This work explores the interaction between tone and stress in contour tone languages. Mandarin tone 2 (hence after T$_2$) sandhi is focused as this sandhi process is uniformly agreed as being induced by stress. In order to disambiguate the basic facts of this process, an experiment was designed and conducted. Based on the analysis of the results, a prosodic licensing account of T$_2$ sandhi is provided and two types of relations between tone and stress in Mandarin are attested. Statistically, there is a relation between high register and prominence, between low register and non-prominence; between rising contour and prominence and between level tone and non-prominence. Dynamically, stress governs the realization of tone.

1. Introduction

It has been reported that there is a relation between prominence and tone in register tone languages: stressed position attracts high tone, and high tone attracts stress (Liberman 1975, Selkirk 1984, 1995, Goldsmith 1987); low tone attracts non-prominence and non-prominent position attracts low tone (de Lacy 1999). However, little is known about the relationship between tone and stress in contour tone languages. This is a work aimed at exploring this understudied topic. Mandarin is focused on because it is perhaps the most extensively studied contour tone language in the world and it has a comprehensive tonal inventory including rising, falling as well as level tones.

To examine the relation between tone and stress, knowledge of two types is crucial. First, how tones change in context; second, how stress is assigned in a certain domain. As the only tone sandhi process uniformly agreed as being triggered by stress, T$_2$ sandhi is the best place to explore the interaction between stress and tone in Mandarin. Unfortunately, controversies exist concerning the facts of this sandhi process and how stress is assigned at every level of the Mandarin prosodic hierarchy, from the foot up to the level of the phonological phrase. Therefore, this work starts from an experiment on T$_2$ sandhi and assumes that Mandarin builds trochaic feet. By providing a prosodic licensing account of T$_2$ sandhi, the interaction between tone and stress in Mandarin is discussed.

Specifically, I propose the following in this paper: (1) T$_2$ sandhi is a process involving both neutralization and assimilation. (2) T$_2$ sandhi applies only when T$_2$ syllables occur in prosodically non-prominent (unstressed) positions. (3) There is a relation between prominence (stress) and high register; between non-prominence (lack of stress) and low register; a relation between prominence and rising contour; and between non-prominence and level tone in contour tone languages. (4) All tone sandhi processes in Mandarin are stress-driven. (5) It is stress that governs the realization of tone: prominence makes tone high and rise; non-prominence makes tone level.

The paper is structured as follows. In section 2, I raise questions, both observational and theoretical, that arise from previous research on T$_2$ sandhi. In section 3, the design and results of the experiment are presented. In section 4, the experimental results are analyzed. A prosodic licensing account of T$_2$ sandhi is provided in section 5. In section 6, the interaction between tone and stress in Mandarin is summarized.

2. Mandarin T$_2$ sandhi

In this section, I raise questions concerning Mandarin T$_2$ sandhi. Before we delve into the details of this process, an introduction to the Mandarin tonal system is necessary.

---

Footnote: This paper is a summary of my PhD evaluation paper from McGill University. My deepest gratitude goes to Heather Goad, who guided me through every stage in conducting this research. Many of the ideas presented here are hers or inspired by the numerous discussions we had in her office. I also would like to thank Glyne Piggott for his thoughtful questions and comments on each draft of my evaluation paper. Thanks also go to the three participants of my experiment for their time and patience. Correspondence: email: chen.qu@mail.mcgill.ca.
2.1. Mandarin tonal system

Mandarin is a contour tone language with five tones in its tonal inventory. They are tone 1 (T₁: Hh), tone 2 (T₂: Hlh), tone 3 (T₃: Lhl), tone 4 (T₄: Hhl) and the neutral tone (T₀: L).¹ In Mandarin, all morphemes are monosyllabic in nature and each morpheme is specified as carrying a tone. T₁ to T₄ are carried by lexical morphemes; T₀ is carried by functional morphemes and lexical morphemes realized as reduced syllables.

There are four tone sandhi processes which are T₂ sandhi, T₃ sandhi, T₀ sandhi and yi-bu-qi-ba sandhi (see Chao 1968, Luo & Wang 1981 for detailed discussion). For the last three tone sandhi processes, shown in (1), almost no disagreement exists concerning the facts of the processes per se.²

(1) a. T₃ sandhi  
   \( T₃ \rightarrow T₂ / \_\_T₁ \) (Lhl \( \rightarrow \) Hlh / \_\_ Lhl)

b. T₀ sandhi
   i. T₀ \( \rightarrow \) mid / T₁ (or T₂) \_\_ (L \( \rightarrow \) Lh / Hh (or Hlh) \_\_)
   ii. T₀ \( \rightarrow \) low / T₄ \_\_ (L \( \rightarrow \) Ll / Hhl \_\_)
   iii. T₀ \( \rightarrow \) mid / T₃ \_\_ (L \( \rightarrow \) Lh / Lhl (or Lhlh?) \_\_)

c. yi-bu-qi-ba sandhi
   i. yi ("one"), qi ("seven"), ba ("eight")  
      T₁ \( \rightarrow \) T₂ / \_\_T₄ (Hlh \( \rightarrow \) Hhl / \_\_ Hhl)
   ii. bu ("not")  
      T₄ \( \rightarrow \) T₂ / \_\_T₄ (Hhl \( \rightarrow \) Hlh / \_\_ Hhl)

As far as T₂ sandhi is concerned, researchers disagree concerning the basic facts of this sandhi process. Up until now, four different observations have been reported by Luo & Wang (1981), Chao (1968), Wu (1984) and Chang (1992). Based on these observations, three stress-based accounts of T₂ sandhi have been proposed by Chang (1992), Duanmu (2000, 2007) and Chen (2000). In what follows, I introduce the research questions addressed in this paper based on a summary of the previous research on T₂ sandhi.

2.2. Research questions

According to Luo & Wang (1981), a T₂ (Hlh) carried by a medial syllable in a tri-syllabic expression is changed into T₁ (Hh) when both the initial and final syllables carry any tone except T₀ (L). They also suggest that T₂ sandhi is induced by stress.

According to Chao (1968), however, the initial syllable can only carry T₁ (Hh) or T₂ (Hlh). Chao also points out that T₂ sandhi is only attested in conversational speech. Chen (2000) and Duanmu (2000, 2007) adopt his observations and suggest further that T₂ sandhi is an assimilatory process in which T₂ (Hlh) is assimilated by the tone carried by the preceding syllable when it is unstressed.

Wu (1984) shares the same observation with Chao (1968) as far as the initial syllable is concerned. However, his observations differ from Chao and Luo & Wang in two respects. First, Wu points out that the final syllable also plays a role in triggering T₂ sandhi, i.e. T₂ (Hlh) is changed into T₄ (Hhl) when the final syllable carries T₂ (Hlh) or T₃ (Llh). Second, he proposes that different hierarchical structures of tri-syllabic expressions play a role in inducing T₂ sandhi, i.e. T₂ sandhi is mostly attested in left-branching structures ([[σ₁ σ₂] σ₃]) rather than right-branching ones ([σ₁ [σ₂ σ₃]]). Here and in the following discussion, both left- and right-branching corresponds to either morpho-syntactic or syntactic branching. In the experiment, the selected tri-syllabic stimuli with left-branching structures are all words, while those with right-branching structures are phrases.

¹ In this paper, by combining Bao’s (1990, 1999) tone structure model and Yip’s (1990) features for Mandarin tones, I propose that the five Mandarin tones are represented as shown. H and L refer to tonal registers; h and l are tonal features. For more discussion on the representation of these tones, see Qu (2007).

² For T₉ sandhi, as the focus in previous research is laid on how T₀ is phonetically realized in different contexts, controversies exist concerning the pitch carried by T₉ syllables. See Qu (2007) for more information.
Chang’s (1992) observations differ from those of all the previously mentioned researchers in three respects. First, Chang states that T₂ sandhi is sometimes attested in disyllabic expressions as well. Second, she points out that an initial T₁ (Lhl) rarely triggers T₂ sandhi, though an initial T₂ (Hhl) does. Third, Chang claims that T₂ sandhi is attested in fast speech, but she does not discuss how her fast speech differs from or is similar to Chao’s (1968) conversational speech. Chang also regards T₂ sandhi as being induced by stress, but she treats this sandhi process as neutralization in which T₁ (Hh) is the default tone given to unstressed syllables. Table 1 overleaf serves as a summary of the previous research concerning T₂ sandhi. Empty cells mean that the particular researcher does not address this issue. The disagreements lead to two types of research questions. The first type of question is observational as is listed in (2) and the second type of question is theoretical as in (3).

(2) Observational questions:
- a. Does T₂ sandhi apply in disyllabic expressions as well as in tri-syllabic expressions?
- b. What is the exact change T₂ undergoes?
- c. In tri-syllabic expressions, what are the tones carried by the initial and final syllables which can trigger T₂ sandhi?
- d. Do speech rate and hierarchical structure of the expression play a role in the application of this sandhi process?

(3) Theoretical questions:
- a. How is stress assigned in Mandarin?
- b. Is T₂ sandhi a process of assimilation or neutralization?

Table 1: Summary of previous research concerning T₂ sandhi

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tri-syllabic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>expressions only</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Initial syllable</td>
<td>T₁ T₂ T₃ T₄</td>
<td>T₁ T₂ T₄</td>
<td>T₁ T₂</td>
<td>T₁ T₂</td>
</tr>
<tr>
<td>Final syllable</td>
<td></td>
<td>T₁ T₂ T₃ T₄</td>
<td>T₁ T₂ T₃ T₄</td>
<td>T₁ T₂ T₄</td>
</tr>
<tr>
<td>Tonal change</td>
<td>T₁</td>
<td>T₁</td>
<td>T₁</td>
<td>T₁ T₄</td>
</tr>
<tr>
<td>Hierarchical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech rate</td>
<td>Fast speech</td>
<td>Conversational speech</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trigger</td>
<td>Stress</td>
<td>Stress</td>
<td>Stress (Chen 2000, Duanmu 2000, 2007)</td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td>Neutralization</td>
<td>Assimilation (Chen 2000, Duanmu 2000, 2007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Default tone: T₁)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. An experiment on T₂ sandhi

3.1. Design

In order to answer the observational questions listed in (2), I designed an experiment in which both language-internal and external factors were controlled. Concerning language-internal factors, both tri-syllabic and disyllabic expressions were included.

First, as far as the tri-syllabic expressions are concerned, the tone carried by the medial syllable was set to T₂. Initial and final syllables carried any tone except T₀. Expressions with either left- or right-branching structures with exactly the same tonal pattern were chosen, as is shown in (4).
(4) Tri-syllabic expressions
\[
[\sigma_1 \sigma_2 \sigma_3] \quad \left[ \sigma_1 \left[ \sigma_2 \sigma_3 \right] \right]
\]
\[
T_1 \ T_2 \ T_1 \quad T_1 \ T_2 \ T_1
T_2 \ T_2 \quad T_2 \ T_2
T_3 \ T_3 \quad T_3 \ T_3
T_4 \ T_4 \quad T_4 \ T_4
\]

Second, disyllabic expressions were of two kinds. The first group consisted of expressions with the initial syllable set to \(T_2\). In the second group, the final syllable was set to \(T_2\) as shown in (5).

(5) Disyllabic expressions:
\[
T_2 \ T_1 \quad T_1 \ T_2
T_2 \quad T_2
T_3 \quad T_3
T_4 \quad T_4
\]

Third, the \(T_2\) syllable under our focus was controlled as well. The onset consonant was limited to aspirated voiceless obstruents and the rhyme was restricted to /i/, /u/ and /a/. In order to control for a possible final lengthening effect, a carrier sentence was used, as can be seen in (6).

(6)
\[
\text{Shu1 shang4 mei2you3 __ __ __ san1 ge4 zi4.}
\]

There are no three characters in the book.

Concerning language-external factors, both subjects and speech rate were controlled. Three female native speakers aged 38yrs from the city of Harbin, China were chosen. They were asked to read each sentence in three speech rates: fast, normal and slow. All the analysis was done using Praat (Boersma and Weenink 2005). The pitch contour of the \(T_2\) syllable was examined in pitch tracking graphs, narrow band spectrograms and pulse intervals in expanded waveforms.

3.2. Results

No \(T_2\) sandhi was found in disyllabic expressions. \(T_2\) sandhi was attested in tri-syllabic expressions only in fast speech within the range of 500 – 690ms. The sandhi form of \(T_2\) is a level pitch. Concerning the hierarchical structure of the expressions, less \(T_2\) sandhi is found in right-branching structures than in left-branching ones. This is consistent with Wu’s observation. As for the tone carried by the initial syllable, \(T_2\) sandhi is attested when this syllable carries \(T_2\) regardless of hierarchical structure. An initial \(T_3\) never triggers \(T_2\) sandhi, consistent with Chang’s observation. Speaker variation is found when the initial syllable

---

3 This is a pilot study and many factors must be controlled, so only three subjects are included. More speakers will be tested in follow-up studies. Note that the productions of the three subjects are quite uniform. The reason that I did not select participants from the city of Beijing is that the inhabitants of Beijing are now heterogeneous in nature due to the increasingly large number of immigrants who have moved there from various places in China. This may have an effect on the dialects spoken in the city and thus on the application of \(T_2\) sandhi. For more information concerning subject selection, see Qu (2007).

4 Subjects were asked to control the three speech rates by themselves. Later in analyzing the data, only stimuli that fell between fixed durations were compared, as can be see in footnote 6.

5 There is one exception in Speaker 1’s fast speech: /teʰii2 pʰau2/ → [teʰii1 pʰau2] “cheong-sam” (women’s gown in Qing dynasty).

6 The three speakers varied as far as speech rate is concerned. Speaker 3 always spoke faster than the other two speakers. In her fast speech, most of the tri-syllabic expressions took less than 500ms, while, for the other two speakers, they were between 500ms and 690ms. In Speaker 3’s fast speech, the medial \(T_2\) in all stimuli underwent tone sandhi. Some of them became a level pitch, while others became falling. Because of this, we have compared her normal speech with the other two speakers’ fast speech. Her fast speech is only taken into consideration when a certain stimulus she produced was close to that produced by the other two speakers in terms of duration.
carries either $T_1$ or $T_4$ and when the final syllable carries $T_3$. The experimental results concerning tri-
 syllabic expressions are summarized in Table 2. (Y indicates that $T_2$ sandhi is attested, V stands for speaker
 variation where $T_2$ sandhi is attested in two speakers’ productions, and an empty cell indicates that no $T_2$
 sandhi is attested or that it is only attested in one speaker’s production.)

<table>
<thead>
<tr>
<th>Initial Syllable</th>
<th>Morpho-syntactic Structure</th>
<th>Final Syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$T_1$</td>
</tr>
<tr>
<td>$T_1$</td>
<td>$[[[\sigma_1, \sigma_2], \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td>$T_2$</td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td>$T_3$</td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td>$T_4$</td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>$[[\sigma_1, \sigma_2, \sigma_3]]$</td>
<td>V</td>
</tr>
</tbody>
</table>

4. Generalizations concerning $T_2$ sandhi

Turning to the theoretical questions in (3), Mandarin stress is a thorny issue. Disagreements exist for every
 level of prosodic structure, from the foot up to the phonological phrase. The prosodic hierarchy I follow is
 in (7) (Nespor & Vogel 1986, Selkirk 1986).

(7) Prosodic hierarchy (partial)
    Phonological Phrase (PPh)
    Prosodic Word (PWd)
    Foot (Ft)
    Syllable ($\sigma$)

By following Yip (1990, 1995) and Duanmu (1990, 1993, 2007), I regard Mandarin as a quantity-
sensitive language. However, different from their proposal, in addition to $T_0$ syllables, I also consider $T_3$
 (Lhl) syllables as light (See Qu 2007 for a detailed discussion on Mandarin quantity-sensitivity). Following
 Goad, White & Steele (2003) and Goad & White (2006), I assume that the Mandarin foot is optimally an
 uneven trochee, in spite of the fact that uneven trochees are highly marked according to Hayes (1995). As
 for word-level stress, I propose that it always falls on the left-most heavy syllable. The intuition of
 indeterminate word level stress in Mandarin, I suggest, is due to the different prosodic structures assigned
 to disyllabic words in combination with a final lengthening effect (see Qu 2007 for detailed discussion). As
 far as phrasal stress is concerned, I assume that it is assigned by End rule right (Hayes 1995).

4.1. Analysis of the experimental results: disyllabic expressions

Please recall that no $T_2$ sandhi is attested in disyllabic expressions except in one token (see footnote 5). The
disyllabic stimuli used in the experiment have either an initial $T_2$ or a final $T_2$. As $T_2$ (Hlh) syllables are
inherently heavy, in both cases, the $T_2$ syllables attract stress, as is shown by the structures in (8a) and
(8bii). Therefore, no $T_2$ sandhi is predicted to apply, consistent with the experimental results. The possible
reasons for the exception will be discussed later.
4.2. Analysis of the experimental results: tri-syllabic expressions

Recall that among the tri-syllabic stimuli, T₂ sandhi is attested more often when the expressions are left-branching than when they are right-branching. Those with left-branching structures are all compounds which are either [[N N] N]ₚ or [[A N] N]ₚ, while those with right-branching structures are all phrases which are either [A [N N]]ₚ or [V [N N]]ₚ. In the experiment, hierarchical structure failed to make a difference when the initial syllable carries T₂ (T₂ sandhi applies across the board) and when the initial syllable carries T₃ (no T₂ sandhi is attested). Speakers vary when the final syllable carries T₃.

4.2.1. Tri-syllabic expressions with left-branching structures

Consider the tri-syllabic expressions with left-branching structures first. As the medial syllable in our experiment is set to T₂, the prosodic structures of disyllabic words relevant to our discussion here are the ones with a final T₂, i.e., the structures in (8b). As far as the prosodic structure in (8b) is concerned, when a mono-syllabic morpheme is added to its right, the extrametrical foot can no longer be extrametrical due to the peripherality restriction on extrametricality (Hayes 1995: 57). If this mono-syllabic morpheme is a heavy syllable, it forms an extrametrical foot by itself as shown by (9) below. The stress clash that arises from the addition of the heavy syllable makes the medial syllable become light and thus results in an uneven trochee. In this case, T₂ sandhi is predicted to apply, consistent with the experimental results in

\[ \text{For T₂ syllables occurring in word-initial position, when the final syllable is light as in (8ai), a single foot is formed as Mandarin optimally builds uneven trochees. The initial T₂ syllable is the head of the foot because of the weight-to-stress principle (Kager 1999: 172). In (8aii) when the final syllable is also heavy, two feet are formed. However, because of the constraint of non-finality (Kager 1999: 165), the final foot is extrametrical. Therefore, word-level stress falls on the left-most heavy syllable.} \]

\[ \text{In (8b), T₂ syllables occur in word-final position. When the initial syllable is also heavy, the final foot is extrametrical as is shown by the structure in (8bi). However, when the initial syllable is light, the final foot cannot be extrametrical due to the non-exhaustivity restriction on extrametricality (Hayes 1995: 58). In both cases, word-level stress falls on the left-most heavy syllable as well. In consequence, the different prosodic structures proposed for Mandarin disyllabic words, together with a possible final lengthening effect, account for native speakers' intuition of indeterminate word-level stress.} \]
which T₂ sandhi is attested when both the initial and final syllables are heavy, i.e. carrying T₁ (Hh), T₂ (Hlh) and T₄ (Hhl). I will discuss later why no T₂ sandhi is attested when the tone pattern is T₃ T₂ T₄.

(9)

When the final syllable is light, i.e. carrying T₃ (Lhl), two prosodic structures are possible as is shown in (10). Since a sequence of light-heavy-light syllables is more rhythmic than a sequence of heavy-light-light syllables, (10a) is more likely to be attested than (10b) where T₂ sandhi is predicted not to apply. This is exactly what was found in the experiment.

(10) a.

As far as the structure in (8bii) is concerned, when a mono-syllabic morpheme is added to its right, no matter whether it is a heavy or light syllable, the second syllable will not be affected as is shown in (11). Thus, no T₂ sandhi is predicted to apply, consistent with the experimental results: no T₂ sandhi is attested when the initial syllable carries T₃ (Lhl).
Please recall from Table 2 that speaker variation is mostly attested when the final syllable carries $T_3$. This is due to the special properties $T_3$ syllables have. They can be realized either as heavy (carrying Lhlh) or light (carrying Lhl) in phrase-final position. As a result, when the final $T_3$ syllable is heavy, i.e. carrying Llh, the medial $T_2$ (Hlh) syllable is predicted to undergo $T_2$ sandhi as is shown in (9). When the final $T_3$ syllable is light, i.e. carrying Lhl, $T_2$ sandhi is predicted not to apply as is shown in (10) and (11b). The predictions are consistent with the experimental results. Speakers who failed to apply $T_2$ sandhi to this group of expressions pronounce the final $T_3$ as Lhl, while speakers who do apply $T_2$ sandhi pronounce the final $T_3$ as Lhlh as shown by the examples in (12).

(12) a. /kʷanHh tʂʰaαHhl sʰλLhl/ → [kʷanHh tʂʰaαHhl sʰλLhlh] (Speakers 2 & 3) → [kʷanHh tʂʰaαHhl sʰλLhl] (Speaker 1) “Office of Observation”

b. /lai4 pʰi2 kǒu3/ → [lai4 pʰiHh kǒuLhlh] (Speakers 1 & 3) → [lai4 pʰiHh kǒuLhl] (Speaker 2) ‘rascal’

4.2.2. Tri-syllabic expressions with right-branching structures

Concerning tri-syllabic stimuli with right-branching structures, the prosodic structures of the disyllabic words relevant to our discussion are in (8a). Please recall that I assume that Mandarin phrasal stress is assigned by End rule right. Therefore, the medial syllable always receives phrasal stress regardless of the status of both the initial and final syllables as can be seen from the structures in (13). As a result, no $T_2$ sandhi is predicted to apply in this group of tri-syllabic expressions.
Turning to the experimental results, no \( T_2 \) sandhi is attested except when the initial syllable carries \( T_2 \) (Hlh) and when the \( T_2 \) syllable is surrounded by two \( T_1 \)'s (Hh), almost consistent with our predictions. Consider the expressions with an initial \( T_2 \) first. According to the markedness scale for tones in (14) (Yip 2002), \( T_2 \) is the most marked tone in Mandarin. Therefore, when two \( T_2 \)'s are next to each other, it is not surprising that the medial \( T_2 \) undergoes change due to the obligatory contour principle (OCP).

(14) Markedness scale for tone (in descending order) (Yip 2002):
Contour > Level, Rising > Falling, High > Low

Turning now to the tone pattern of \( T_1 \) \( T_2 \) \( T_1 \), I regard the tonal change in this case, shown in (15a), as phonetic. In fact, a low tone becoming high when surrounded by high tones is a widely attested process in African languages (see (15b)) and has been argued by Schuh (1978) to be due to physiological constraints.

(15) a. \( T_2 \rightarrow T_1 / T_1 \_ \_ \_ \_ T_1 \) (Hlh \( \rightarrow \) Hh / Hh \_ \_ Hh)  
   b. Compare: low tone \( \rightarrow \) high / high \_ \_ high

Recall that no \( T_2 \) sandhi is attested in the expression with left-branching structures when \( T_2 \) is surrounded by two \( T_8 \). \( T_2 \) sandhi is, however, predicted to apply in this case. I consider this to be a phonetic effect as well. The phonetic pitch of the medial \( T_2 \) syllable is influenced by the pitch of its surrounding tones as schematized in (16), thereby, disguising the application of the phonological process of \( T_2 \) sandhi.

(16) \( T_4 \) \( T_2 \) \( T_4 \)

\[ \text{H = tone register, C = contour, and h, l = tonal features. Please note that the dependents of C are on the same tier.} \]
Note that speaker variation is predicted not to exist in expressions with right-branching structures, but it is attested when the initial syllable carries T₄ and the final syllable carries either T₁ or T₁. I consider this also as a phonetic effect. As is well-known, the vowel /i/ is cross-linguistically shorter than other vowels. In the expressions under concern, the initial syllables contain a diphthong and the medial syllable contains /i/. To stress the medial syllable means that it must be lengthened, because duration has been experimentally shown to be the principal phonetic correlate of stress in Mandarin (Lin, Yan & Sun 1984). As a result, special emphasis must be placed on the second syllable in order to lengthen the vowel /i/ when it is adjacent to a syllable containing a diphthong. In fact, this is exactly what happened in Speaker 2’s production: T₂ sandhi fails to apply and special emphasis is attested on the medial syllable as can be seen in (17a-b). This influence of the surrounding vowels on the application of T₂ sandhi also accounts for the exception attested in the disyllabic stimuli (see footnote 5), as is shown in (17c).

(17) a. /tʃʰaiHhl tʃʰiiHlh puu3/ → i. [tʃʰaiHhl tʃʰiiHlh puu3] (Speakers 1 & 3)
ii. [tʃʰaiHhl tʃʰiiHlh puu3] (Speaker 2)
   “used chess manual”

b. /pʰaiHhl tʃʰiiHlh piŋHhl/ → i. [pʰaiHhl tʃʰiiHlh piŋHhl] (Speakers 1 & 3)
ii. [pʰaiHhl tʃʰiiHlh piŋHhl] (Speaker 2)
   “to send soldiers by surprise”

c. /tʃʰiiHlh pʰauHlh/ → [tʃʰiiHlh pʰauHlh] (Speaker 1)
   “cheong-sam” (women’s gown in Qing dynasty)

Since overriding phonetic considerations make it look like T₂ sandhi has applied where it has not and make it look like the process has not applied where it has, Table 3 shows where T₂ sandhi applies as a phonological process by removing the effects of the overriding phonetic considerations. The tone patterns subject to change are dotted.

Table 3: Application of T₂ sandhi as a phonological process

<table>
<thead>
<tr>
<th>Initial Syllable</th>
<th>Morpho-syntactic Structure</th>
<th>Final Syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T₁</td>
</tr>
<tr>
<td>T₁</td>
<td>[[σ₁, σ₂] σ₁]</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>[σ₁, σ₂]</td>
<td>*</td>
</tr>
<tr>
<td>T₂</td>
<td>[[σ₁, σ₂] σ₁]</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>[σ₁, σ₂]</td>
<td>Y</td>
</tr>
<tr>
<td>T₃</td>
<td>[[σ₁, σ₂] σ₁]</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>[σ₁, σ₂]</td>
<td></td>
</tr>
<tr>
<td>T₄</td>
<td>[[σ₁, σ₂] σ₁]</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>[σ₁, σ₂]</td>
<td>*</td>
</tr>
</tbody>
</table>

Recall from Table 1 that disagreement exists concerning what kind of process T₂ sandhi is. Some researchers have proposed that T₂ sandhi involves neutralization (Chang 1992), while others have suggested that it involves assimilation (Duannu 2000, 2007, Chen 2000). I propose that it is actually both. As far as the process of neutralization is concerned, when a syllable is unstressed, the structure below its contour node is de-linked resulting in a bare contour node. Then the bare contour node triggers spreading of the rightmost tonal feature of the preceding syllable resulting in a level pitch, either high or mid (Hh after T₁(Hh)/T₂(Hlh), Hl after T₄(Hhl)), which captures the process of assimilation. The two processes are shown in (18).
5. A prosodic licensing account of T₂ sandhi

In Mandarin, only T₂ (Hlh) but not T₁ (Hh) or T₄ (Hhl) undergoes tone sandhi when unstressed. No stress account can offer a satisfactory explanation to this fact. A prosodic licensing account of T₂ sandhi must be assumed. According to Harris (1997), prosodic heads support more complex structures than non-heads. Dresher and van der Hulst (1998) propose that the dependent in any constituent cannot be more complex than its head. Concerning T₂ sandhi, T₂ is the most marked tone in Mandarin. As a result, it can only be licensed by a prosodic head. Building on Yip’s (2002) markedness scale for tones provided in (14), I propose the markedness scale for Mandarin in (19), as there is no evidence in Mandarin to suggest that T₄ (Hhl) is more marked than T₁ (Hh).

(19) Markedness scale for Mandarin tones (in descending order)
T₂ (Hlh) > T₄ (Hhl), T₁ (Hh) > T₃ (Lhl) > T₀ (L)

The distribution of the five Mandarin tones within a foot shows why no other tone sandhi processes are attested in the same environment where T₂ sandhi is found.

(20) Distribution of Mandarin tones within a foot:

\[
\begin{array}{c}
\text{Ft} \\
\sigma_1 \\
\sigma_2 \\
T₂ (Hlh) \\
T₃ (Hhl) \\
T₁ (Hh) \\
T₃ (Lhl) \\
T₀ (L)
\end{array}
\]

As no low register tones (T₃ and T₀) occur in prosodic head position, T₂ (Hlh) is the only possible tone which, when occurring in non-head position, could result in a structure more complex than its head. Consequently, only T₂ must undergo sandhi when occurring in non-head position.

6. Relationship between stress and tone in contour tone languages

The relationship between stress and tone in contour tone languages has already been made evident by the distribution of Mandarin tones within a foot. Since only high register tones occur in prosodic head position and low register tones never do, a static relation is attested between high register and prominence, and between low register and non-prominence. Note that we are not taking T₃ (Llh) into account, as it is a complex contour though in low register and it never occurs in head position within a foot. There is also a relation held between rising contour and prominence and between level tone and non-prominence, as the complex contour of T₂ (Hlh) can only be licensed by prosodic head position and it becomes level when in non-head position. It is without a doubt, then, that T₂ sandhi is stress-driven. In fact, other tone sandhi processes in Mandarin are also stress-driven.

Recall from Section 2.1 that, besides T₂ sandhi, there are three other tone sandhi processes in Mandarin. As far as T₃ sandhi is concerned, it is a process in which the unstressed T₃ syllable acquires its pitch from the prosodic head within its local domain. As for T₃ sandhi, when two T₃ (Lhl) syllables are next to each other, the first T₃ becomes T₂ (Hlh) resulting in a heavy syllable and thus an uneven trochee due to the prosodic head position this word-initial syllable occupies. In fact, not only T₃ (Lhl), but also T₁ (Hh)
and $T_4$ (Hhl) are sometimes realized as $T_2$ (Hlh) depending on the tones carried by the following syllable as in yi-bi-qi-ba sandhi. Thus, it is stress which regulates the realization of each tone in Mandarin. The dynamic relation held between tone and stress in Mandarin is: prominent position makes pitch high and rising; non-prominent position makes pitch level.

In conclusion, both static and dynamic relations between tone and stress are attested in Mandarin. Statically, there is a relation between tone register (high register) and prominence (stress) and between contour (rising tone) and prominence (stress). Dynamically, it is prominence (stress) which determines tonal realizations. Prominent (stressed) positions make tone high and rising and non-prominent positions make tone level.

References

De Lacy, Paul. 1999. Tone and prominence. MS. University of Massachusetts, Amherst. ROA#333.