ON THE INTERACTION OF TONE AND STRESS
IN CONTOUR TONE LANGUAGES[*]

Chen Qu
McGill University

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. Statically, 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.

2.1. Mandarin tonal system

Mandarin is a contour tone language with five tones in its tonal inventory. They are tone 1 (T_1: Hh), tone 2 (T_2: Hlh), tone 3 (T_3: Lhl), tone 4 (T_4: Hhl) and the neutral tone (T_0: L).[1] In Mandarin, all morphemes are monosyllabic in nature and each morpheme is specified as carrying a tone. T_1 to T_4 are carried by lexical morphemes; T_0 is carried by functional morphemes and lexical morphemes realized as reduced syllables. There are four tone sandhi processes which are T_2 sandhi, T_3 sandhi, T_0 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. T_3 sandhi
> $T_3 \rightarrow T_2$ / ___T_3 (Lhl $\rightarrow$ Hlh / ___ Lhl)
>
>   b. T_0 sandhi
>   i. $T_0 \rightarrow$ mid / T_1 (or T_2) ___ (L $\rightarrow$ Lh / Hh (or Hlh) ___)
>   ii. $T_0 \rightarrow$ low / T_4 ___ (L $\rightarrow$ Ll / Hhl ___)
>   iii. $T_0 \rightarrow$ mid / T_3 ___ (L $\rightarrow$ Lh / Lhl (or Lhlh?) ___)
>
>   c. yi-bu-qi-ba sandhi
>   i. yi ("one"), qi ("seven"), ba ("eight")
> $T_1 \rightarrow T_2$ / ___T_4 (Hh $\rightarrow$ Hlh / ___ Hhl)
>   ii. bu ("not")
> $T_4 \rightarrow T_2$ / ___T_4 (Hhl $\rightarrow$ Hlh / ___ Hhl)

As far as T_2 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 T_2 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 T_2 sandhi.

2.2. Research questions

According to Luo & Wang (1981), a T_2 (Hlh) carried by a medial syllable in a tri-syllabic expression is changed into T_1 (Hh) when both the initial and final syllables carry any tone except T_0 (L). They also suggest that T_2 sandhi is induced by stress.

According to Chao (1968), however, the initial syllable can only carry T_1 (Hh) or T_2 (Hlh). Chao also points out that T_2 sandhi is only attested in conversational speech. Chen (2000) and Duanmu (2000, 2007) adopt his observations and suggest further that T_2 sandhi is an assimilatory process in which T_2 (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 T_2 sandhi, i.e. T_2 (Hlh) is changed into T_4 (Hhl) when the final syllable carries T_2 (Hlh) or T_3 (Lhl). Second, he proposes that different hierarchical structures of tri-syllabic expressions play a role in inducing T_2 sandhi, i.e. T_2 sandhi is mostly attested in left-branching structures ([[σ_1 σ_2] σ_3]) rather than right-branching ones ([σ_1 [σ_2 σ_3]]). 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 T_0 sandhi, as the focus in previous research is laid on how T_0 is phonetically realized in different contexts, controversies exist concerning the pitch carried by T_3 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 T_2 sandhi is sometimes attested in disyllabic expressions as well. Second, she points out that an initial T_3 (Lhl) rarely triggers T_2 sandhi, though an initial T_4 (Hhl) does. Third, Chang claims that T_2 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 T_2 sandhi as being induced by stress, but she treats this sandhi process as neutralization in which T_1 (Hh) is the default tone given to unstressed syllables. Table 1 overleaf serves as a summary of the previous research concerning T_2 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 T_2 sandhi apply in disyllabic expressions as well as in tri-syllabic expressions?
   b. What is the exact change T_2 undergoes?
   c. In tri-syllabic expressions, what are the tones carried by the initial and final syllables which can trigger T_2 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 T_2 sandhi a process of assimilation or neutralization?

Table 1: Summary of previous research concerning T_2 sandhi

	Luo & Wang (1981)	Chang (1992)	Chao (1968)	Wu (1984)
Tri-syllabic expressions only	Yes	No	Yes	Yes
Initial syllable	$T_1 T_2 T_3 T_4$	$T_1 T_2 T_4$	$T_1 T_2$	$T_1 T_2$
Final syllable	$T_1 T_2 T_3 T_4$	$T_1 T_2 T_3 T_4$	$T_1 T_2 T_3 T_4$	$T_1 T_4$ ($T_2 \rightarrow T_1$) $T_2 T_3$ ($T_2 \rightarrow T_4$)
Tonal change	T_1	T_1	T_1	$T_1 T_4$
Hierarchical structure				$[[\sigma_1\ \sigma_2]\ \sigma_3]$
Speech rate		Fast speech	Conversational speech	
Trigger	Stress	Stress	Stress (Chen 2000, Duanmu 2000, 2007)	
Process		Neutralization (Default tone: T_1)	Assimilation (Chen 2000, Duanmu 2000, 2007)	

3.   An experiment on T_2 sandhi

3.1.   Design

In order to answer the observational questions listed in (2), I designed an experi-ment 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 T_2. Initial and final syllables carried any tone except T_0. 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

$[[\sigma_1\ \sigma_2]\ \sigma_3]$		$[\sigma_1\ [\sigma_2\ \sigma_3]]$	
$T_1\ T_2\ T_1$		$T_1\ \ T_2\ T_1$	
T_2	T_2	T_2	T_2
T_3	T_3	T_3	T_3
T_4	T_4	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:

$T_2\ T_1$	$T_1\ T_2$
T_2	T_2
T_3	T_3
T_4	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) 书　上　　没 有 ＿＿ ＿ 三　个　字。
Shu1 shang4　mei2you3 ＿ ＿ ＿ san1 ge4 zi4.
[ʂuʊʂaŋ　　məɨ jəu ＿ ＿ ＿ san kə t͡sii]
Book up no have ＿ ＿ ＿ three CL character
"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.[3] They were asked to read each sentence in three speech rates: fast, normal and slow.[4] 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.[5] T_2 sandhi was attested in tri-syllabic expressions only in fast speech within the range of 500 – 690ms.[6] 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: /tɕʰii2 pʰaʊ2/ → [tɕʰii1 pʰaʊ2] "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.
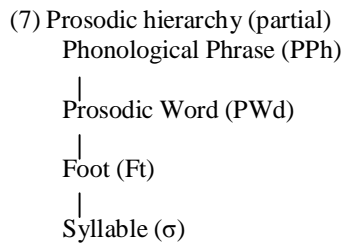
carries either T_1 or T_4 and when the final syllable carries T_3. The experimental results concerning tri-syllabic expressions are summarized in Table 2. (Y indicates that T_2 sandhi is attested, V stands for speaker variation where T_2 sandhi is attested in two speakers' productions, and an empty cell indicates that no T_2 sandhi is attested or that it is only attested in one speaker's production.)

Table 2: Summary of the experimental results

Initial Syllable	Morpho-syntactic Structure	Final Syllable			
		T_1	T_2	T_3	T_4
T_1	$[[\sigma_1 \, \sigma_2] \, \sigma_3]$	Y	Y	V	Y
	$[\sigma_1 \, [\sigma_2 \, \sigma_3]]$	Y			
T_2	$[[\sigma_1 \, \sigma_2] \, \sigma_3]$	Y	Y	Y	Y
	$[\sigma_1 \, [\sigma_2 \, \sigma_3]]$	Y	Y	Y	Y
T_3	$[[\sigma_1 \, \sigma_2] \, \sigma_3]$				
	$[\sigma_1 \, [\sigma_2 \, \sigma_3]]$				
T_4	$[[\sigma_1 \, \sigma_2] \, \sigma_3]$	Y	Y	V	
	$[\sigma_1 \, [\sigma_2 \, \sigma_3]]$	V		V	

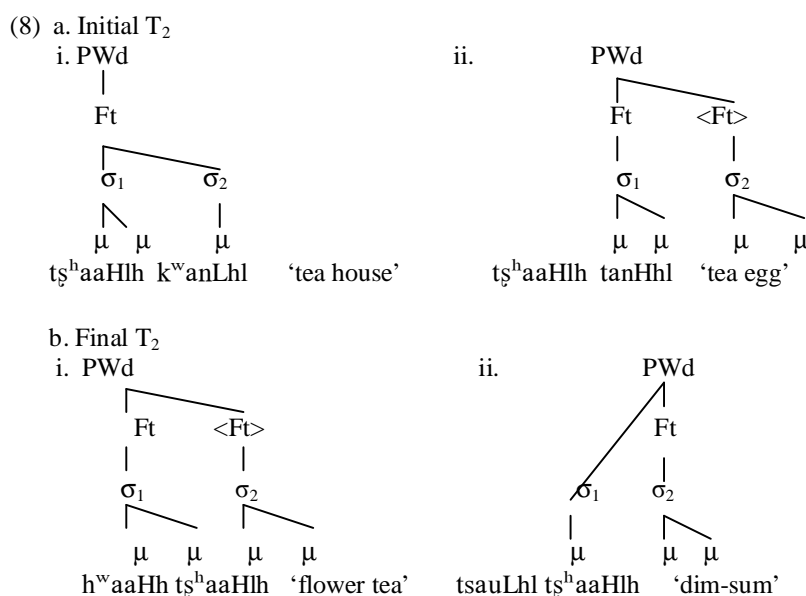## 4. Generalizations concerning T_2 sandhi

Turning to the theoretical questions in (3), Mandarin stress is a thorny issue. Disa-greements 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 (σ)

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.

(8) a. Initial T_2

  i. PWd
    |
  Ft
    σ_1 σ_2
    μ μ μ

tṣʰaaHlh kʷanLhl 'tea house'

  ii.   PWd
    Ft   <Ft>
    σ_1 σ_2
    μ μ μ μ

tṣʰaaHlh tanHhl 'tea egg'

  b. Final T_2

  i. PWd
    Ft   <Ft>
    σ_1 σ_2
    μ μ μ μ

hʷaaHh tṣʰaaHlh 'flower tea'

  ii.   PWd
         Ft
    σ_1 σ_2
    μ μ μ

tsauLhl tṣʰaaHlh 'dim-sum'

For T_2 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 T_2 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 (Kager1999:165), the final foot is extrametrical. Therefore, word-level stress falls on the left-most heavy syllable.[7] In (8b), T_2 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.

4.2. Analysis of the experimental results: tri-syllabic expressions

Recall that among the tri-syllabic stimuli, T_2 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]$_N$ or [[A N] N]$_N$, while those with right-branching structures are all phrases which are either [A [N N]]$_{np}$ or [V [N N]]$_{vp}$. In the experiment, hierarchical structure failed to make a difference when the initial syllable carries T_2 (T_2 sandhi applies across the board) and when the initial syllable carries T_3 (no T_2 sandhi is attested). Speakers vary when the final syllable carries T_3.
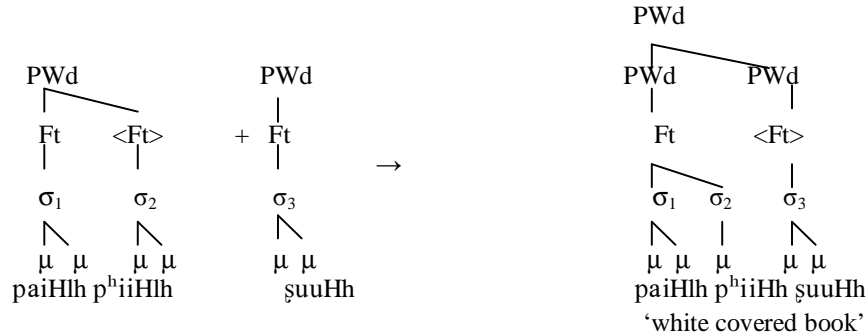
4.2.1. Tri-syllabic expressions with left-branching structures

Consider the tri-syllabic expressions with left-branching structures first. As the medial syllable in our experiment is set to T_2, the prosodic structures of disyllabic words relevant to our discussion here are the ones with a final T_2, i.e. the structures in (8b). As far as the prosodic structure in (8bi) 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 mor-pheme 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, T_2 sandhi is predicted to apply, consistent with the experimental results in

---

[7] Recall that in the experimental design, no T_0 syllables were included. Therefore, the final light syllable in (8ai) only carries T_3 (Lhl). In (8aii), the final syllable can also carry T_1, T_2 and T_3 when heavy (see below in text).
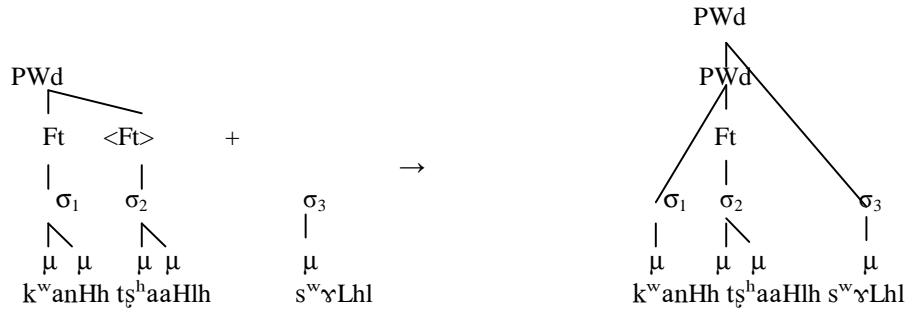
which T_2 sandhi is attested when both the initial and final syllables are heavy, i.e. carrying T_1 (Hh), T_2 (Hlh) and T_4 (Hhl). I will discuss later why no T_2 sandhi is attested when the tone pattern is T_4 T_2 T_4.

(9)

```
      PWd              PWd                              PWd
     /  \               |                           /        \
   Ft   <Ft>   +        Ft           →           PWd          PWd
    |     |              |                        |            |
   σ₁    σ₂             σ₃                        Ft          <Ft>
   /\    /\             /\                       /  \           |
  μ μ   μ μ            μ μ                      σ₁   σ₂         σ₃
  paiHlh pʰiiHlh       ṣuuHh                    /\   |          /\
                                              μ μ   μ         μ μ
                                            paiHlh pʰiiHh ṣuuHh
                                             'white covered book'
```
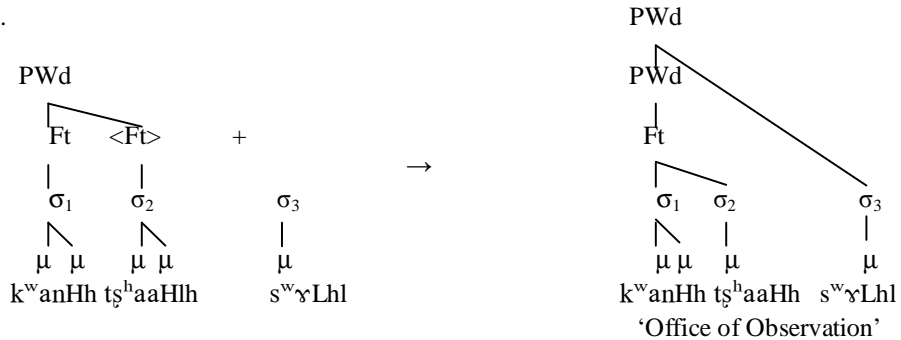
When the final syllable is light, i.e. carrying T_3 (Lhl), two prosodic structures are possible as is shown in (10). Since a sequence of light-heavy-light syllables is more rhythmic than a sequence of heavy-light-light syllables, (10a) is more likely to be attested than (10b) where T_2 sandhi is predicted not to apply. This is exactly what was found in the experiment.

(10) a.

```
      PWd                                          PWd
     /   \                                        /   \
   Ft   <Ft>    +                              PWd     \
    |     |                                     |       \
   σ₁    σ₂            σ₃         →            Ft        \
   /\    /\            |                       |          \
  μ μ   μ μ           μ                       σ₁  σ₂       σ₃
 kʷanHh tṣʰaaHlh    sʷɤLhl                    |  /\        |
                                             μ  μ μ        μ
                                           kʷanHh tṣʰaaHlh sʷɤLhl
```

b.

```
      PWd                                          PWd
     /   \                                        /    \
   Ft   <Ft>    +                              PWd      \
    |     |                                     |         \
   σ₁    σ₂            σ₃         →            Ft          \
   /\    /\            |                      /  \          \
  μ μ   μ μ           μ                      σ₁   σ₂         σ₃
 kʷanHh tṣʰaaHlh    sʷɤLhl                  /\   |           |
                                          μ μ   μ           μ
                                        kʷanHh tṣʰaaHh  sʷɤLhl
                                          'Office of Observation'
```
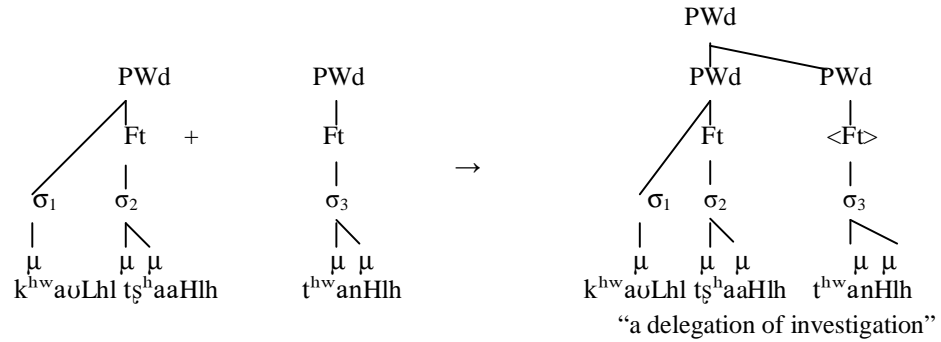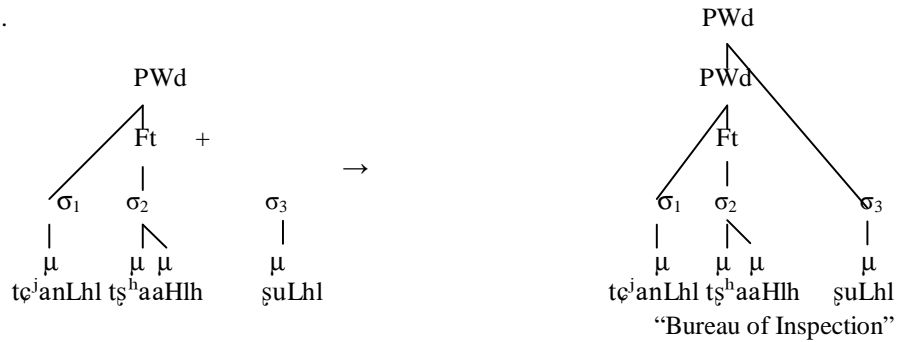
As far as the structure in (8bii) is concerned, when a mono-syllabic morpheme is added to its right, no matter whether it is a heavy or light syllable, the second syllable will not be affected as is shown in (11). Thus, no T_2 sandhi is predicted to apply, consistent with the experimental results: no T_2 sandhi is attested when the initial syllable carries T_3 (Lhl).

(11) a.

```
        PWd          PWd                                      PWd
         |            |                                      /    \
        /Ft    +      Ft              →              PWd        PWd
         |            |                             /Ft         <Ft>
       σ₁  σ₂        σ₃                          σ₁  σ₂          σ₃
        |   /\        /\                          |   /\          /\
        μ  μ μ       μ μ                          μ  μ μ         μ μ
   kʰʷaʊLhl tʂʰaaHlh  tʰʷanHlh           kʰʷaʊLhl tʂʰaaHlh  tʰʷanHlh
                                              "a delegation of investigation"
```

b.

```
        PWd                                          PWd
         |                                          /   \
        /Ft    +                              PWd
         |                     →              /Ft
       σ₁  σ₂      σ₃                       σ₁  σ₂        σ₃
        |   /\      |                        |   /\        |
        μ  μ μ      μ                        μ  μ μ        μ
   tɕʲanLhl tʂʰaaHlh  ʂuLhl          tɕʲanLhl tʂʰaaHlh  ʂuLhl
                                           "Bureau of Inspection"
```
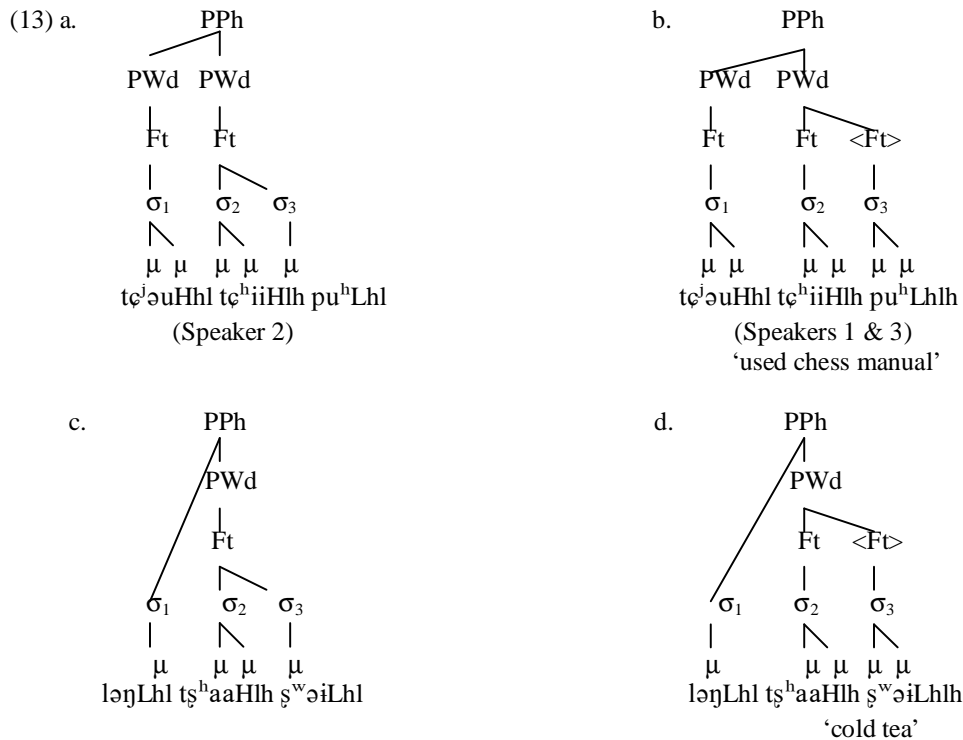
Please recall from Table 2 that speaker variation is mostly attested when the final syllable carries T_3. This is due to the special properties T_3 syllables have. They can be realized either as heavy (carrying Lhlh) or light (carrying Lhl) in phrase-final position. As a result, when the final T_3 syllable is heavy, i.e. carrying Lhlh, 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 tʂʰaaHlh sʷɤLhl/ → [kʷanHh tʂʰaaHh sʷɤLhlh] (Speakers 2 & 3)

→ [kʷanHh tʂʰaaHlh sʷɤLhl] (Speaker 1)

"Office of Observation"

b. . /lai4 pʰii2 kəu3/ → [lai4 pʰiHh kəuLhlh] (Speakers 1 & 3)

→ [lai4 pʰiiHlh kəuLhl] (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.
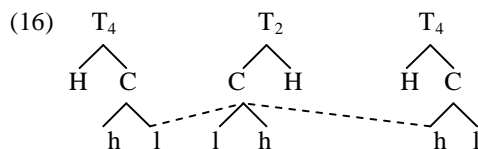
(13) a.

PPh
PWd PWd
Ft Ft
σ₁ σ₂ σ₃
μ μ μ μ μ

tɕʲəuHhl tɕʰiiHlh puʰLhl
(Speaker 2)

b.

PPh
PWd PWd
Ft Ft <Ft>
σ₁ σ₂ σ₃
μ μ μ μ μ μ

tɕʲəuHhl tɕʰiiHlh puʰLhlh
(Speakers 1 & 3)
'used chess manual'

c.

PPh
PWd
Ft
σ₁ σ₂ σ₃
μ μ μ μ

ləŋLhl tʂʰaaHlh ʂʷəɨLhl

d.

PPh
PWd
Ft <Ft>
σ₁ σ₂ σ₃
μ μ μ μ μ

ləŋLhl tʂʰaaHlh ʂʷəɨLhlh
'cold tea'

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_1s (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_2s 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 → Hh / Hh ___ Hh)
b. Compare: low tone → 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_4s. 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.[8]

(16) T₄ T₂ T₄
H C C H H C
h l l h h l

---

[8] 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 T_4 and the final syllable carries either T_3 or T_1. 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: T_2 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 T_2 sandhi also accounts for the exception attested in the disyllabic stimuli (see footnote 5), as is shown in (17c).

(17) a. /tɕʲəʊHhl tɕʰiiHlh puu3/ → i. [tɕʲəʊHhl tɕʰiHh puu3] (Speakers 1 & 3)
ii. [tɕʲəʊHhl tɕʰiiHlh puu3] (Speaker 2)
"used chess manual"

b. /pʰaɨHhl tɕʰiiHlh pɨŋHh/ → i. [pʰaɨHhl tɕʰiHh pɨŋHh] (Speakers 1 & 3)
ii. [pʰaɨHhl tɕʰiiHlh pɨŋHh] (Speaker 2)
"to send soldiers by surprise"

c. /tɕʰiiHlh pʰaʊHlh/ → [tɕʰiiHh pʰaʊHlh] (Speaker 1)
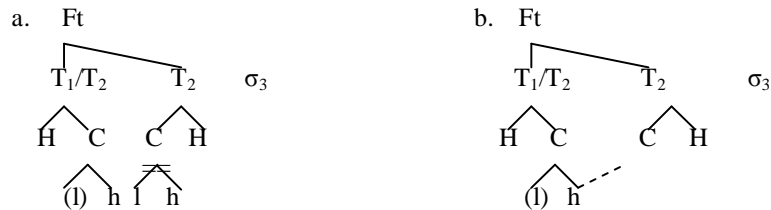"cheong-sam" (women's gown in Qing dynasty)

Since overriding phonetic considerations make it look like T_2 sandhi has applied where it has not and make it look like the process has not applied where it has, Table 3 shows where T_2 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 T_2 sandhi as a phonological process

Initial Syllable	Morpho-syntactic Structure	Final Syllable			
		T_1	T_2	T_3	T_4
T_1	[[σ₁ σ₂] σ₃]	Y	Y	V	Y
	[σ₁ [σ₂ σ₃]]	•			
T_2	[[σ₁ σ₂] σ₃]	Y	Y	Y	Y
	[σ₁ [σ₂ σ₃]]	Y	Y	Y	Y
T_3	[[σ₁ σ₂] σ₃]				
	[σ₁ [σ₂ σ₃]]				
T_4	[[σ₁ σ₂] σ₃]	Y	Y	V	• Y
	[σ₁ [σ₂ σ₃]]	•		•	

Recall from Table 1 that disagreement exists concerning what kind of process T_2 sandhi is. Some researchers have proposed that T_2 sandhi involves neutralization (Chang 1992), while others have suggested that it involves assimilation (Duanmu 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 T_1(Hh)/T_2(Hlh), Hl after T_4(Hhl)), which captures the process of assimilation. The two processes are shown in (18).

(18) T_2 sandhi: neutralization + assimilation
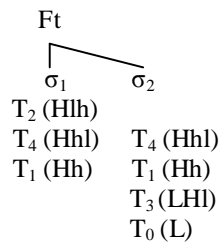

5. A prosodic licensing account of T_2 sandhi

In Mandarin, only T_2 (Hlh) but not T_1 (Hh) or T_4 (Hhl) undergoes tone sandhi when unstressed. No stress account can offer a satisfactory explanation to this fact. A prosodic licensing account of T_2 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 T_2 sandhi, T_2 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 T_4 (Hhl) is more marked than T_1 (Hh).

(19) Markedness scale for Mandarin tones (in descending order)
T_2 (Hlh) > T_4 (Hhl), T_1 (Hh) > T_3 (Lhl) > T_0 (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 T_2 sandhi is found.

(20) Distribution of Mandarin tones within a foot:

	σ_1	σ_2
	T_2 (Hlh)	
	T_4 (Hhl)	T_4 (Hhl)
	T_1 (Hh)	T_1 (Hh)
		T_3 (LHl)
		T_0 (L)

As no low register tones (T_3 and T_0) occur in prosodic head position, T_2 (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 T_2 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 T_3 (Lhlh) 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 T_2 (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 T_2 sandhi is stress-driven. In fact, other tone sandhi processes in Mandarin are also stress-driven.

Recall from Section 2.1 that, besides T_2 sandhi, there are three other tone sandhi processes in Mandarin. As far as T_0 sandhi is concerned, it is a process in which the unstressed T_0 syllable acquires its pitch from the prosodic head within its local domain. As for T_3 sandhi, when two T_3 (Lhl) syllables are next to each other, the first T_3 becomes T_2 (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 T_3 (Lhl), but also T_1 (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 attes-ted in Mandarin. Statically, there is a relation between tone register (high register) and prominence (stress) and between contour (rising tone) and prominence (stress). Dynami-cally, it is prominence (stress) which determines tonal realizations. Prominent (stressed) positions make tone high and rising and non-prominent positions make tone level.

References

Bao, Zhiming. 1990. On the nature of tone. PhD dissertation. MIT.
    1999. The structure of tone. Oxford: Oxford University Press.
Boersma, P. & Weenink, D. 2005. Praat: Doing phonetics by computer (version 4.3.22) [Computer program]. http://www.praat.org/.
Chao, Yuenren. 1968. A grammar of spoken Chinese. Berkeley and Los Angeles: University of California Press.
Chang, Laura Mei-Chih. 1992. A prosodic account of tone, stress, and tone sandhi in Chinese languages. PhD disserta-tion. University of Hawaii. [Published by Crain, Taipei (1993)].
Chen, Matthew Y. 2000. Tone Sandhi: Patterns across Chinese dialects. Cambridge: Cambridge University Press.
De Lacy, Paul. 1999. Tone and prominence. MS. University of Massachusetts, Amherst. ROA#333.
Duanmu, San. 1990. A formal study of syllable, tone, stress and domain in Chinese languages. PhD dissertation, MIT.
    1993. Rime length, stress, and association domains. Journal of East Asian Linguistics 2: 1–44.
    2000. The Phonology of Standard Chinese. Oxford: Oxford University Press.
    2007. The Phonology of Standard Chinese (2nd edition). Oxford: Oxford University Press.
Dresher, E. & H. G. van der Hulst. 1998. Head-dependent asymmetries in prosodic phonology: visibility and complex-ity. Phonology 15: 317–352.
Goad, H, White, L & Steele, J. 2003. Missing inflection in L2 acquisition: defective syntax or L1-constrained prosodic representations? Canadian Journal of Linguistics 48: 243–263.
Goad, H & White, L. 2006. Ultimate attainment of L2 inflection: Effects of L1 prosodic structure. Second Language Research 22: 243–268.
Goldsmith, J. 1987. Tone and accent, and getting the two together. BLS 13: 88–104.
Harris, John. 1997. Licensing inheritance: an integrated theory of neutralization. Phonology 14: 315–370.
Hayes, Bruce. 1995. Metrical Stress Theory: Principles and Case Studies. University of Chicago Press.
Kager, René. 1999. Optimality Theory. Cambridge University Press.
Liberman, M. 1975. The intonational system of English. Cambridge: MIT Press.
Lin, Maocan, Yan, Jingzhu & Sun, Guohua. 1984. Beijinghua Liangzizu Zhengchang Zhongyin de Chubu Shiyan [A preliminary experiment on normal stress in Beijing Mandarin bi-syllabic expressions.]. Fangyan 1: 57–73.
Luo, Changpei & Wang, Jun. 1981. (2nd edition). Putong Yuyinxue Gangyao. [Outline of general phonetics]. Beijing: Shangwu Yinshuguan.
Nespor, M & Vogel, I. 1986. Prosodic phonology. Foris, Dordrecht.
Qu, Chen. 2007. A prosodic licensing account of Mandarin tone 2 sandhi: an interaction between tone and stress. McGill Evaluation Paper in Linguistics.
Selkirk, E. O. 1984. Phonology and syntax: The relation between sound and structure. Cambridge: MIT Press.
    1986. On derived domains in sentence phonology. Phonology Yearbook 3: 371–405.
    1995. The prosodic structure of function words. In Beckman, J. Dickey, L. W. & Urvanczyk, S. (eds.) Papers in Optimality Theory. Amherst, MA: GLSA, pp. 439–470.
Schuh, R. G. 1978. Tone rules. In Fromkin, V. A. Tone: A linguistic survey. pp. 77–112. New York: Academic Press, INC.
Wu, Zongji. 1984. Putonghua sanzizu biandiao guilü. [Tone sandhi of tri-syllabic expressions in standard Mandarin]. Chinese Linguistics Journal 2: 70–92.
Yip, Moira. 1990. The Tonal phonology of Chinese. New York: Garland Publishing 1980a. PhD Dissertation, MIT.
    1995. Tone in East Asian languages. In Goldsmith (ed.) The handbook of phonological theory. pp. 476-494. Oxford: Blackwell.
    2002. Tone. Cambridge: Cambridge University Press.