan^4) *to fabricate*	1.43	22	-
4	Target	赌友 (du^3 you^3) *gambler friend*	0.48	16	4.37
	Toneme Competitor	堵车 (du^3 che^1) *traffic jam*	2.09	15	-
	Contour Competitor	独资 (du^2 zi^1) *exclusive investment*	0	19	-
	Segmental Competitor	度假 (du^4 jia^4) *to take a vacation*	2.85	20	-

5	Target	俯首 (fu^3 shou^3) *to be obedient*	1	19	8.12
	Toneme Competitor	辅修 (fu^3 xiu^1) *to minor in*	0.3	20	-
	Contour Competitor	伏贴 (fu^2 tie^1) *to fit perfectly*	0	15	-
	Segmental Competitor	附录 (fu^4 lu^4) *appendix*	0.9	15	-
6	Target	抚养 (fu^3 yang^3) *to bring up*	2.59	16	7.85
	Toneme Competitor	俯身 (fu^3 shen^1) *to bend down*	1.49	17	-
	Contour Competitor	服装 (fu^2 zhuang^1) *clothes*	2.96	20	-
	Segmental Competitor	负债 (fu^4 zhai^4) *to debt*	1.58	16	-
7	Target	雨水 (yu^3 shui^3) *rainwater*	1.86	12	5.07
	Toneme Competitor	雨滴 (yu^3 di^1) *raindrop*	1.34	22	-
	Contour Competitor	鱼虾 (yu^2 xia^1) *fish and shrimps*	0.3	17	-
	Segmental Competitor	预购 (yu^4 gou^4) *to purchase in advance*	0.3	18	-
8	Target	主管 (zhu^3 guan^3) *person in charge*	2.69	19	4.09
	Toneme Competitor	主攻 (zhu^3 gong^1) *main attack*	1.3	12	-
	Contour Competitor	竹签 (zhu^2 qian^1) *bamboo stick*	0	19	-
	Segmental Competitor	注释 (zhu^4 shi^4) *annotation*	1.49	20	-
9	Target	虎口 (hu^3 kou^3) *jaws of death*	1.46	11	4.03
	Toneme Competitor	虎威 (hu^3 wei^1) *power and prestige*	0	17	-
	Contour Competitor	湖边 (hu^2 bian^1) *lakeside*	2.39	17	-
	Segmental Competitor	互助 (hu^4 zhu^4) *to help each other*	1.8	11	-

10	Target	火把 (huo³ ba³) *torch*	1.91	11	3.14
	Toneme Competitor	火锅 (huo³ guo¹) *hot pot*	1.34	16	-
	Contour Competitor	活捉 (huo² zhuo¹) *capture alive*	1.92	19	-
	Segmental Competitor	货架 (huo⁴ jia⁴) *goods shelf*	1.63	17	-
11	Target	咫尺 (zhi³ chi³) *very short*	1.15	13	14.10
	Toneme Competitor	指尖 (zhi³ jian¹) *finger tip*	1.75	15	-
	Contour Competitor	职称 (zhi² cheng¹) *professional title*	0.48	21	-
	Segmental Competitor	制定 (zhi⁴ ding⁴) *to formulate*	2.51	16	-
12	Target	眼睑 (yan³ jian³) *eyelids*	1.45	23	8.18
	Toneme Competitor	眼珠 (yan³ zhu¹) *eyeball*	1.83	21	-
	Contour Competitor	严师 (yan² shi¹) *strict teacher*	0	13	-
	Segmental Competitor	厌恶 (yan⁴ wu⁴) *to dislike*	2.61	16	-
13	Target	使馆 (shi³ guan³) *embassy*	1.57	19	4.32
	Toneme Competitor	始发 (shi³ fa¹) *departure*	0.6	13	-
	Contour Competitor	时钟 (shi² zhong¹) *clock*	2	16	-
	Segmental Competitor	事件 (shi⁴ jian⁴) *event*	3.44	14	-
14	Target	土改 (tu³ gai³) *land reform*	0	10	5.63
	Toneme Competitor	土丘 (tu³ qiu¹) *mound*	0.7	8	-
	Contour Competitor	图钉 (tu² ding¹) *drawing pin*	1.3	15	-
	Segmental Competitor	兔肉 (tu⁴ rou⁴) *rabbit meat*	1.04	14	-

15	Target	统管 (tong^3 guan^3) *overall administration*	0.3	23	0.06
	Toneme Competitor	统称 (tong^3 cheng^1) *general term*	0.7	19	-
	Contour Competitor	同桌 (tong^2 zhuo^1) *deskmate*	0.78	16	-
	Segmental Competitor	痛恨 (tong^4 hen^4) *to hate*	2.32	21	-
16	Target	仰角 (yang^3 jiao^3) *angle of elevation*	0.3	13	1.84
	Toneme Competitor	养生 (yang^3 sheng^1) *to maintain good health*	0.7	14	-
	Contour Competitor	羊羔 (yang^2 gao^1) *lamb*	1.91	16	-
	Segmental Competitor	样式 (yang^4 shi^4) *pattern*	1.81	16	-
17	Target	指导 (zhi^3 dao^3) *to guide*	2.68	15	7.43
	Toneme Competitor	纸巾 (zhi^3 jin^1) *tissue*	2.33	10	-
	Contour Competitor	直说 (zhi^2 shuo^1) *to speak frankly*	2.73	17	-
	Segmental Competitor	制造 (zhi^4 zao^4) *to manufacture*	3.2	18	-
18	Target	起码 (qi^3 ma^3) *at least*	2.99	18	6.61
	Toneme Competitor	起初 (qi^3 chu^1) *in the beginning*	2.39	17	-
	Contour Competitor	骑车 (qi^2 che^1) *to cycle*	2.33	15	-
	Segmental Competitor	气愤 (qi^4 fen^4) *furious*	2.06	16	-

IV-2. No-Merger Sandhi targets (T1T1 and T4T1) and respective competitors.

Stimuli			Log Freq.	N. of Stroke	Bigram Mutual Info.

1	Target	僵尸 (jiang^1 shi^1) *mummy*	2.63	18	9.44
	Toneme Competitor	江南 (jiang^1 nan^2) *the southern Yangtze region*	1.38	15	-
	Contour Competitor	奖牌 (jiang^3 pai^2) *medal*	1.85	21	-
	Segmental Competitor	降落 (jiang^4 luo^4) *to land*	2.84	20	-
2	Target	尖刀 (jian^1 dao^1) *sharp knife*	0.9	8	7.25
	Toneme Competitor	奸贼 (jian^1 zei^2) *traitors*	0	16	-
	Contour Competitor	简明 (jian^3 ming^2) *brief*	0.6	21	-
	Segmental Competitor	渐变 (jian^4 bian^4) *to change graudually*	0.3	19	-
3	Target	低压 (di^1 ya^1) *low pressure*	1.23	13	2.24
	Toneme Competitor	低垂 (di^1 chui^2) *droop*	0.78	15	-
	Contour Competitor	抵达 (di^3 da^2) *to arrive*	2.54	14	-
	Segmental Competitor	地势 (di^4 shi^4) *topography*	1.3	14	-
4	Target	肩章 (jian^1 zhang^1) *shoulder badge*	1.15	19	4,72
	Toneme Competitor	兼容 (jian^1 rong^2) *compatible*	1.4	20	-
	Contour Competitor	简洁 (jian^3 jie^2) *concise*	1.8	22	-
	Segmental Competitor	健将 (jian^4 jiang^4) *athelete*	1.32	19	-
5	Target	交心 (jiao^1 xin^1) *to lay one's heart bare*	1.4	10	-0.99
	Toneme Competitor	胶皮 (jiao^1 pi^2) *rubber*	0.9	15	-
	Contour Competitor	角膜 (jiao^3 mo^2) *cornea*	1.48	21	-
	Segmental Competitor	教具 (jiao^4 ju^4) *teaching aids*	0.78	19	-

6	Target	收心 (shou¹ xin¹) *to get serious*	0	10	-0.28
	Toneme Competitor	收容 (shou¹ rong²) *to take in*	1.72	16	-
	Contour Competitor	守恒 (shou³ heng²) *conservation*	0.6	15	-
	Segmental Competitor	授课 (shou⁴ ke⁴) *to teach*	1.32	21	-
7	Target	优生 (you¹ sheng¹) *healthy birthing*	0.3	11	0.88
	Toneme Competitor	悠扬 (you¹ yang²) *melodious*	0.9	17	-
	Contour Competitor	友人 (you³ ren²) *friend*	1.36	6	-
	Segmental Competitor	诱变 (you⁴ bian⁴) *mutagenesis*	0	17	-
8	Target	收编 (shou¹ bian¹) *to incorporate armies*	0.3	18	3.78
	Toneme Competitor	收盘 (shou¹ pan²) *to close*	0.6	17	-
	Contour Competitor	手足 (shou³ zu²) *brothers*	1.48	11	-
	Segmental Competitor	受贿 (shou⁴ hui⁴) *to take bribes*	1.7	18	-
9	Target	搭车 (da¹ che¹) *to get a lift*	1.11	16	3.46
	Toneme Competitor	搭桥 (da¹ qiao²) *to build a bridge*	1.48	22	-
	Contour Competitor	打磨 (da³ mo²) *to polish*	1.53	21	-
	Segmental Competitor	大麦 (da⁴ mai⁴) *barley*	1.32	10	-
10	Target	世间 (shi⁴ jian¹) *world*	2.12	12	4.2
	Toneme Competitor	试航 (shi⁴ hang²) *trial voyage*	0	18	-
	Contour Competitor	实情 (shi² qing²) *truth*	2.38	19	-
	Segmental Competitor	史册 (shi³ ce⁴) *historical records*	1.78	10	-

11	Target	汽车 (qi⁴ che¹) *car*	3.12	11	9.38
	Toneme Competitor	汽油 (qi⁴ you²) *gasoline*	2.67	15	-
	Contour Competitor	旗袍 (qi² pao²) *cheongsam*	0.48	24	-
	Segmental Competitor	起义 (qi³ yi⁴) *revolt*	1.91	13	-
12	Target	器官 (qi⁴ guan¹) *organ*	2.48	24	6.27
	Toneme Competitor	气球 (qi⁴ qiu²) *balloon*	2.47	15	-
	Contour Competitor	棋牌 (qi² pai²) *chess and cards*	0.78	24	-
	Segmental Competitor	乞丐 (qi³ gai⁴) *beggar*	2.12	7	-
13	Target	浴缸 (yu⁴ gang¹) *bathtub*	2.44	19	11.22
	Toneme Competitor	预言 (yu⁴ yan²) *to predict*	2.81	17	-
	Contour Competitor	鱼雷 (yu² lei²) *torpedo*	1.99	21	-
	Segmental Competitor	语录 (yu³ lu⁴) *quotation*	1.15	17	-
14	Target	市区 (shi⁴ qu¹) *urban area*	2.4	9	5.59
	Toneme Competitor	市容 (shi⁴ rong²) *appearance of a city*	0.78	15	-
	Contour Competitor	识别 (shi² bie²) *to identify*	2.53	14	-
	Segmental Competitor	使坏 (shi³ huai⁴) *to play dirty*	0.78	15	-
15	Target	误工 (wu⁴ gong¹) *absent from work*	0.3	12	3.4
	Toneme Competitor	务农 (wu⁴ nong²) *to be a farmer*	0.7	11	-
	Contour Competitor	无穷 (wu² qiong²) *infinite*	2.08	11	-
	Segmental Competitor	舞剧 (wu³ ju⁴) *dance drama*	0.78	24	-

16	Target	异心 (yi⁴ xin¹) *disloyalty*	0.48	10	-1.19
	Toneme Competitor	异同 (yi⁴ tong²) *similarities and differences*	0.6	12	-
	Contour Competitor	疑难 (yi² nan²) *difficulty*	3.02	24	-
	Segmental Competitor	椅背 (yi³ bei⁴) *back of chair*	1.11	21	-
17	Target	滞销 (zhi⁴ xiao¹) *unsaleable*	0.3	24	9.08
	Toneme Competitor	致癌 (zhi⁴ ai²) *carcinogenic*	1.34	27	-
	Contour Competitor	植皮 (zhi² pi²) *skin grafting*	0.78	17	-
	Segmental Competitor	纸币 (zhi³ bi⁴) *paper money*	1.61	11	-
18	Target	面纱 (mian⁴ sha¹) *veil*	1.77	16	3.46
	Toneme Competitor	面条 (mian⁴ tiao²) *noodle*	2.47	16	-
	Contour Competitor	绵羊 (mian² yang²) *sheep*	1.96	17	-
	Segmental Competitor	免税 (mian³ shui⁴) *duty-free*	1.79	19	-

IV-3. No Sandhi targets (T4T4, T3T2 and T3T4) and respective competitors.

Stimuli			Log Freq.	N. of Stroke	Bigram Mutual Info.
1	Target	继续 (ji⁴ xu⁴) *to continue*	3.67	21	11.76
	Toneme Competitor	季节 (ji⁴ jie²) *season*	2.49	13	-
	Contour Competitor	机床 (ji¹ chuang²) *machine tool*	0.3	13	-
	Segmental Competitor	挤压 (ji³ ya¹) *to press*	1.82	15	-

2	Target	意愿 (yi4 yuan4) *wish*	2.45	27	2.3
	Toneme Competitor	意图 (yi4 tu2) *intention*	2.5	21	-
	Contour Competitor	衣橱 (yi1 chu2) *closet*	2.37	22	-
	Segmental Competitor	乙肝 (yi3 gan1) *hepatitis B*	0.6	8	-
3	Target	事例 (shi4 li4) *example*	1.11	16	2.14
	Toneme Competitor	试图 (shi4 tu2) *to attemp*	3.42	16	-
	Contour Competitor	师徒 (shi1 tu2) *master and apprentice*	0	16	-
	Segmental Competitor	史诗 (shi3 shi1) *epic*	1.72	13	-
4	Target	教义 (jiao4 yi4) *religious regulations*	1.81	14	3.2
	Toneme Competitor	教程 (jiao4 cheng2) *tutorial*	1.23	23	-
	Contour Competitor	交流 (jiao1 liu2) *to communicate*	2.99	16	-
	Segmental Competitor	脚尖 (jiao3 jian1) *tiptoe*	1.93	17	-
5	Target	去世 (qu4 shi4) *to pass away*	3.02	10	2.55
	Toneme Competitor	去除 (qu4 chu2) *to eliminate*	1.97	14	-
	Contour Competitor	屈服 (qu1 fu2) *to surrender*	2.4	16	-
	Segmental Competitor	取消 (qu3 xiao1) *to cancel*	3.28	18	-
6	Target	上市 (shang4 shi4) *to go on market*	1.97	8	0.48
	Toneme Competitor	上扬 (shang4 yang2) *to raise*	1.11	9	-
	Contour Competitor	伤员 (shang1 yuan2) *the wounded*	1.98	13	-
	Segmental Competitor	赏光 (shang3 guang1) *to honor with presence*	1.45	18	-

7	Target	检查 (jian³ cha²) *to check*	3.37	20	11.24
	Toneme Competitor	检验 (jian³ yan⁴) *to test*	2.59	21	-
	Contour Competitor	监狱 (jian¹ yu⁴) *jail*	3.67	19	-
	Segmental Competitor	见解 (jian⁴ jie³) *opinion*	2.01	17	-
8	Target	简直 (jian³ zhi²) *simply*	3.41	21	8.86
	Toneme Competitor	简历 (jian³ li⁴) *resumé*	2.42	17	-
	Contour Competitor	监控 (jian¹ kong⁴) *to monitor*	3.01	21	-
	Segmental Competitor	舰长 (jian⁴ zhang³) *captain*	2.2	14	-
9	Target	主题 (zhu³ ti²) *topic*	2.77	17	4.5
	Toneme Competitor	主妇 (zhu³ fu⁴) *housewife*	2.51	11	-
	Contour Competitor	猪肉 (zhu¹ rou⁴) *pork*	2.37	17	-
	Segmental Competitor	住址 (zhu⁴ zhi³) *home address*	2.33	14	-
10	Target	吻合 (wen³ he²) *to match*	2.55	13	5.6
	Toneme Competitor	稳重 (wen³ zhong⁴) *steady*	1.51	23	-
	Contour Competitor	温度 (wen¹ du⁴) *tempareture*	2.63	21	-
	Segmental Competitor	问好 (wen⁴ hao³) *to greet*	2.64	12	-
11	Target	妥协 (tuo³ xie²) *to compromise*	2.39	13	11.25
	Toneme Competitor	妥善 (tuo³ shan⁴) *appropriate*	1.72	19	-
	Contour Competitor	托运 (tuo¹ yun⁴) *to consign for shipment*	1.42	13	-
	Segmental Competitor	拓展 (tuo⁴ zhan³) *to develop*	1.72	18	-

12	Target	指明 (zhi^3 ming^2) *to point out*	1.93	17	0.46
	Toneme Competitor	纸币 (zhi^3 bi^4) *paper money*	1.61	11	-
	Contour Competitor	枝叶 (zhi^1 ye^4) *branches and leaves*	1.11	13	-
	Segmental Competitor	质朴 (zhi^4 pu^3) *simple*	1	14	-
13	Target	史册 (shi^3 ce^4) *historical records*	1.73	10	6.21
	Toneme Competitor	史前 (shi^3 qian^2) *prehistory*	1.84	14	-
	Contour Competitor	失陪 (shi^1 pei^2) *excuse me*	3.08	15	-
	Segmental Competitor	试纸 (shi^4 zhi^3) *dipstick*	1.08	15	-
14	Target	舞会 (wu^3 hui^4) *Prom*	2.76	20	3.98
	Toneme Competitor	舞台 (wu^3 tai^2) *stage*	3.28	19	-
	Contour Competitor	巫婆 (wu^1 po^2) *witch*	2.38	18	-
	Segmental Competitor	物产 (wu^4 chan^3) *resource*	0.7	14	-
15	Target	享誉 (xiang^3 yu^4) *to enjoy fame*	1.04	21	6.36
	Toneme Competitor	响铃 (xiang^3 ling^2) *alarm*	1.42	19	-
	Contour Competitor	香浓 (xiang^1 nong^2) *grumous*	0	18	-
	Segmental Competitor	巷口 (xiang^4 kou^3) *end of alley*	0.6	12	-
16	Target	手锯 (shou^3 ju^4) *hand saw*	0.48	17	1.3
	Toneme Competitor	守则 (shou^3 ze^2) *regulation*	2.11	12	-
	Contour Competitor	收集 (shou^1 ji^2) *to collect*	2.96	18	-
	Segmental Competitor	瘦小 (shou^4 xiao^3) *thin and small*	1.63	17	-

17	Target	暑假 (shu^3 jia^4) *summer vacation*	2.25	23	11.06
	Toneme Competitor	薯条 (shu^3 tiao^2) *French fries*	2.56	23	-
	Contour Competitor	抒情 (shu^1 qing^2) *to express one's emotions*	1.76	18	-
	Segmental Competitor	术语 (shu^4 yu^3) *terminology*	2.24	14	-
18	Target	简历 (jian^3 li^4) *resumé*	2.22	17	4.6
	Toneme Competitor	简洁 (jian^3 jie^2) *concise*	1.8	22	-
	Contour Competitor	兼容 (jian^1 rong^2) *compatible*	1.4	20	-
	Segmental Competitor	健美 (jian^4 mei^3) *fitness*	1.62	19	-

Appendix V

The full list of target sentences in different boundary conditions in Chapter 6. The critical Tx+N(NN)+T1+T1/T2 structure is marked in bold.

V-1. Below-NP Boundary condition.

Tx+N+T1T1:	
1	他说**他的猫偷**走了那条大鱼。 (Ta^{1} shuo^{1} **ta^{1} de mao^{1} tou^{1}** zou^{3} le na^{4} tiao^{2} da^{4} yu^{2}.) *He said his cat stole that big fish.*
2	他说**姨的猫听**到了那个声音。 (Ta^{1} shuo^{1} **yi^{2} de mao^{1} ting^{1}** dao^{4} le na^{4} ge sheng^{1} yin^{1}.) *He said auntie's cat heard that sound.*
3	他说**姐的猫关**上了那个小门。 (Ta^{1} shuo^{1} **jie^{3} de mao^{1} guan^{1}** shang^{4} le na^{4} ge xiao^{3} men^{2}.) *He said elder sister's cat shut that little door.*
4	他说**妹的猫吃**完了那袋粮食。 (Ta^{1} shuo^{1} **mei^{4} de mao^{1} chi^{1}** wan^{2} le na^{4} dai^{4} liang^{2} shi.) *He said younger sister's cat ate up that bag of grain.*
Tx+N+T1T2:	
5	他说**他的猫离**开了那座房子。 (Ta^{1} shuo^{1} **ta^{1} de mao^{1} li^{2}** kai^{1} le na^{4} zuo^{4} fang^{2} zi.) *He said his cat left that house.*
6	他说**姨的猫夺**回了那条大鱼。 (Ta^{1} shuo^{1} **yi^{2} de mao^{1} duo^{2}** hui^{2} le na^{4} tiao^{2} da^{4} yu^{2}.) *He said auntie's cat took back that big fish.*
7	他说**姐的猫迷**上了那个游戏。 (Ta^{1} shuo^{1} **jie^{3} de mao^{1} mi^{2}** shang^{4} le na^{4} ge you^{2} xi^{4}.) *He said elder sister's cat was crazy about that game.*
8	他说**妹的猫学**会了如何爬树。 (Ta^{1} shuo^{1} **mei^{4} de mao^{1} xue^{2}** hui^{4} le ru^{2} he^{2} pa^{2} shu^{4}.) *He said younger sister's cat learned how to climb a tree.*
Tx+NN+T1T1:	
9	他说**妈妈的猫抓**住了那只老鼠。 (Ta^{1} shuo^{1} **ma^{1} ma de mao^{1} zhua^{1}** zhu^{4} le na^{4} zhi^{1} lao^{3} shu^{3}.) *He said mom's cat caught that mouse.*
10	他说**爷爷的猫拉**住了那根绳子。 (Ta^{1} shuo^{1} **ye^{2} ye de mao^{1} la^{1}** zhu^{4} le na^{4} gen^{1} sheng^{2} zi.) *He said grandpa' cat tugged that rope.*
11	他说**姐姐的猫叼**走了那只小鸟。 (Ta^{1} shuo^{1} **jie^{3} jie de mao^{1} diao^{1}** zou^{3} le na^{4} zhi^{1} xiao^{3} niao^{3}.) *He said elder sister's cat held that bird in the mouth.*

12	他说妹妹的猫喝完了那盒牛奶。 (Ta^1 shuo^1 **mei^4 mei de mao^1 he^1** wan^2 le na^4 he^2 niu^2 nai^3.) *He said younger sister's cat drank up that box of milk.*
Tx+NN+T1T2:	
13	他说**妈妈的猫拿**走了那个玩具。 (Ta^1 shuo^1 **ma^1 ma de mao^1 na^2** zou^3 le na^4 ge wan^2 ju^4.) *He said mom's cat took that toy.*
14	他说**爷爷的猫爬**上了那棵老树。 (Ta^1 shuo^1 **ye^2 ye de mao^1 pa^2** shang^4 le na^4 ke^1 lao^3 shu^4.) *He said grandpa's cat climbed up that old tree.*
15	他说**姐姐的猫玩**坏了那个玩具。 (Ta^1 shuo^1 **jie^3 jie de mao^1 wan^2** huai^4 le na^4 ge wan^2 ju^4.) *He said elder sister's cat destroyed that toy.*
16	他说**妹妹的猫停**下了它的脚步。 (Ta^1 shuo^1 **mei^4 mei de mao^1 ting^2** xia^4 le ta^1 de jiao^3 bu^4.) *He said younger sister's cat stopped walking.*
Tx+NNN+T1T1:	
17	他说**妈妈们的猫踢**乱了那个线球。 (Ta^1 shuo^1 **ma^1 ma men de mao^1 ti^1** luan^4 le na^4 ge xian^4 qiu^2.) *He said moms' cat kicked that thread ball into a mess.*
18	他说**爷爷们的猫推**倒了那个鱼缸。 (Ta^1 shuo^1 **ye^2 ye men de mao^1 tui^1** dao^3 le na^4 ge yu^2 gang^1.) *He said grandpas' cat push down that fish tank.*
19	他说**姐姐们的猫扑**向了那只蝴蝶。 (Ta^1 shuo^1 **jie^3 jie men de mao^1 pu^1** xiang^4 le na^4 zhi^1 hu^2 die^2.) *He said elder sisters' cat leapt on that butter.*
20	他说**妹妹们的猫追**上了那只兔子。 (Ta^1 shuo^1 **mei^4 mei men de mao^1 zhui^1** shang^4 le na^4 zhi^1 tu^4 zi.) *He said younger sisters' cat caught up with that rabbit.*
Tx+NNN+T1T2:	
21	他说**妈妈们的猫咳**出了那根鱼刺。 (Ta^1 shuo^1 **ma^1 ma men de mao^1 ke^2** chu^1 le na^4 gen^1 yu^2 ci^4.) *He said moms' cat coughed up that fish bone.*
22	他说**爷爷们的猫赢**过了那条野狗。 (Ta^1 shuo^1 **ye^2 ye men de mao^1 ying^2** guo^4 le na4 tiao2 ye3 gou3.) *He said grandpas' cat defeated that wild dog.*
23	他说**姐姐们的猫衔**起了它的孩子。 (Ta^1 shuo^1 **jie^3 jie men de mao^1 xian^2** qi^3 le ta^1 de hai^2 zi.) *He said elder sisters' cat held its kid in the mouth.*
24	他说**妹妹们的猫回**到了它的小窝。 (Ta^1 shuo^1 **mei^4 mei men de mao^1 hui^2** dao^4 le ta^1 de xiao^3 wo^1.) *He said younger sisters's cat went back to its little nest.*

V-2. Subject-Predicate Boundary condition.

Tx+N+T1T1:	
1	他说**妈妈收听**了那个节目。 (Ta1 shuo1 **ma1 ma shou1 ting1** le na4 ge jie2 mu4.) He said mom listened to that program.
2	他说**爷爷搜刮**了很多钱财。 (Ta1 shuo1 **ye2 ye sou1 gua1** le hen3 duo1 qian2 cai2.) He said grandpa took a lot of money.
3	他说**姐姐招收**了几个学生。 (Ta1 shuo1 **jie3 jie zhao1 shou1** le ji3 ge xue2 sheng1.) He said elder sister recruited several students.
4	他说**妹妹发挥**了她的实力。 (Ta1 shuo1 **mei4 mei fa1 hui1** le ta1 de shi2 li4.) He said younger sister gave play to her power.
Tx+N+T1T2:	
5	他说**妈妈宣传**了这次活动。 (Ta1 shuo1 **ma1 ma xuan1 chuan2** le zhe4 ci4 huo2 dong4.) He said mom publicized this action.
6	他说**爷爷缺乏**了一些斗志。 (Ta1 shuo1 **ye2 ye que1 fa2** le yi4 xie1 dou4 zhi4.) He said grandpa lacked the will to fight.
7	他说**姐姐发觉**了这种现象。 (Ta1 shuo1 **jie3 jie fa1 jue2** le zhe4 zhong3 xian4 xiang4.) He said elder sister discovered this phenomenon.
8	他说**妹妹追求**了她的理想。 (Ta1 shuo1 **mei4 mei zhui1 qiu2** le ta1 de li3 xiang3.) He said elder sister pursued her dreams.
Tx+NN+T1T1:	
9	他说**妈妈们听说**了那件事情。 (Ta1 shuo1 **ma1 ma men ting1 shuo1** le na4 jian4 shi4 qing2.) He said moms heard of that situation.
10	他说**爷爷们张贴**了一个广告。 (Ta1 shuo1 **ye2 ye men zhang1 tie1** le yi2 ge4 guang3 gao4.) He said grandpas posted an advertisement.
11	他说**姐姐们接收**了那些病人。 (Ta1 shuo1 **jie3 jie men jie1 shou1** le na4 xie1 bing4 ren2.) He said elder sisters took over those patients.
12	他说**妹妹们牺牲**了很多利益。 (Ta1 shuo1 **mei4 mei men xi1 sheng1** le hen3 duo1 li4 yi4.) He said younger sisters sacrificed a lot of interest.
Tx+NN+T1T2:	
13	他说**妈妈们安排**了一次聚餐。 (Ta1 shuo1 **ma1 ma men an1 pai2** le yi2 ci4 ju4 can1.) He said moms arranged a dinner.

14	他说**爷爷们消除**了一些误会。 (Ta^{1} shuo^{1} **ye^{2} ye men xiao^{1} chu^{2}** le yi^{4} xie^{1} wu^{4} hui^{4}.) He said grandpas eliminated some misunderstandings.
15	他说**姐姐们经营**了一家餐厅。 (Ta^{1} shuo^{1} **jie^{3} jie men jing^{1} ying^{2}** le yi^{4} jia^{1} can^{1} ting^{1}.) He said elder sisters operated a restaurant.
16	他说**妹妹们坚持**了那个信念。 (Ta^{1} shuo^{1} **mei^{4} mei men jian^{1} chi^{2}** le na^{4} ge xin^{4} nian^{4}.) He said younger sisters stuck to that belief.
Tx+NNN+T1T1:	
17	他说**妈妈们的增加**了三百块钱。 (Ta^{1} shuo^{1} **ma^{1} ma men de zeng^{1} jia^{1}** le san^{1} bai^{3} kuai^{4} qian^{2}.) He said moms' increased by 300 yuan.
18	他说**爷爷们的开发**了新的产品。 (Ta^{1} shuo^{1} **ye^{2} ye men de kai^{1} fa^{1}** le xin^{1} de chan^{3} pin^{3}.) He said grandpas' developed new products.
19	他说**姐姐们的分发**了那堆课本。 (Ta^{1} shuo^{1} **jie^{3} jie men de fen^{1} fa^{1}** le na^{4} dui^{1} ke^{4} ben^{3}.) He said elder sisters' distributed that pile of book.
20	他说**妹妹们的装修**了这个房子。 (Ta^{1} shuo^{1} **mei^{4} mei men de zhuang^{1} xiu^{1}** le zhe^{4} ge^{4} fang^{2} zi.) He said younger sisters' decorated this house.
Tx+NNN+T1T2:	
21	他说**妈妈们的收集**了一堆古董。 (Ta^{1} shuo^{1} **ma^{1} ma men de shou^{1} ji^{2}** le yi^{4} dui^{1} gu^{3} dong^{3}.) He said moms' collected a lot of antiques.
22	他说**爷爷们的发扬**了这项传统。 (Ta^{1} shuo^{1} **ye^{2} ye men de fa^{1} yang^{2}** le zhe^{4} xiang^{4} chuan^{2} tong^{3}.) He said grandpas' promoted the tradition.
23	他说**姐姐们的遵循**了这种规律。 (Ta^{1} shuo^{1} **jie^{3} jie men de zun^{1} xun^{2}** le zhe^{4} zhong^{3} gui^{1} lü^{4}.) He said elder sisters' followed this law.
24	他说**妹妹们的收藏**了很多珠宝。 (Ta^{1} shuo^{1} **mei^{4} mei men de shou^{1} cang^{2}** le hen^{3} duo^{1} zhu^{1} bao^{3}.) He said younger sisters' collected meny jewelleries.

CURRICULUM VITAE

Qian Li was born in Nanjing in the People's Republic of China on 27 september 1985. In 2003, she attended Beijing Language and Culture University in Beijing. Her major in college was Teaching Chinese as a Second Language. In 2006, she joined the Phonetics Lab at Beijing Language and Culture University, where she was working as a phonetician. She obtained her Bachelor's degree in 2007 and Master's degree in 2010, both at Beijing Language and Culture Univeristy. In 2010, she moved to Leiden and joined Leiden University Center for Linguistics as a PhD researcher. This dissertation is the result of her research. In 2014, she was also working as a part-time researcher in the Informational Structure group SFB 632 at Potsdam University in Germany.