The Phonetics and Phonology of Tone Mapping in a Constraint-Based Approach

by

Zhiqiang Li

B.A., Tianjin Normal University (1992)
M.A., Beijing Foreign Studies University (1995)

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

at the

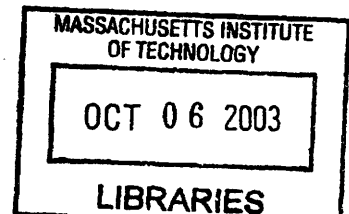MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2003

Author ....
／
Department of Linguistics and Philosophy
August 19, 2003

Certified by... ..
Michael Kenstowicz
Professor of Linguistics
Thesis Supervisor

Accepted by .....
Alec Marantz
Head, Department of Linguistics and Philosophy

The Phonetics and Phonology of Tone Mapping in a Constraint-Based Approach

by

Zhiqiang Li

Abstract

This dissertation concerns both phonetic and phonological aspects of tone mapping in various Chinese languages. The central issue addressed is the role of contrast and positional prominence and neutralization in the realization of tone. The inventory of tonal contrasts constrains the outputs of contextual neutralization as well as the location of pitch targets in phonetic implementation. Two prominent phonological positions in the tone sandhi domain are distinguished: peripheral (initial and final) positions and metrically strong positions. Input tones occupying different prominent positions in the input are preserved in the output; their realization in the output can be determined by the location of stress. A typology of diverse patterns of tone preservation and realization emerge from the interaction of positional faithfulness and positional markedness constraints. The research findings reported here have implications for both phonetics and phonoloy.

Acknowledgments

Contents

Chapter 1 Overview

This dissertation concerns phonetic and phonological aspects of tone mapping in various Chinese languages. The central issue addressed is the role of contrast and positional prominence and positional neutralization in the realization of tone. An effort is also made to connect phonetic pitch target timing patterns with phonological patterns of contour tone re-association.

Phonological contrast has been shown to play important roles in constraining phonological neutralization and regulating phonetic contextual variations (Manuel 1990, Flemming 1995, 1997, Padgett 1997, Lubowicz 2003). Various phonological processes involving neutralization or preservation of underlying contrasts are seen as emerging from the interaction of contrast preservation constraints with markedness constraints. Phonetic variations found in vowel-to-vowel co-articulation in various languages are attributable to the language specific inventory of phonological contrasts. This dissertation examines the role of tonal contrast in constraining the outputs of contextual neutralization in the realization of tone as well as the location of pitch targets in phonetic implementation. A model of phonetic implementation is proposed in the Optimality Theory framework. The detailed phonetic timing of the f0 peak in Mandarin rising tone is modeled as emerging from the interaction of phonological constraints on preserving tonal contrasts with phonetic constraints on contour tone durations. In addition, constraints on tonal contrast are also active in allotonic selection in Mandarin Chinese.

Positional prominence and positional asymmetry of phonological processes have received a great deal of attention in the recent phonological literature. As Beckman (1998) points out, certain phonological positions are singled out by the phonology as being privileged. She suggests that prominent positions include stressed syllables, root-initial syllables, long vowels, syllable onsets and possibly final syllables. These positions enjoy either perceptual or phonetic advantage in contrast with other non-prominent positions. There are distinct patterns of phonological asymmetry which can be used as a diagnostic of positional prominence or privilege. For example, prominent positions license phonological contrasts which are neutralized elsewhere; they also resist phonological processes which apply elsewhere (Beckman 1998:1-2).

The question of position prominence and positional neutralization is relevant in the realization of tone. In the context of Chinese tone languages, the prevailing view sees a direct link between tone preservation and stress. It is assumed in most analyses of tone sandhi in Wu dialects, particularly Shanghai, that only a stressed syllable (leftmost in Shanghai) keeps its underlying tone within a tone sandhi domain; other syllables lose their underlying tones (Yip 1980, Duanmu 1993, 1995, Chen 2000 among others). The retained tone, if a contour tone, may undergo a phonological process of contour re-association in which the contour tone spreads over the first two syllables.

The aforementioned single-prominence view (i.e. stress) is inadequate in that it cannot handle cases of tone movement from one position to another. Tone movement involves two phonological positions: a tone occupying a certain position in the input is realized in a different position in the output. The lexical tone sandhi in Zhenhai, as described by Rose (1990), represents such a case. An adequate account of tone

14

movement calls for reference to at least two prominent positions. In the case of Zhenhai, the initial tone is preserved in a lexical compound, and it moves to the stressed syllable, supplanting the underlying tone of the latter. Following a proposal made by Zoll (1997), we distinguish the peripheral positions (initial and final) and a metrically strong position (i.e. stressed position) in a prosodic domain. The interaction of these two kinds of prominent positions gives rise to complex surface tonal patterns in Zhenhai.

In the Optimality Theory framework (Prince and Smolensky 1993, McCarthy and Prince 1995), there has been a debate regarding the role of positional faithfulness and positional markedness constraints. This dissertation argues for admission of both positional faithfulness and positional markedness constraints into the phonology in order to explain the complex interaction of different prominent positions in deriving tone movement and contour tone re-association.

This dissertation provides some new perspectives on positional faithfulness and positional markedness. Beckman's (1998) theory of positional faithfulness is keyed to the idea of positional prominence in phonology. Tone movement in Zhenhai is remarkable in that the surface properties which render the initial syllable prominent (i.e. longer syllable rime duration and initial tone preservation) do not license the realization of the preserved tone. Rather the preserved tone is realized in a position which is phonologically stressed. Contrasting tonal contours, originating in the initial position, are maintained on the stressed syllable. Stressed syllables fail to retain their underlying tones. The altruistic nature of the initial position in Zhenhai calls for a re-appraisal of positional faithfulness. In Zhenhai W-S disyllabic tone sandhi, the most unfaithful mapping takes

15

place: both syllables acquire tonal outputs that are different from the input. The initial tone moves to the second syllable, supplanting the latter's underlying tone.

Contour tone distribution has been argued to be grounded in licensing, or positional markedness (e.g. Zoll 1996, 2003). Zhang (2001) puts forth a careful proposal of contour tone distribution in which phonetic duration (and sonority) of a syllable rime plays a vital role. His theory predicts that a longer syllable is a better licensor of contour tones than a shorter syllable. The contour tone distribution pattern observed in Zhenhai poses a serious challenge to Zhang's phonetically-grounded approach to contour tone distribution. For example, when a phonological short syllable (i.e. checked syllable) occurs in the initial stressed position, the following phonological long syllable (i.e. smooth syllable) still loses its underlying tone.

Contour tone re-association observed in Wu dialects has been an issue of intensive interest. It is taken as evidence for the composite nature of contour tones, i.e. a contour tone consists of a sequence of level tones (Woo 1969, Goldsmith 1976, Edwin 1976). However, two questions remain. (1) Why does contour tone re-association happen in the first place? (2) Why is re-association local (i.e. the second tone segment seeks out the second syllable, but not the third)? Both questions are approached by adopting contour tone licensing constraints which restrict the distribution of contour tones to certain phonological positions. In particular, a parallel will be drawn between phonetic pitch target timing patterns and phonological patterns of contour tone re-association.

The remainder of this dissertation is organized as follows.

Chapter 2 discusses the role of tonal contrast in allotonic selection and phonetic timing of the f0 peak in Mandarin rising tone. It is proposed that constraints on preserving tonal contrast have to be invoked to explain tonal processes in phonology and phonetics.

Chapter 3 presents an analysis of neutral tone in Mandarin Chinese. Special reference is made to its relation to boundary intonation in Mandarin. It is proposed that the low tone target on the last neutral tone syllable is a boundary tone aligned with the right edge of a prosodic domain. The idea of a boundary tone is applied to an account of tone mapping in various Chinese Wu dialects such as Zhenhai and Shanghai in later chapters.

Chapter 4 lays out a theory of multiple prominent positions and their interaction in tone mapping. It is proposed that both stress and edge in a prosodic domain are relevant and their interaction gives rise to diverse patterns of tone-prominence typology.

Chapter 5 presents a formal theory of tone movement, illustrating the necessity of admitting both positional faithfulness and positional markedness constraints in the phonology.

Chapter 6 offers a formal analysis of diverse patterns of contour tone re-association and spreading. It demonstrates the relevance of phonetic alignment patterns of f0 peak and positional effects on contour tone distribution. In particular, the so-called edge-in association will be shown to be an epiphenomenon of contour tone licensing.

Chapter 7 summarizes the dissertation.

Chapter 2 Tonal Contrast in Tone Mapping

2.1   Introduction

There has been a growing body of research on the status of contrast in both phonology and phonetic implementation (Keating 1985, Manuel 1990, Kirchner 1997, Flemming 1995, Flemming 1997, to appear, Padgett 1997, 2001, 2003, Sanders 2002, Lubowicz 2003, among others). Although its importance has never been underestimated in phonology and phonetics, it is only recently that the role of phonological contrast has been vigorously pursued in phonological theories, especially in the Optimality Theory framework (Flemming 1995, Padgett 1997, Sanders 2002, Lubowicz 2003). For example, it is claimed in the dispersion theory of contrast that, with respect to inventories, the selection of phonological contrasts is subject to three functional goals: maximization of contrast, minimization of articulatory effort and maximization of the number of contrasts (Flemming 1995, to appear). These functional considerations are formalized as a family of rankable and violable constraints on contrasts. A somewhat different model of phonological contrast is proposed by Lubowicz (2003) who suggests that contrast preservation exists as an independent principle in the grammar and that preservation or neutralization of underlying contrasts results from the interaction of contrast preservation constraints with markedness constraints. Despite the conceptual differences which do not concern us here, both theories assume that wellformedness must be evaluated with direct reference to the system of contrasts a form enters into.

Phonological contrast has also been claimed to constrain how underlying phonological forms are phonetically realized. It has been shown in the literature that phonetic variations are not random. They are often constrained by the phonological system of the language in question (Keating 1985, Manuel 1990). One factor that has been found to regulate contextual variation is the distinctiveness of phonological contrasts (Flemming 1997). The often cited example is Manuel (1990) in which she found that vowel-to-vowel coarticulation is more limited in more crowded vowel inventories, where neighboring vowel phonemes are less distinct acoustically. Connell (2002) found that tonal contrast controls intrinsic f0 variations in vowels in a number of African tone languages.

In this chapter, we examine the effect of tonal contrast on both phonological and phonetic fronts: phonological allotonic selection and phonetic timing of f0 targets, drawing on evidence from Mandarin Chinese. We will demonstrate that both processes are subject to constraints on preserving tonal contrast. They result from the interaction of contrast preservation constraints and positional markedness constraints on contour tone distribution.

2.2   Allotonic Selection in Mandarin

2.2.1   Tones in Mandarin Chinese

Mandarin Chinese distinguishes four lexical tones on stressed syllables in citation form. They are illustrated in the following familiar paradigm. When a syllable is unstressed, it does not carry a lexical tone, but is said in "neutral tone" (Chao 1968). A neutral tone

syllable is generally perceived as short and lax. More details about Mandarin neutral tone will be presented in section 2.3.2 and also in chapter 3. Tones are transcribed using Chao's tone digits on a five-point scale (Chao 1930), on which '1' indicates the lowest pitch and '5' the highest pitch. The beginning and ending pitch is marked. The turning point is also marked if it is a complex contour tone (i.e. convex or concave). Following the standard convention of pinyin, four tone diacritics (¯ ´ ˘ `) appear on top of the syllables.

(1)   Four tones in Mandarin Chinese

tone	example	pitch value	gloss
1	fēi	55	'to fly'
2	féi	35	'obese'
3	fěi	214	'bandit'
4	fèi	51	'to waste'

In terms of contour shape, tone 1 is high level, tone 2 is high rising, tone 3 is low dipping, and tone 4 is high falling. The figure below shows the average f0 curve of the syllable /ma/ said in four tones, obtained by averaging over four tokens produced by a male speaker recruited for our experiment, to be reported in section 2.3, and plotted as a function of absolute time. The boundary between /m/ and /a/ is marked by '+'. Tone 1 remains at a high f0 value throughout the syllable. The rising f0 in tone 2 begins one third of the way into the syllable. Tone 3 starts at a mid f0. It falls to the lowest value of the four tones right in the middle of the syllable, then rises sharply to the end. Note that there is a striking resemblance between the rising f0 in tone 2 and tone 3, differing only in their relative location: higher for tone 2, and lower for tone 3. Tone 4 starts with a high f0 value. It hits the highest f0 value among the four tones at the beginning of the

vowel, and then falls to the end. These f0 patterns match those reported in Xu (1997). They reflect the phonetic shapes of tones produced in isolation.

In terms of duration, tone 4 is the shortest while tone 3 is the longest, with tone 1 and tone 2 in between. These patterns agree with those reported in Lin (1988) and Xu (1997). The relevance of these duration patterns to tonal distribution will be discussed in section 2.2.4.

(2)   Mean f0 curves of the four tones in Mandarin Chinese
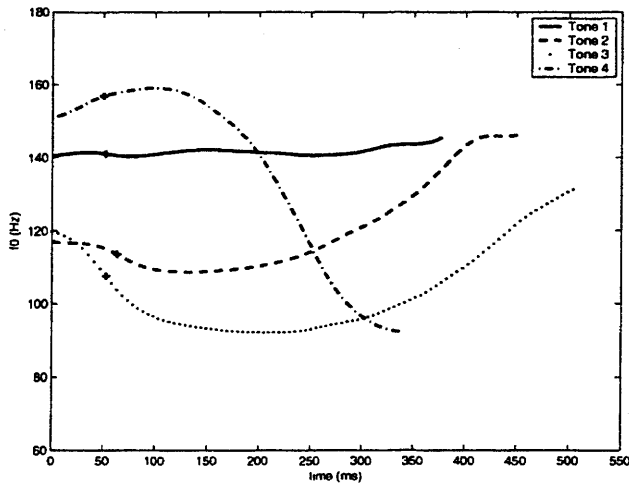
Figure 1: Mean f0 curves of the four tones in /ma/ read in isolation by a male native speaker.

2.2.2   Allotonic Variants of Tones

Some of the four tones have well-known allotones, depending on tonal and prosodic contexts:[1]

---

[1]   We do not consider the contextual variants due to assimilatory tonal coarticulation. For example, when tone 1 is preceded by a tone with low f0 ending, there is a period of rising f0 before the high-level pitch target of tone 1 is reached (Xu 1993, 1997). We will return to this issue in section 2.3.

22

(3)   Allotonic forms of four tones

tone	citation form	allotonic forms
1	55	55
2	35	35, 55
3	214	214, 35, 21
4	51	53

Tone 2 is high rising in citation form, but becomes 55 in a specific tonal context. Chao

(1968:27-8) describes this as "a tone sandhi of minor importance": in a three-syllable

group (words or phrases), tone 2 in the middle changes to tone 1 when the first syllable is

tone 1 or tone 2 (i.e. both have high f0 ending), and the third syllable is in any of the four

tones (i.e. cannot be in neutral tone). This tonal process applies only at a conversational

(i.e. relatively fast) speed. Two examples are given below. Tones on different syllables

are separated by a hyphen. Target tones by this process are underlined.

(4)   a.   H-<u>LH</u>-T → H-<u>H</u>-T, LH-<u>LH</u>-T → LH-<u>H</u>-T   (T = any lexical tone)

   b. cong you bing   "onion oil cake"
       55 - <u>35</u> - 214   citation form
       55 - <u>55</u> - 214   surface form

   c. shei neng fei   "Who can fly?"
       35 - <u>35</u> - 55
       35 - <u>55</u> - 55

The first syllable is in tone 1 (b) and tone 2 (c) in the above examples. The third syllable

can be in any tone. The beginning of the rise on the second syllable assimilates to the

ending of the preceding tone in the process.

It is proposed that this process is conditioned by weak stress on the second

syllable in a trisyllabic word (Yip 1980, Zhang 1988, Shih and Sproat 1992). In

Mandarin, the second syllable has the least stress in a trisyllabic word. According to

Chao (1968:38), in a string of full-toned syllables, the last has most stress, the first has slightly weaker stress, and the intermediate ones have the least stress. Yip (1980:291) provides the following rule. She suggests that this is clearly a rule of tonal assimilation, which applies only when the tones are dominated by a weak node of the stress tree (i.e. non-finally).[2]

(5)

$$L \rightarrow H / H \underline{\quad\quad} H$$

with a w node dominating the blank and the final H.

Apart from the arbitrariness of the rule in (5), it suffers from other problems. As Duanmu (2000:223) points out, it is unclear why assimilation does not occur in other trisyllabic combinations, such as HL-H-LH (tone 4 - tone 1 - tone 2), which contains two HLH sequences. In addition, why does rising tone, but not falling tone, undergo this process? The comparable context would be HL-HL-HL (three tone 4's).

We outline a theory which draws on Zhang's (2001) insight on contour tone licensing. First, the two conditions (least stress and fast speech rate) combine to lead to the conclusion that the rising tone is being modified because of duration reduction. Since the target syllable is the middle syllable in a three-syllable group, it has the least stress, which correlates with shorter duration. The duration is further reduced at a relatively fast speech rate. Second, in terms of the relation between duration and contour tone distribution, a rising tone is more marked than falling tone and level tones because it requires a more stringent condition on duration. Sundberg (1979) and Xu and Sun (2002) document that it takes longer time to implement a pitch rise than a pitch fall with the

---

[2]   In her theory, stress is final at the phrasal level (Yip 1980:144).

same pitch excursion. This can be seen in figure 1 as well, where the rising tone and the dipping tone are much longer than the falling tone.

The tonal process involving the rising tone in Mandarin shows that the more marked rising tone is neutralized to high level (as a result of assimilation to the preceding tone) in the context where the host syllable is not long enough to accommodate the rise. High and fall are not targeted by this process because they are less marked than a rise.

We have pointed out that LH is more marked than H and HL. However, tone 3 is the most marked tone in terms of duration because it is the longest. Tone 3 has three variants: 214 in final position and citation, 35 before another tone 3, and 21 before any other tones. It is also 21 before neutral tone (to be discussed in chapter 3). Although these allophonic processes are well described, we should ask why they take place in the first place. Apparently they are distinct from the assimilatory process we first discussed. Specifically, the following questions can be raised:

(6) a. Why does 214 only appear in final position and citation?

     b. Why does 214 become 35 before another 214, but not elsewhere?

     c. What determines the choice between 21, 14 and 44 before other tones?[3]

Before we leave this preliminary section, we briefly describe the variants of tone 4. Tone 4 alternates between 51 and 53. Chao (1968:28-29) points out that when tone 4 is followed by another tone 4, the sequence becomes 53-51, rather than 51-51. In fact, tone 4 is realized as 53 whenever it is in non-final position (i.e. followed by another tone). 51 may thus be due to the combined effect of final lowering and final lengthening (cf.

---

[3] We assume that in principle 214 can be simplified as 21 (L), 14 (R), or 44 (H), corresponding to the first half, second half and ending point of 214.

25

Liberman and Pierrehumbert 1984): the falling tone falls to 51 when there is more time as in final position.

2.2.3 Third Tone Sandhi

The well-known third tone sandhi applies to two syllables in tone 3 and changes the first one into tone 2. The process is stated below with two examples:

(7)   a. tone 3 → tone 2 / _____ tone 3; alternatively 214 → 35 / _____ 214

     b. 'good wine'   'buy horse'
        hao jiu   mai ma
        214 214   214 214   citation tone
        <u>35</u> 214   <u>35</u> 214   sandhi tone

The third tone sandhi applies within a domain. For a review of the literature on the formation of tone sandhi domain, the reader is referred to Duanmu (2000, chapter 11) and Chen (2000, chapter 9).

Xu (1993) found that there is a slight separation of the f0 curves for the derived tone 2 (from tone 3) and the underlying tone 2.[4] However, the perception study of Wang and Li (1967) shows that, as far as the listener is concerned, tone 2 derived from tone 3 is indistinguishable from the underlying tone 2.

There have been attempts to uncover the phonological and perceptual motivations for the third tone sandhi. Cheng (1973) and Yip (1980) consider it a case of dissimilation, which disallows two identical tones in succession. For example, Yip (2002), in continuation with Yip (1980), takes L as the underlying form of tone 3, and attributes the

---

[4]   The derived tone 2 has lower f0 than the underlying tone 2. The difference is smaller than 10 Hz, as estimated from the f0 graphs reported in Xu (1997:69).

third tone sandhi to OCP: L.L → LH.L (p.180). In contrast, Milliken (1989) takes the citation form as basic, and proposes that tone 3 is underlying L followed by a floating H, i.e. L(H). The floating H associates with its host syllable to derive the third tone sandhi: L(H).L(H) → LH.L(H). In a more phonetic approach, Huang (2001) shows that perceptually tone 3 and tone 2 are most confusable because of their similar phonetic shapes.

There is also a historical basis for the third tone sandhi in Mandarin Chinese. There are four tonal categories in Middle Chinese (approximately from AD 200 to 900), as reflected in the pronouncing dictionary *Qieyun* (AD 601, compiled by Lu Fayan). They are referred to by their traditional nomenclature as *ping* "level", *shang* "rising", *qu* "departing" and *ru* "entering". Each of the four MC tonal categories is split into *yin* and *yang* registers, conditioned by the voicing contrast in the onset. The *yin* register with a voiceless onset has a higher pitch value than the corresponding *yang* register with a voiced onset. Although the correlation between tone register and voiced/voiceless contrast in the syllable onset is still transparent in Wu dialects (like Shanghai), it is not always reflected in the synchronic systems in other dialects. The correspondence between the four tones in Mandarin Chinese and the MC tonal categories is depicted below (see Chen 1976 for details):

(8)

register	tone			
	I *ping* "level"	II *shang* "rising"	III *qu* "departing"	IV *ru* "entering"
a. *yin*	tone 1 (55)	tone 3 (214)	tone 4 (51)	tone 1, 2, 3, 4
b. *yang*	tone 2 (35)			

The tonal category IV is the so-called "checked tone", occurring on syllables with consonantal endings. It completely disappears in Mandarin Chinese, and splits irregularly into the four tones. In historical terms, the third tone sandhi is often described as "*shang (II)* + *shang (II)* → *yang ping (Ib)* + *shang (II)*". According to Mei (1977), this process was observed as early as the sixteenth century in northern Mandarin dialects. It is still observed in many northern dialects despite the fact that the exact tonal values for the two synchronic correspondents of MC tonal categories *shang* and *yang ping* vary from dialect to dialect. In the following table, we compare the four tones of three Mandarin dialects with reference to the MC tonal categories. The third tone sandhi applies in all of them.

(9)   Tonal systems in Beijing (Mandarin Chinese), Tianjin and Jinan

MC tonal categories	*yin ping*	*yang ping*	*shang*	*qu*
Beijing	55	35	214	51
Tianjin[5]	21	45	213	53
Jinan[6]	213	42	55	21

The list can be made longer by including more Mandarin dialects. The upshot is that the third tone sandhi observed in Mandarin Chinese is the residue of a historical change that has synchronic ramifications in many northern dialects. That is why this tonal process is sensitive to the MC tonal categories corresponding to the tones in a synchronic tonal system. Since we have nothing new to contribute to the current understanding of the third tone sandhi, we simply refer to this process as a descriptive constraint THIRDTONESANDHI (cf. Duanmu 1999:29). Returning to Mandarin Chinese, the change

---

[5]   Li and Liu (1985)
[6]   Z. Qian (1995)

from 214 to 35 before another 214 will be considered as resulting from this constraint regardless of whether it is a process of dissimilation or a more arbitrary tonal substitution.

2.2.4   Positional Asymmetry in Tonal Distribution

We have noted earlier that tone 3 is realized as 214, a complex contour tone, only in citation or final position. Elsewhere it is either 35 (compelled by the THIRDTONESANDHI), or 21. The other three tones do not have this positional asymmetry in terms of tonal distribution. They (i.e. 55, 35 and 51) are free to occur in both final and non-final positions. Zhang (2001) suggested that the final-lengthening effect, often observed in final stressed syllables, is responsible for the restriction of 214 to final position. His insights are formalized as the following constraints:

(10)   a. *COMPLEXCONTOUR/NON-FINAL σ́: a complex contour tone is not allowed on a

non-final stressed syllable.

b. *COMPLEXCONTOUR/FINAL σ́: a complex contour tone is not allowed on a final

stressed syllable.

c. *COMPLEXCONTOUR/UNSTRESSED σ: a complex contour tone is not allowed on

an unstressed syllable.

The complex contour tone is defined as having two inflection points and also significantly longer phonetic duration than simple contour tones in citation form. In Mandarin 214 is considered a complex contour tone whereas 35 is a simple contour tone. Since unstressed syllables have shorter duration than non-final stressed syllables, which in turn have shorter duration than final stressed syllables (due to the final-lengthening

effect), the following ranking of the three constraints obtains in light of the durational requirement for the realization of a complex contour tone: *COMPLEXCONTOUR/UNSTRESSED σ >> *COMPLEXCONTOUR/NON-FINAL σ́ >> *COMPLEXCONTOUR/FINAL σ́. These constraints, in interaction with faithfulness constraints, restrict the complex contour tone to a specific position, phrase-final in this case.

(11) *COMPLEXCONTOUR/NON-FINAL σ́ >> FAITH(T) >> COMPLEXCONTOUR/FINAL σ́

A. tone 3 in non-final position

214-35	*COMPLEXCONTOUR/ NON-FINAL σ́	FAITH(TONE)	COMPLEXCONTOUR/ FINAL σ́
a. 214-35	*!		
b. ☞ 21-35		*	

B. tone 3 in final position

35-214	*COMPLEXCONTOUR/ NON-FINAL σ́	FAITH(TONE)	COMPLEXCONTOUR/ FINAL σ́
a. 35-21		*!	
b. ☞ 35-214			*

The positional asymmetry in the distribution of tone 3 is accounted for by ranking the tonal faithfulness constraint FAITH(TONE) ("do not change a tone") lower than the positional markedness constraint referring to non-final stressed syllables. It has to be ranked higher than *COMPLEXCONTOUR/FINAL σ́ so that the complex contour tone 214 is allowed on final stressed syllables. In the first tableau (A), candidate (a) fatally violates the high-ranking constraint prohibiting a complex contour on non-final stressed syllables. (b) does not violate *COMPLEXCONTOUR/NON-FINAL σ́, so the unfaithful mapping is the

output. In the second tableau (B), the faithful mapping wins because the complex contour occurring in the final position only violates the low-ranking COMPLEXCONTOUR/FINAL σ́.

However, when we consider a larger candidate set, indeterminacy appears given the ranking established above.

(12) tone 3 in non-final position

214-35	*COMPLEXCONTOUR/ NON-FINAL σ́	FAITH(TONE)
a. 214-35	*!	
b. ☞ 21-35		*
c. ☞ 14-35		*
d. ☞ 44-35		*

Candidate (b) is the actual output. The unattested mappings from the input to the candidates in (c) and (d) also go through in the above tableau because none of them violates the tonal markedness constraint and all of them violate the faithfulness constraint. It deserves mentioning that FAITH(TONE) is used as a cover term for MAX(TONE) ("do not delete a tone") and DEP(TONE) ("do not insert a tone"). One may contend that when the real faithfulness constraints are employed the candidates in (b), (c) and (d) could possibly be distinguished. We work it out in the following tableau. Since there is no tone insertion, we only consider MAX(TONE). Assuming that the mora is the tone bearing unit in Mandarin (Duanmu 1990), tone 3 can be represented as being associated with three moras, the first two linked with a L and the third one with a H. Tone 2 is associated with two moras, linked with L and H respectively. The representational difference captures the fact that tone 3 is a complex contour tone with longer duration and its final rise is

much later in the syllable. All candidates in (12) are translated into the representational forms.

(13)

σ σ /|\ /\ m m m m m \/ | | | L H L H	*COMPLEXCONTOUR/ NON-FINAL σ	MAX(TONE)
a. m m m $ m m \/ | | | L H L H	*!	ʼ
b. ☞ m m $ m m \/ | | L L H		*!
c. ☜ m m $ m m | | | | L H L H		
d. m m $ m m \/ | | H L H		*!

($ = syllable boundary)

Candidate (a) is faithful to the input, but violates the high-ranking constraint. The three remaining candidates all delete one association line from the input. (c) does not delete any tone from the input, so it satisfies MAX(TONE). (b) and (d) each deletes one tone from the input. Therefore, (c) comes out as the output instead of (b). Further decomposing MAX(TONE) into MAX(H) and MAX(L) would not help because (c) does not violate MAX(TONE) at all.

Several fixes can be entertained. First, (c) may be ruled out by invoking a constraint banning rise in non-final position. This is not a viable option because such a constraint would incorrectly rule out 35 in non-final position. One can distinguish high rise (35) and low rise (14), and have a tonal markedness constraint that rules out the low rise, but allows the high rise. Such a move seems ad hoc and unmotivated given that

there is no other known tonal processes that refer to the register distinction between high

rise and low rise in Mandarin Chinese. By the same token, (d) cannot be ruled out by

invoking a tonal markedness constraint (like *H/NON-FINAL ό) because it will incorrectly

rule out tone 1 (55) in non-final position. Second, one can invoke the notion of "nuclear

tone", introduced by Yip (2002). She suggests that within the syllable the first (or only)

tone is the nuclear or head tone (Yip 2002:176). If we consider 21 as the head tone of

214, then FAITHNUCLEARTONE will pick (b) over (c) and (d).

(14)   *COMPLEXCONTOUR/NON-FINAL ό, FAITHNUCLEARTONE

	214-35	*COMPLEXCONTOUR/ NON-FINAL ό	FAITHNUCLEARTONE
a.	214-35	*!	
b. ☞	21-35		
c.	14-35		*!
d.	44-35		*!

In the next section, we consider an alternative analysis, which takes into

consideration the system of tonal contrast in Mandarin Chinese.

2.2.5   Tonal Contrast

We have shown that the four lexical tones in Mandarin Chinese are distinguishable by

contrasting contour shapes in citation form. Also, tone 3 is neutralized with tone 2 in a

specific tonal context, as triggered by the THIRDTONESANDHI. In terms of tonal contrast,

neutralization occurs when it is compelled by tonal markedness constraints (e.g.

THIRDTONESANDHI). Other than that, tonal contrast has to be maintained. Following

Flemming (1995) and Padgett (2003), we define a family of constraints that have the

effect of maintaining tonal contrasts in the tone system, defined below:

(15) PRESERVECONTRAST(T_i/T_j): T_i and T_j have to be perceptually distinct.

Constraints in this family monitor tonal contrasts in the system, and penalize tone mappings which will lead to neutralization. More about this constraint family will be presented in section 2.3.4.

Now we consider the mappings of tone 3 again:

(16) A. 214 → 35, neutralizing, compelled by THIRDTONESANDHI

B. 214 → 21, non-neutralizing, compelled by *COMPLEXCONTOUR/NON-FINAL σ

and PRESERVECONTRAST(T_i/T_j)

	214-35[7]	tonal contrast neutralized?
a.	214-35	no, T1: 55, T2: 35, T3: 214, T3: 51
b. ☞	21-35	no, T1: 55, T2: 35, T3: 21, T3: 51
c.	14-35	yes, T1: 55, T2: 35, T3: 14, T4: 51
d.	44-35	yes, T1: 55, T2:35, T3: 44, T4: 51

When tone 3 is followed by another tone 3, 214 is mapped to 35, which is the tonal value of tone 2. The mapping is a neutralization process, compelled by THIRDTONESANDHI (16A). When tone 3 is followed by any other tone, 214 becomes 21 (15B). In the above table (16B), we consider various mappings of 214 in non-final position and evaluate them for their effect on the system of tonal contrast. Recall that in citation form the four lexical tones are clearly distinguished by contrasting tonal contours. In (a), tone 3 is mapped to 214, which is its citation form. After this mapping has taken place, the system of tonal contrast (labeled as T1 through T4 in the table) is maintained; there is no neutralization of tonal contrast. In (b), tone 3 is mapped to 21. Although it is not faithful to the input tone value, the mapping does not perturb the system of tonal contrast since 21

---

[7] We do not consider candidates like 41 which involve metathesis of individual tonal features. It will be ruled out by the undominated LINEARITY constraint (McCarthy and Prince 1995).

34

is still distinct from the other three tones in the system. In (d), if we treat 44 and 55 as phonologically identical, the mapping from 214 to 44 will induce neutralization between tone 3 and tone 1 (shaded region). The mapping in (c) deserves some comments. It gives rise to 14 for tone 3, which is confusable with tone 2 (35). The question is whether 14 and 35 neutralize. Conceivably, there is a register difference between 14 (low rise) and 35 (high rise). We argue that even if 14 and 35 are not phonetically neutralizing, they are perceptually confusable. This is because Mandarin Chinese and most northern dialects do not use register as a distinctive tonal feature. They are different from Wu dialects, like Shanghai, in which a phonological distinction is made between a high rising tone and a low rising tone. In Mandarin Chinese, tonal register is modified as a result of stress or focus. Xu (1999) shows that when a syllable is focused, its tonal register is shifted up and also expanded. Therefore, we conclude that the mapping from 214 to 14 violates the constraint on contrast preservation.

In order to explain allotonic selection of tone 3, we offer the following ranking arguments (16-xx) for the constraints on tonal contrast.

(17) THIRDTONESANDHI >> PRESCONTRAST(T_i/T_j)

	214-214	THIRDTONESANDHI	PRESCONTRAST(T_i/T_j)
a.	214-214	*!	
b.	21-214	*!	
c. ☞	35-214		*, tone 3 / tone 2

The neutralizing third tone sandhi motivates the ranking of THIRDTONESANDHI over the constraint on preserving tonal contrast. We simply use THIRDTONESANDHI as a descriptive constraint. It specifies both structural description and structural change. It

requires that the sequence of 214-214 be changed to 35-214. In the tableau, the mapping from 214 to 35 in (c) neutralizes the contrast between tone 2 and tone 3.

(18) PRESCONTRAST(T_i/T_j), *COMPLEXCONTOUR/NON-FINAL σ́

214-35		PRESCONTRAST(T_i/T_j)	*COMPLEXCONTOUR/ NON-FINAL σ́
a.	214-35		*!
b. ☞	21-35		
c.	14-35	*!, tone 3 / tone 2	
d.	44-35	*!, tone 3 / tone 1	

In this tableau, PRESCONTRAST(T_i/T_j) excludes (c) and (d) because both mappings give rise to neutralization of tonal contrast, as we have discussed in (16). (c) violates the tonal markedness constraint for having a complex contour in non-final position. (b) is the winner. Since the winner does not violate either constraint, the two constraints in (18) are not crucially ranked. We have established that *COMPLEXCONTOUR/NON-FINAL σ́ outranks FAITH(TONE) in (12). There does not seem to be a crucial ranking between PRESCONTRAST(T_i/T_j) and FAITH(TONE).

The forgoing discussion shows that the constraint on contrast preservation has to be included in the phonology. In the case of Mandarin Chinese, it correctly predicts which allophonic form should be selected for tone 3. Its role in the phonology cannot be subsumed by either tonal faithfulness constraints or contour tone licensing constraints.

2.2.6   Summary

To summarize, we have shown that constraints on contrast preservation determine the selection of allotones of tone 3 when they are interacting with constraints on contour tone

distribution. In the next section, we provide further evidence for constraints on tonal contrast by turning to the phonetic timing of f0 targets in Mandarin Chinese.

2.3   Phonetic Timing of F0 Target[8]

2.3.1   Motivation of the Study

The impetus for the investigation to be reported in this section rests on two broad issues in the phonology-phonetics interface: one theoretical and the other experimental. On the theoretical side, a central question is how the abstract phonological structure is physically realized in real time and space. In this connection, a distinction is often made between phonological processes, which are dealt with in the phonological component of the grammar, and phonetic processes, which are dealt with in the phonetic implementation component (Keating 1996, Cohn 1990). It has been observed that mapping from the output of phonology to observable surface phonetic events is not straightforward. Rather the phonetic processes exhibit both contextual and cross-linguistic variations. As we discussed earlier, the distinctiveness of phonological contrasts has been shown to regulate contextual variation (Flemming 1997). In this section, we demonstrate a case in which constraints on tonal contrast operates to restrict the placement of f0 targets.

On the experimental side, a number of recent studies conclude that the precise temporal coordination of f0 events with phonetic segments is quite complex and that tonal alignment does not follow directly from phonological association. In phonology, the association of tone or pitch accent with specific elements of the segmental string (i.e.

---

[8]   This section is based on Li (2002) with modifications.

tone-bearing unit) is specified by autosegmental links: Chinese tones are associated with moraic segments (or syllable rimes), English pitch accents are associated with stressed syllables, and Japanese word accents are associated with a specific mora.

Such an association relation is often obscured in the phonetic implementation process where tonal targets are realized as f0 events. A cross-linguistically common pattern is that the f0 peak of a tone or a pitch accent may appear after the tone-bearing unit (mora, rime or syllable) with which it is phonologically associated. This is sometimes called "peak delay". For example, Silverman and Pierrehumbert (1990) found that in English the f0 peak in prenuclear H* pitch accent often occurs after the accented syllable that bears the pitch accent. Prieto, van Santen and Hirschberg (1995) reported that in Mexican Spanish the f0 peak in H* accent is delayed unless in the presence of a following accented syllable. In their study of rising prenuclear accent in Modern Greek, Arvaniti, Ladd and Mennen (1998) found that the H target is consistently aligned just after the beginning of the first post-accentual vowel. Xu (2001) presented similar peak delay phenomena for the Mandarin rising tone whose f0 peak mostly occurs just inside the following onset consonant in the context of a following low tone. The situation is further complicated by speech rate and the prosodic position of the tonal target in question (Steele 1986, Silverman and Pierrehumbert 1990). In addition, there seem to be consistent, but also fine-grained cross-linguistic variations in tonal target alignment. In a number of recent papers by Ladd and his colleagues (*Modern Greek*: Arvaniti, Ladd and Mennen 1998, *British English*: Ladd, Faulkner, Faulkner and Schepman 1999, *Dutch*: Ladd, Mennen and Schepman 2000, *German*: Atterer and Ladd 2002), they report consistent alignment patterns of tonal targets in the languages they studied; those

languages exhibit small but significant alignment differences from one another. Specifically, although the low target is aligned with the onset of the stressed syllable in both British English and Modern Greek prenuclear rises, the high target is aligned earlier in the former case, typically late in the immediately following consonant. In German, the low target of the prenuclear rise is "aligned well within the initial consonant of the stressed syllable or even early in the stressed vowel" (Atterer and Ladd 2002), and Northern speakers align the low target earlier than Southern speakers.

The immediate precursor of our experimental study is Xu's extensive studies of tonal alignment with segmental strings in Mandarin Chinese (Xu 1998, 1999, 2001), in particular his study of f0 peak delay. He found that regardless of internal syllable structure, f0 curves for all four tones maintain a consistent alignment to their host syllables. He also found that when the tonal context is appropriate (e.g. L__L), f0 peaks appear in tone 1 (H), tone 2 (LH) and tone 4 (HL), and they exhibit distinct patterns of alignment with segmental strings. Specifically, the f0 peak in a rising tone (R) occurs late in the syllable, and mostly in the following onset consonant; the f0 peak in a falling tone (F) occurs early in the vowel or near the middle of the syllable; and the f0 peak in a high tone (if the peak is observable) occurs close to but before the end of the syllable. The f0 peak is delayed regularly in R regardless of speech rate, but it is only delayed half of the time in H at fast speech rate, with almost no delay at all at normal and slow speech rate. To account for variations in the occurrence of peak delay in Mandarin, Xu and colleagues proposed an interactive account which attributes peak delay as resulting from the interaction of the underlying pitch targets and their articulatory implementation that determines peak alignment (Xu 2001, Xu and Wang 2001, Xu and Sun 2002). According

39

to this theory, tonal target in **H** is [high], thus f0 starts to rise right at the syllable beginning (e.g. in L.**H**). In contrast, the target in **R** is [rise] with a low starting f0, thus f0 has to go down first (e.g. in H.**R**). Peak delay occurs regularly in R because the beginning of rising f0 occurs about in the middle of syllable rime (i.e. late beginning of rise), and the f0 peak has to run over into the following segment to complete the rise before L. Peak delay occurs rarely in H because the beginning of rising f0 transition occurs right at the syllable boundary (i.e. early beginning of rise). As a result, there is enough time to finish the rising f0 transition before the end of the syllable.

Our experimental study is motivated by an empirical difference between the f0 peak alignment in Modern Greek rising accent and Mandarin rising tone. As reported in Xu (2001), the f0 peak in Mandarin rising tone mostly occurs just inside the following onset segment. In Greek, however, the f0 peak in a rising accent is aligned on average 15-20 ms after the beginning of the following unstressed vowel (Arvaniti, Ladd and Mennen 1998). In addition, the onset of the rise is also aligned differently: in Greek, it occurs right before the onset of the syllable whereas in Mandarin, it occurs in the middle of the syllable rime. The difference could possibly arise from "tonal crowding" or "stress clash" in Mandarin because in Xu's studies the rising tone was followed by a stressed syllable carrying a low tone (i.e. tone 3) while in Arvaniti et al.'s speech materials there were always at least two unstressed syllables on either side of the test stressed syllable carrying the rising accent. Therefore, the first goal of this section is to understand what happens to the f0 peak in Mandarin rising tone when its host syllable is followed by unstressed syllables. If the early peak in Mandarin (i.e. in the following consonant compared to at the beginning of the following vowel in Modern Greek) were due to the

presence of the following stressed syllable, then we would expect a later f0 peak in R when the following syllable is unstressed. The second goal is to provide an Optimality-Theoretic analysis (Prince and Smolensky 1993, McCarthy and Prince 1995) of the positional asymmetries in the f0 peak alignment in Mandarin. Based on our experimental findings, we will show that constraints on contrast preservation play an important role in regulating the placement of f0 peaks.

2.3.2 Language Background and Experimental Method

The experiment was designed to examine the effect of the following metrical and tonal contexts on f0 peak alignment. The main factor to manipulate is stress vs. non-stress in the following syllable(s). Before we present the speech data, a note on stress and its relation to tone in Mandarin is in order since the experimental design employs the distinction and its tonal reflections.

Mandarin distinguishes four lexical tones on stressed syllables. A stressed syllable carries one of the four tones. In contrast, when a syllable is unstressed, it is in neutral tone or toneless. Some syllables, like suffixes and sentential particles, are always unstressed except in citation form (where tone 1 will be used). When stressed syllables become unstressed in certain lexical or prosodic contexts, they lose their underlying tones. Unstressed syllables are much shorter in duration than stressed syllables (Lin and Yan 1980). Unlike full-toned stressed syllables which can occur in any position of a word, unstressed syllables only occur in restricted positions. They are usually cliticized to the preceding stressed syllable. Their phonetic pitch varies with the preceding tone (Chao 1968). See chapter 3 for further discussion of neutral tone.

41

(19) Phonetic pitch of neutral tone

context	pitch	example	gloss
after tone 1 (55)	2	tā de	'his'
after tone 2 (35)	3	huáng de	'yellow one'
after tone 3 (214)	4	nǐ de	'yours'
after tone 4 (51)	1	dà de	'big one'

For present purposes, neutral tone syllables are H after tone 3, and L after the other three tones. In other words, they carry a short H pitch target after tone 3, and a short L target after any other tone. They are called short H or L because of the shorter duration of their host syllables. In a string of neutral tone syllables, only the last one gets a short L target. The others do not have pitch targets except the one after tone 3, which gets a short H target. An example is shown in (20) where a syllable in tone 1 (H) is followed by one, two and three neutral tone syllables. A neutral tone syllable is indicated as N.

(20) a. σ N   b. σ N N   c. σ N N N
         | | | | | |
         H L   H   L   H   L

The speech data were constructed in such a way that the four target tones (H, R, short H and F) are followed by one, two or three unstressed syllables (i.e. neutral tone syllables), which are either in utterance-medial or -final positions. A short H occurs in an unstressed syllable after tone 3. All three target tones are preceded by a syllable either in L or F to ensure that f0 peak can be clearly identified. In order to compare with peak delay in R before L, data were also recorded with R followed by H, L and F. Sentences recorded are given below. They are classified according to tonal context and metrical condition, and are transcribed in pinyin. The syllables carrying target tones all have CV

structure and they appear in comparable prosodic positions, initial in a prosodic word, in order to minimize the effect of prosody. Target words (which consist of target tones plus the following context) are underlined, so they will be read with focus.

(21) Sentences Recorded in the Experiment

(a) LRL Fǔyīn de chángdù tōngcháng yǐ <u>háomiǎo</u> jìsuàn. Yīshēng jiànyì <u>niúnǎi</u> měitiān dōuyào hē. Tā bù tài rènde <u>luómǎ</u> shùzì.	The duration of consonants is usually measured in ms. Doctor suggests that one drink milk daily. He does not quite know Roman digits.
(b) LRN/NN/NNN Huángróng jiào <u>máomao/men</u> lái bāngmáng. Huángróng jiào <u>máomao/men</u> mài zìhuà. Guōjìng zhǎodào <u>máomaole</u>. Guōjìng zhǎodào <u>máomaomenle</u> cái húilái. Guōjìng zhǎodào <u>máomaomenle jiù</u> húilái Guōjìng zhǎodào <u>máomaomenle</u>.	Huangrong asks Maomao/'s for help. Huangrong asks Maomao/'s to sell paintings. Guojing has found Maomao. Guojing returned after having found Maomao's. Guojing will return after having found Maomao's. Guojing has found Maomao's.
(c) LHN/NN/NNN Huángróng jiào <u>māma/men</u> lái bāngmáng. Huángróng jiào <u>māma/men</u> mài zìhuà. Guōjìng zhǎodào <u>māmale</u>. Guōjìng zhǎodào <u>māmamenle</u> cái lúilái. Guōjìng zhǎodào <u>māmamenle jiù</u> lúilái. Guōjìng zhǎodào <u>māmamenle</u>.	Huangrong asks mother/'s for help. Huangrong asks mother/'s to sell paintings. Guojing has found mother. Guojing returned after having found mothers. Guojing will return after having found mothers. Guojing has found mothers.
(d) Lh/N/NN (h = short H) Huángróng jiào <u>nǎinai/men</u> lái bāngmáng. Huángróng jiào <u>nǎinai/men</u> mài zìhuà. Guōjìng zhǎodào <u>nǎinaile</u>. Guōjìng zhǎodào <u>nǎinaimenle</u> cái lúilái. Guōjìng zhǎodào <u>nǎinaimenle jiù</u> lúilái. Guōjìng zhǎodào <u>nǎinaimenle</u>.	Huangrong asks granny/'s for help. Huangrong asks granny/'s to sell paintings. Guojing has found granny. Guojing returned after having found grannies. Guojing will return after having found grannies. Guojing has found grannies.
(e) LFN/NN/NNN Huángróng jiào <u>mèimei/men</u> lái bāngmáng. Huángróng jiào <u>mèimei/men</u> mài zìhuà. Guōjìng zhǎodào <u>mèimeile</u>. Guōjìng zhǎodào <u>mèimeimenle</u> cái lúilái. Guōjìng zhǎodào <u>mèimeimenle jiù</u> lúilái. Guōjìng zhǎodào <u>mèimeimenle</u>.	Huangrong asks sister/'s for help. Huangrong asks sister/'s to sell paintings. Guojing has found sister. Guojing returned after having found sisters. Guojing will return after having found sisters. Guojing has found sisters.
(f) LRF/H Tā míngtiān yào qù <u>luómàn</u> shāngshà. Guōjìng gǎndào shǒujiǎo <u>mámù</u> cái tíngxiàlái. Zhèchūxì jiào <u>límáohuàntàizǐ</u>. <u>Máolā</u> shì duì yīsīlán xuézhě de zūnchēng.	He is going to Luoman department store tomorrow. G.J. stopped when he felt numb in his hands and feet. The play is called "limaohuantaizi". "Maola" is an honorable title for Islamic scholars.

43

Three native speakers of Mandarin Chinese, one female and two males, all affiliated with MIT, were recruited to record the data. Recording was conducted in a sound-treated room in the MIT Speech Communication Group. The speech signal was digitized at 16 kHz at real time using the Marsha program running on a PC. The subjects read test sentences in Chinese characters displayed on a computer monitor at normal speech rate. They were asked to put focus on targets words, which are underlined. There are four repetitions for each sentence by each speaker.

The f0 extraction followed the procedures developed in Xu (1999, 2001). The signal was first transferred to the Linux server and then converted into appropriate file format readable by programs in the ESPS/waves+ signal processing software package. For each individual test sentence, the ESPS *epochs* program was used to mrak every pitch period in the target word. The marked signals were then manually edited to correct spurious vocal pulse markings due to pitch halfing or pitch doubling. Segmentation labels were added to mark segment boundaries in the target word. The vocal pulse markings and segmentation labels for each sentence were saved in text files, which were then processed by a set of C programs.[9] The program first converted the duration of pitch periods into f0 values, and then smoothed the resulting f0 curves using a trimming algorithm (Xu 1999) that eliminates f0 outliers that are likely to occur at C-V and V-C boundary.

The acoustic measurements, taken algorithmically using another C program, were based on the smoothed f0 curves. The first measurement is duration of the first two

---

[9]   Thanks to Yi Xu for providing the crucial C functions.

syllables in the target word. The first syllable is the one that bears the target tone. The syllable duration is computed from the temporal locations of segment labels at the beginning and end of the relevant syllable. The second measurement is peak-to-C2, which is the location of f0 peak relative to the end of the syllable bearing the target tone. A positive value indicates that the f0 peak occurs before the end of the syllable and a negative value indicates otherwise. Other measurements include duration and excursion size of rising and falling f0 curves in the target word. The f0 maximum is identified as the end of a rising f0 and at the same time, the beginning of a falling f0. The beginning of a rising f0 may be defined as the f0 minimum right before the rise. However, as noted in Xu (1998), there are cases in which the f0 minima seem too far away from the point at which the f0 curves take a sharp turn upward. Such a point is where the f0 curve reaches its greatest acceleration. Mathematically, the moment of greatest acceleration corresponds to the local maximum in the second derivation of the f0 curve. Another C program was written to compute the first and second derivatives of the f0 curves and then to locate the moment of greatest acceleration in the derivatives. Before the derivatives were taken, the smoothed f0 curves were further smoothed using a triangular window function. It has turned out that the beginning of rising f0 thus obtained was too "conservative", being too late in many cases. Finally it was decided to take the mid point between the f0 minima and moment of greatest acceleration as the beginning of rising f0. Visual inspection of the f0 curves shows that the mid point is closest to where the real rise is. The end of a falling f0 is taken to be the f0 minimum after the f0 peak.

2.3.3   Results and Discussion

2.3.3.1   Syllable Duration

The averaged duration of the first two syllables in the target word is shown in (22).

(22) Averaged durations of the first two syllables in the target word (ms.)
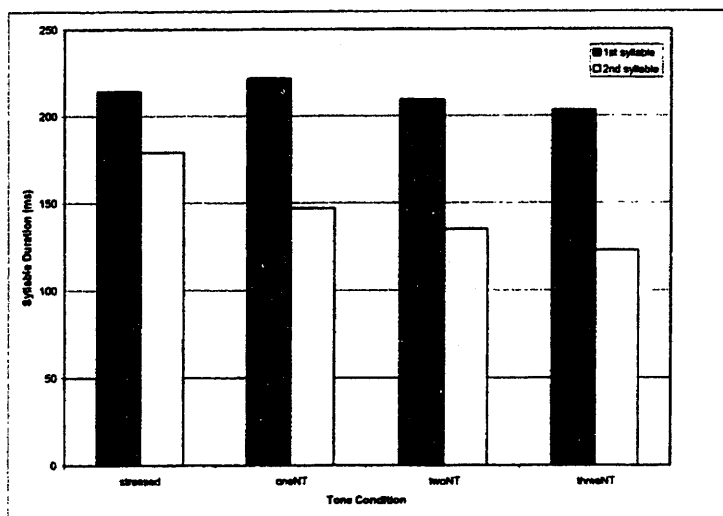
The first syllable (black bar) is always stressed in all four conditions. When the second syllable (white bar) is also stressed ("stressed" condition), the first syllable is longer than the second. When the second syllable is the only neutral tone syllable, its duration is significantly shorter than the preceding stressed syllable ("oneNT" condition). In the case of multiple neutral tone syllables, the second syllable is the first neutral tone syllable. Its duration is further reduced as the number of neutral tone syllables increase ("twoNT" condition and "threeNT" condition). The first syllable duration does not vary much when it is followed by different numbers of neutral tone syllables. The only significant difference is between duration of the first stressed syllable and the second unstressed syllable. The relevant statistics are summarized in the following table.

(23) Statistics of durations of the first two syllables in the target word

	stressed			oneNT			twoNT			threeNT		
	σ_1	σ_2	p	σ_1	σ_2	p	σ_1	σ_2	p	σ_1	σ_2	p
Tone1				213.6	133.5	*	200.0	125.3	*	194.1	114.5	*
Tone2	214.0	179.1	%	225.2	152.2	*	216.8	143.6	*	210.6	131.6	*
Tone3				226.0	155.6	*	211.2	136.0	*	205.1	122.1	*

*p values are based on one-factor ANOVAs. * significant at 0.001; % at 0.05.*

2.3.3.2 Positional Asymmetry in Tonal Target Alignment

The major finding of the experiment is that the alignment of the f0 peak in tone 2 exhibits variations as a function of the following metrical context. So the first goal of this section is to examine the placement of f0 peak in rising tone when its host syllable is followed by neutral tone syllables. To graphically examine tone-syllable alignment, the smoothed f0 curves were averaged across four repetitions using the following procedure developed in Xu (1999). First, the mean duration was computed for each segment over four repetitions. Next, a normalized mean f0 curve for each segment was computed by taking a fixed number of points at equal time intervals and averaging each point over four repetitions. The mean f0 curves of all segments were plotted as a function of time using Matlab. The figure in (23) displays mean f0 curves of four lexical tones as they are followed by three neutral tone syllables, marked by capital letter N1, N2 and N3. H, L, R and F label the relevant portion of f0 curves corresponding to four tones.

(24) Mean f0 curves of four tones followed by three neutral tone syllables (speaker 1)
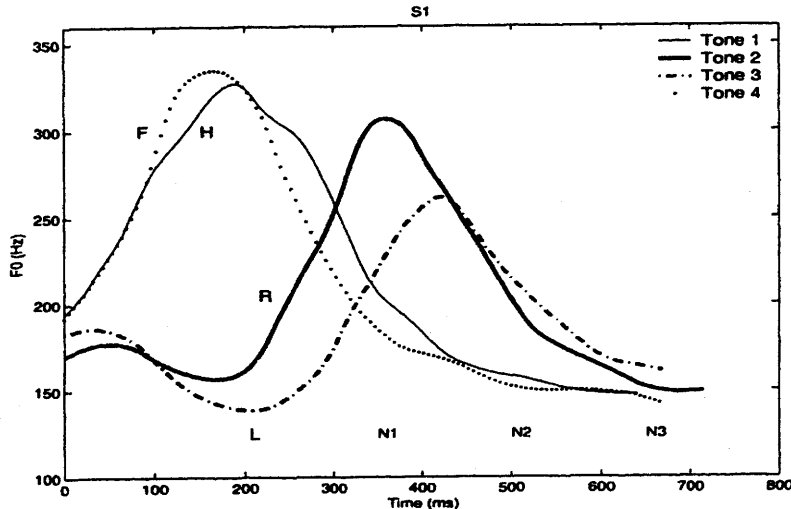
The figure shows that the canonical shapes of the four tones are modified in context. For example, the initial rising f0 in H and F is due to the influence of preceding tonal context. It is observed that f0 curves over the three neutral tone syllables fall almost linearly to the end, indicating the presence of a low pitch target at the end. The neutral tone immediately following tone 3 (L) has a H pitch target. The rising f0 in tone 2 (LH) starts late in the syllable and it has a much delayed f0 peak (in the middle of the following syllable N1) when followed by three neutral tone syllables.

The following figures display mean f0 curves of tone 2 and tone 1 respectively in different contexts from all three speakers: (a) followed by tone 3 (only for tone 2), (b) by one neutral tone, (c) by two neutral tones, (d) by three neutral tones, and (e) by two and three neutral tones combined in utterance-final position. In the figures, the dotted region on each curve corresponds to the sonorant onset of the syllable following the target tones (i.e. the second syllable in the target word). It is used as a segmental landmark against which the f0 peak alignment in tone 2 and tone 1 can be identified. For example, if the f0

peak is found inside the dotted region, it means the f0 peak occurs in the onset of the following syllable. In other words, the f0 peak delay occurs. The end of the second syllable is marked by capital letter X. The second syllable starts from the beginning of the dotted region and ends at the temporal location marked by X.

(25) a. Tone 2 + Tone 3 (LH + L)
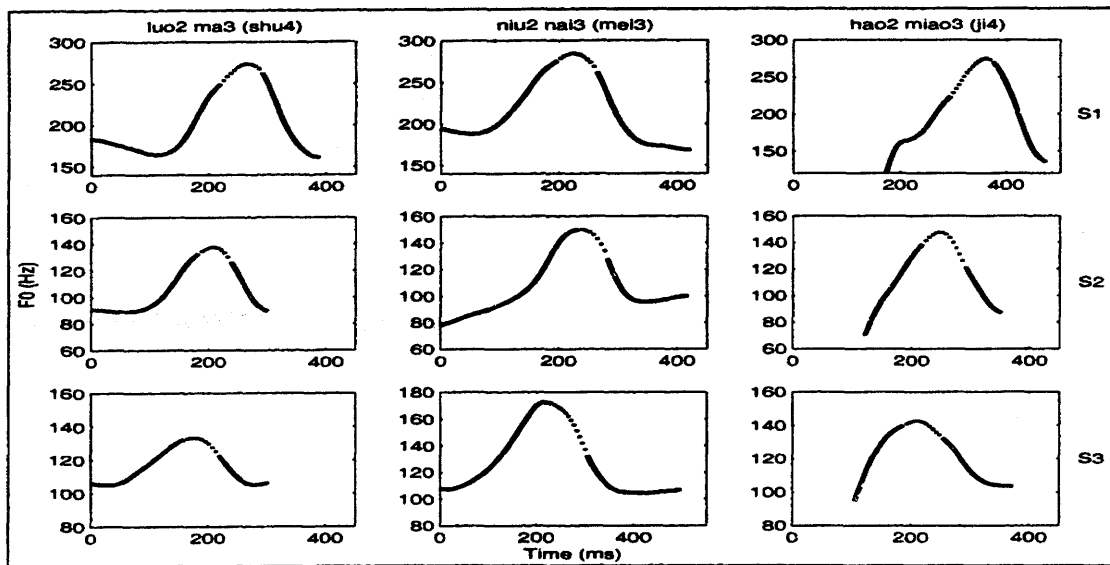
The f0 curves corresponding to tone 2 (LH) + tone 3 (L) are shown in (25a). The numbers in pinyin transcription indicate which tone the syllable carries. Each speaker's data are presented on each row, indicated by S1, S2 and S3. There are three target words for this tonal combination, corresponding to three columns in the figure. It is worth noting that the second target word *niu2 nai3* 'milk' is followed by *mei3 tian1* 'everyday'. This is a context in which the third tone sandhi can apply to two adjacent tone 3's, turning the first tone 3 into tone 2. This does not happen in our data because of the following two reasons. First, the third tone sandhi applies optionally at a major prosodic boundary (e.g. subject − predicate) (Duanmu 2000), which is the case for the carrier

sentence in which the target word *niu2 nai3* is embedded. Second, all target words were read with focus, which makes them less likely to undergo tone sandhi if the second tone 3 belongs to the following word. For all three speakers, the f0 peak of the rise mostly occurs in the following syllable onset (i.e. dotted region). There are almost no instances of f0 peak encroaching on the following syllable rime (i.e. to the right of the dotted region). This pattern of peak delay agrees with Xu's (2001) finding.

(25) b. Tone 2 + oneNT (LH + short L)
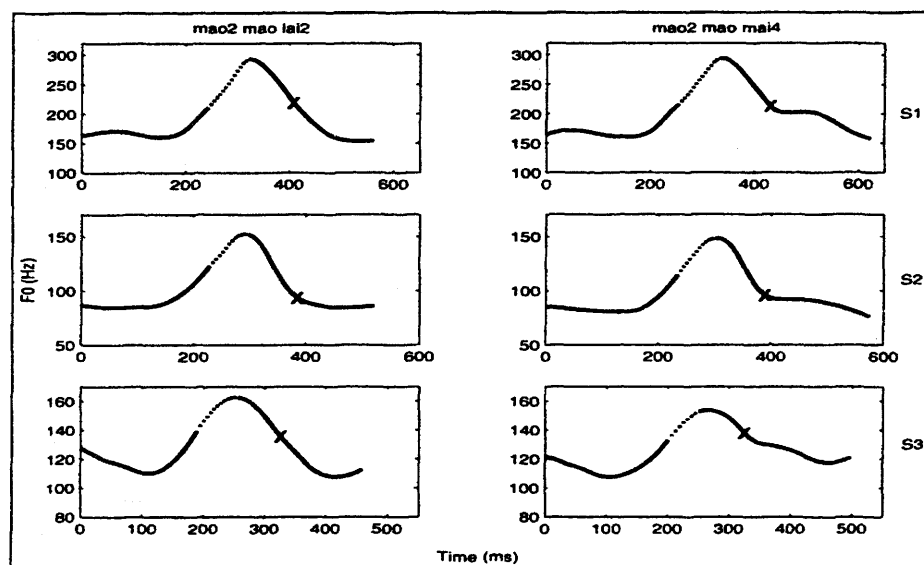
The f0 curves corresponding to tone 2 (LH) + one neutral tone (short L) are shown in (25b). X marks the end of the neutral tone syllable. The target word *mao2 mao* is followed by a syllable in tone 2 (LH) and tone 4 (HL) respectively. Notice that the short L target on neutral tone syllables is clearly identifiable when it is followed by HL (in the second column) despite the fact that the fall is considerably reduced and lowered in register because of the post-focus effect (the target word is focused) (Xu 1999). The short L target coincides with the temporal location marked by X. In the first column, the

neutral tone is followed by LH. Since the post-focus rise is also reduced and lowered in register, the short L target on neutral tone syllables are not identifiable at least for speakers S1 and S3. However, comparing the f0 values reached at the end of neutral tone syllables (i.e. the temporal location marked by X) in the first and second columns, we find that they are almost identical for each speaker. Therefore, when the L target is not identifiable, we simply take the f0 at X as its value. The tonal contexts are identical in (25a) and (25b): a rising target followed by a low target. The metrical contexts are different: the second syllable is stressed in (25a), hence longer in duration, while it is unstressed in (25b). The metrical difference will be shown to have an impact on how the f0 peak is aligned in these two contexts. Right now, we observe that the f0 peak is found at the end of the following onset (i.e. end of the dotted region) or the beginning of the following rime. A quick question which can be asked is whether there is more peak delay when the rise is followed by a neutral tone syllable (25b) than when it is followed by a stressed syllable in tone 3 (25a). One may argue, by assuming that duration of the rise is constant in these two contexts[10], that the more f0 peak delay is only an artifact of the shorter syllable duration on the following unstressed syllable. In order to contravene this argument, one has to show that the rise is lengthened when it is followed by a neutral tone, and thus that the duration of the rise is sensitive to the following context rather than being fixed regardless of context. This is indeed the case, as we will show later.

---

[10] Two further assumptions that have to be made are the implementation of the rise begins at the left edge of its host syllable, and the following context has no effect on where the implementation of the rise begins. See Xu and Wang (2001) for relevant discussions.

(25) c. Tone 2 + twoNT (LH + N + short L)
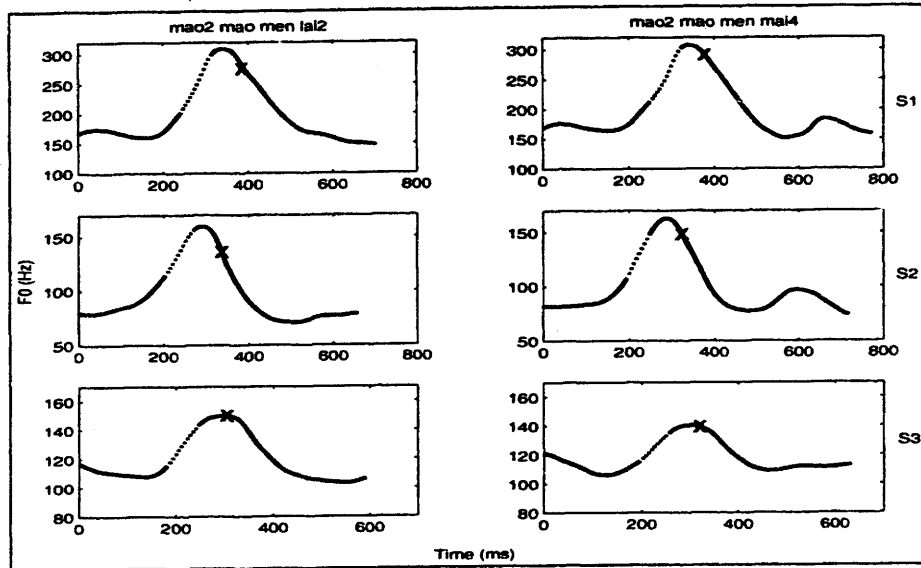
The f0 curves corresponding to tone 2 (LH) + two neutral tones are shown in (25c). X marks the end of the first neutral tone syllable. Since there are two neutral tone syllables, the short L target appears on the second. It is identifiable when followed by a fall as in the second column. When it is sometimes not quite identifiable (in the first column, followed by a rise), the f0 value at the end of the second neutral tone syllable (not shown in the figure) is taken to be the target value of L. What is striking about this pattern is that the f0 peak of the rise occurs in the middle of the following rime, and in a few cases, even near its end. There are no instances in which the f0 peak crosses over the first neutral tone and migrates all the way into the second.

(25) d. Tone 2 + threeNT (LH + N + N + short L)

mao2 mao men le (cai2)   mao2 mao men le (jiu4)
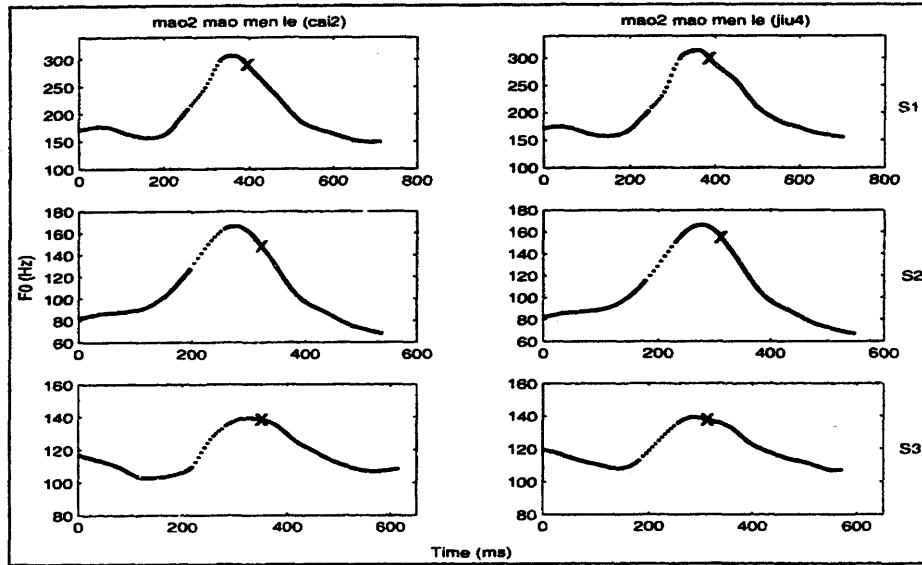
The f0 curves corresponding to tone 2 (LH) + three neutral tones are shown in (25d). The short L target is found on the final neutral tone syllable. That is why we observe a much longer and less steep fall from the f0 peak to the end. The target words *mao2 mao men le* are followed by syllables in tone 2 and tone 4 in the two columns. The f0 curves corresponding to the following tone 2 and tone 4 are not shown in the figure. The occurrence of the f0 peak follows the pattern in (25c): it appears in the middle of the following syllable rime (first neutral tone syllable), and sometimes near its end. Again, there is no f0 peak migrating into the second neutral tone syllable.

(25) e. Tone 2 + twoNT / threeNT in utterance-final position
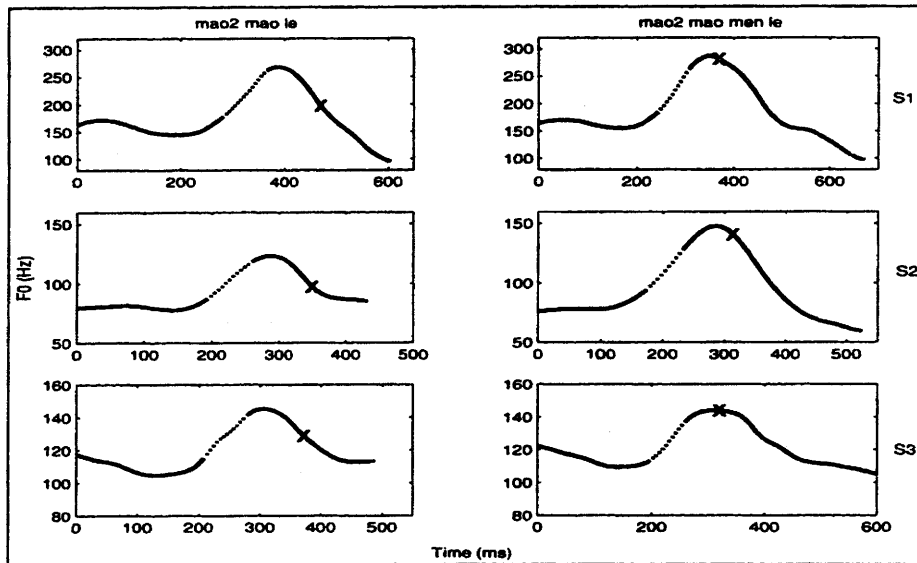
The f0 curves corresponding to tone 2 (LH) + two and three neutral tones are shown in (25e). The target words occur in utterance-final position. X marks the end of the first neutral tone syllable. It is observed that the f0 peak alignment patterns are similar to (25c, d) where the target words are in utterance-medial position.

(26) a. Tone 1 + oneNT (H + short L)
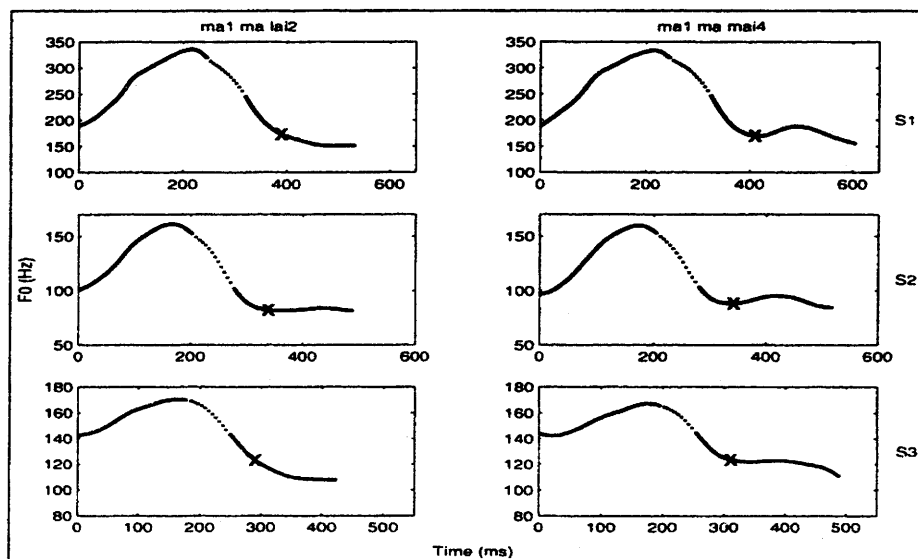
b. Tone 1 + twoNT (H + N + short L)
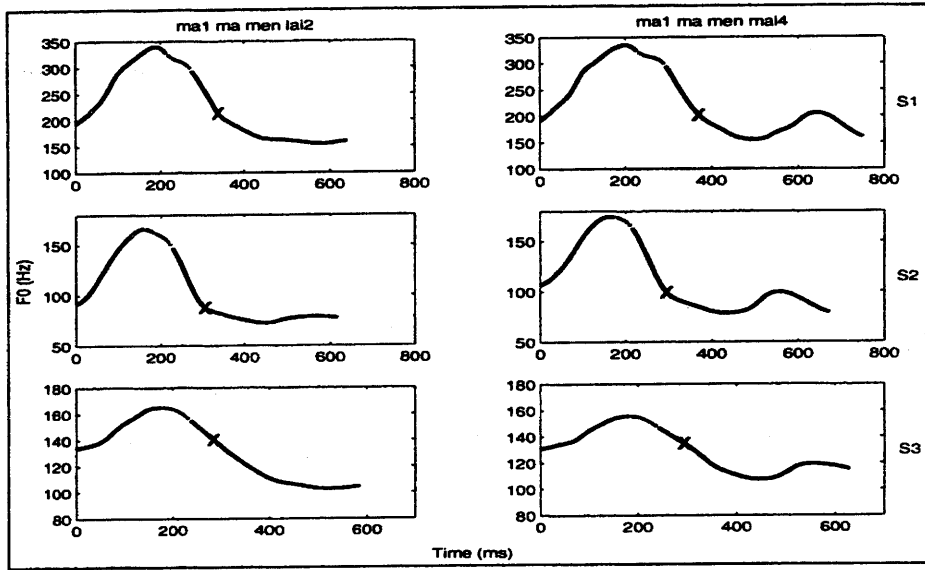
c. Tone 1 + threeNT (H + N + N + short L)
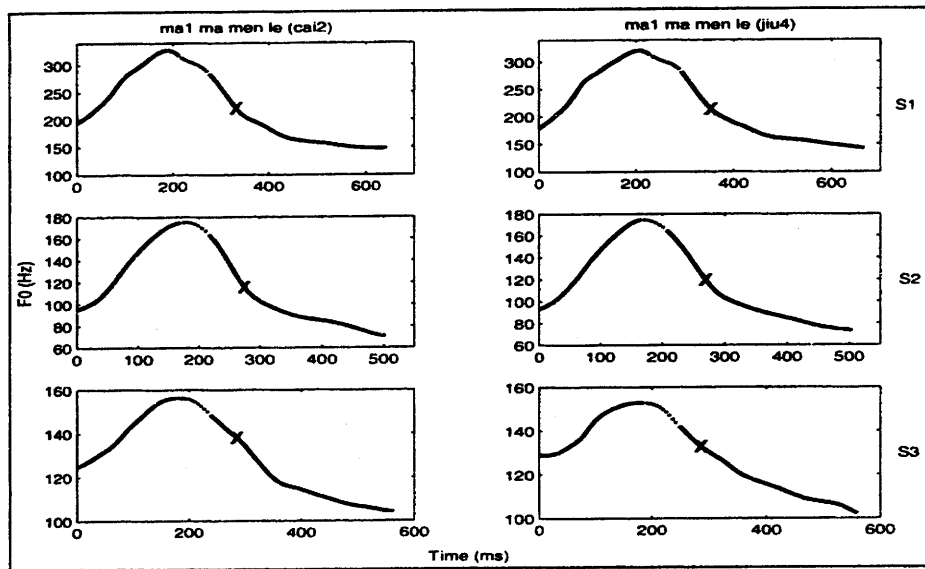
d. Tone 1 + twoNT / threeNT in utterance-final position
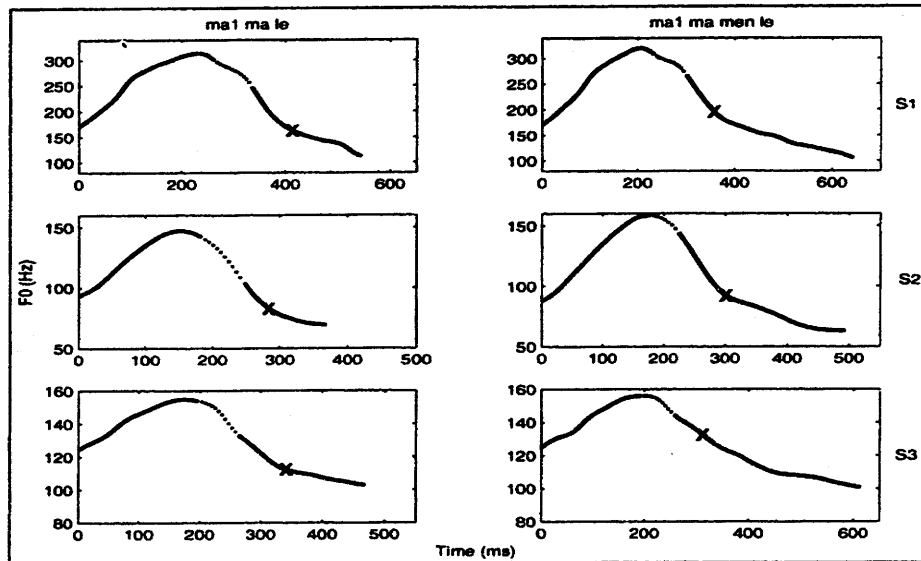
The f0 curves corresponding to tone 1 (H) in different metrical contexts are shown in (26a-d). There is a rising f0 transition before the f0 peak due to the assimilatory effect of the preceding tonal context (Xu 1997). It starts right at the beginning of the syllable in H tone. It differs from the delayed rising f0 observed in a rise in (25a-e) where it starts almost in the middle of its host syllable. The consequence of such an alignment difference will be explored later. In all contexts, there is no f0 peak delay for H: the f0 peak is always found toward the end of its host syllable, but before the following syllable onset (i.e. the dotted region).

On the basis of the above discussion, the f0 peak alignment in tone 2 (LH) and tone 1 (H) can be summarized as follows. First, the f0 peak in tone 2 exhibits variations in its alignment as a function of the following metrical contexts. Second, there is no peak delay in tone 1 regardless of the following contexts.

It is yet to be shown that the variations in f0 peak alignment in tone 2 are not an artifact of the reduced duration of neutral tone syllables. One may contend that duration of rise is held constant in all contexts, but when the following syllable becomes shorter, it looks like the f0 peak is delayed more.

The measurement of peak-to-C2 is used to assess this hypothesis. Recall that peak-to-C2 is the location of the f0 peak relative to the end of its host syllable. A positive value indicates that the f0 peak occurs before the end of the syllable and a negative value indicates peak delay. It is illustrated below.

(27) Illustration of peak-to-C2: delay = $\text{time}_P - \text{time}_{C2}$

a. no peak delay: delay ≤ 0 b. peak delay: delay > 0

σ_1=H σ_2=N σ_1=LH σ_2=N

P   P

C1   C2   C1   C2   time
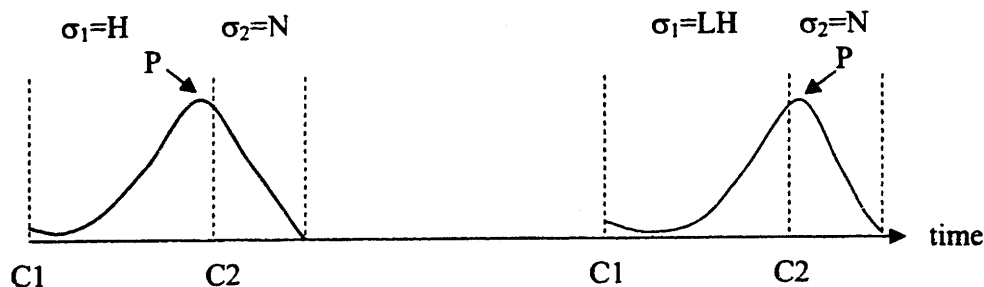
C1 marks the beginning of the syllable in tone 1 (H) and tone 2 (LH). C2 marks the beginning of the neutral tone syllable (N). The f0 peak is indicated by P. The amount of peak delay can be measured by $\text{time}_P - \text{time}_{C2}$, i.e. the temporal location of P subtracted by the temporal location of C2. The first syllable duration can be obtained by C2 – C1.

The absolute peak-to-C2 values are presented in (27). The four black bars represent the values of peak delay in four different contexts. They are, from left to right, peak delay in R when followed by a stressed syllable in L ("stressed"), a neutral tone syllable ("oneNT"), two neutral tone syllables ("twoNT") and three neutral tone syllables

("threeNT"). The peak delay in tone 1 is represented by three gray bars, corresponding to oneNT to threeNT contexts. ("NT" stands for neutral tone.) The numbers on top of or under the bars are the values of peak delay measured in milliseconds. Positive values mean that the f0 peak occurs before the end of its host syllable, hence, no delay. Negative values mean there is peak delay.

(28)   Averaged values of peak delay in different metrical contexts
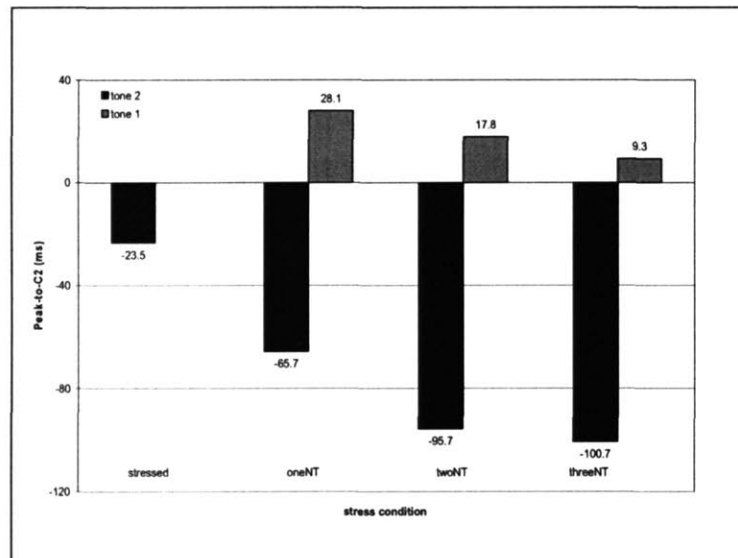
As shown in (28), for tone 1 all peak-to-C2 values are positive, indicating that on average the f0 peak in H occurs within its host syllable. The values are slightly decreasing from "oneNT" to "threeNT" condition, indicating that on average the f0 peak is moving closer to the right syllable boundary when there are more neutral tone syllables. The decrement is on the order of 10 ms. A different picture emerges for tone 2 with all negative values, which indicates the presence of extensive f0 peak delay. A closer examination of tone 2 data reveals the following three observations: (1) the peak delay value is significantly larger when the syllable after tone 2 is unstressed (i.e. in neural tone)

than when it is stressed (one-factor ANOVA, $p<0.001$), (2) the peak delay value is also significantly larger when there are two or three unstressed syllables than when there is only one (one-factor ANOVA, $p<0.001$), and (3) there is no significant difference when there are two or three unstressed syllables (one-factor ANOVA, $p=0.13$).

It is worth noting that the measurement peak-to-C2 is independent of whether the following syllable is stressed or unstressed, and having long or short duration. It is only affected by where the f0 peak is located and where its host syllable ends (i.e. C2). Given that the syllable duration is not significantly varied in different metrical contexts (i.e. C2 is a constant relative to C1), it is reasonable to assume that when the absolute value of peak-to-C2 is larger, there is more peak delay. We conclude that the variations in the f0 peak alignment are not an artifact of the reduced duration of the following unstressed syllable. The f0 peak is actually delayed more when the following syllable is unstressed.

Here we observe an asymmetry: when the following syllable has longer duration (stressed, in tone 3), the peak delay is shorter; when the following syllable has shorter duration (unstressed, in neutral tone), the peak delay is longer. Such an asymmetry is unexpected in stress clash theory in that the f0 peak is aligned significantly later when there are two or three unstressed syllables than when there is only one (see section 2.3.1 for a brief discussion of stress class theory). That theory would predict a dichotomy just between stressed and unstressed following contexts. The interactive account does not explicitly capture the fact that the following tonal context also affects the placement of f0 peak in R. It predicts that peak delay happens whenever there is a sharp f0 rise just before the syllable offset (Xu 2001), but it does not explicitly predict peak alignment

difference as a function of the following context. It does not predict where the f0 peak would be aligned.

The table in (29) shows the frequency of peak delay in tone 2 and tone 1. Without exception, the f0 peak in tone 2 occurs after the end of its host syllable when followed by neutral tone syllables. The frequency of peak delay increases monotonically in tone 1 as the number of following neutral tone syllables increases. Such a pattern is reflected in the quantitative values of peak delay in (28).

(29) Frequency of peak delay

	stressed	oneNT	twoNT	threeNT
Tone 2	81%	100%	100%	100%
Tone 1		4%	14%	25%

2.3.3.3 Duration and Excursion Size of F0 Contours

Duration and f0 excursion size of the rising and falling contours are also measured in the experiment. The measurements are summarized in tabular form below.

(30) Averaged duration (ms) and f0 excursion (Hz) of rising and falling contours

			stressed	oneNT	twoNT	threeNT
Tone 2	Rise	Duration	132.3	152.7	163.5	164.6
		Excursion	76.4	85.7	80.6	86.1
	Fall	Duration	125.4	127.7	195.6	287.7
		Excursion	82.1	70.6	92.6	107.3
Tone 1	Rise	Duration		182.0	181.6	182
		Excursion		85.6	87.3	90.0
	Fall	Duration		159.4	282.5	372.2
		Excursion		97.2	113.3	118.8

Since tone 2 is realized as a delayed rise, starting nearly in the middle of its host syllable, rise in tone 2 in the above table refers to the real rising part of f0, defined in section 2.3.2. Rise in tone 1 refers to the rising f0 transition before the f0 peak is reached. Fall in tone

1 and tone 2 refers to the falling f0 from the f0 peak in tone 1 and tone 2 to the next L target. For example, when tone 1 is followed by a stressed syllable in tone 3 (L), there is a falling contour between the f0 peak and the L target of tone 3. Similarly, when tone 1 is followed by three neutral tone syllables, there is a long falling contour from the f0 peak to the L target on the last neutral tone syllable.

Several observations can be made of the duration and f0 excursion data given in (30). First, there is an increase in the duration of rising f0 (rise time) in tone 2 from "stressed" to "oneNT" condition and from "oneNT" to "twoNT" condition. The rise time does not differ in "twoNT" and "threeNT" conditions, though. Second, the rise time in tone 1 does not differ in the context of unstressed syllables. It is longer than the rise time in tone 2. Third, the f0 excursion size in rising f0 tends to increase with rise time, but this is not always the case (for example, rise excursion in "twoNT" condition for tone 2). Fourth, the fall duration increases with the number of unstressed syllables. This is because fall duration is the duration from the f0 peak to the L target on the *last* neutral tone syllable. Fifth, the f0 excursion size in falling f0 does increase with fall time. The following plot, based on the table above, illustrate the observations.

(31) Averaged duration (ms) and f0 excursion (Hz) of rising and falling contours

61
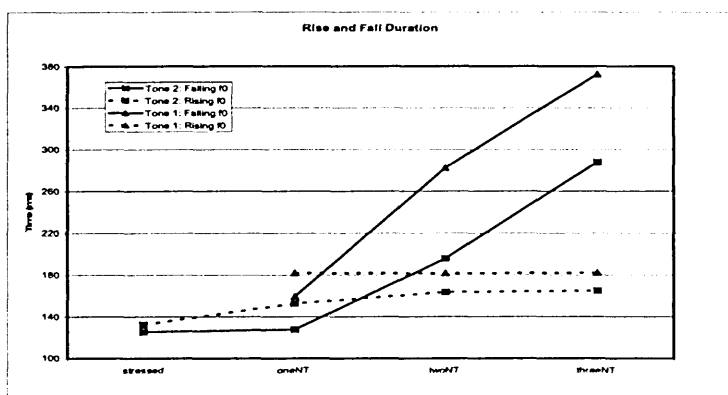
2.3.3.4 Summary

The results on the alignment of f0 peaks have revealed a strong effect of the following metrical context. For f0 peak in tone 2, it occurs in the following syllable onset when followed by a stressed syllable in tone 3, near the beginning of the following syllable rime when followed by an unstressed syllable, and in the middle of the following rime when followed by two or three unstressed syllables. In terms of mean values of peak delay, the following pattern emerges: "stressed" condition < "oneNT" < "twoNT/threeNT". Accordingly, the rise time follows the same pattern: it becomes longer when the rise is followed by neutral tone syllables. On average, f0 peak in tone 1 occurs before the end of its host syllable in the context of unstressed syllables. However, the effect of following context is also seen in terms of the frequency of peak delay, as shown in (29). Both Xu's (2001) theory and "stress clash" theory have to be modified to capture these findings.

2.3.4 Analysis

We propose a formal analysis of tonal alignment to account for the asymmetries found in the f0 peak alignment in Mandarin. Assuming a distinction of phonology and phonetic implementation, the analysis we are proposing models the phonetic implementation of tonal alignment. We will first identify the phonological and phonetic constraints that pertain to tonal alignment and then determine their rankings to derive the alignment differences.

2.3.4.1 Perceptual Distinctiveness of Tonal Contrast

We have demonstrated the use of the constraint on tonal contrast in the selection of the correct allotone of tone 3 in non-final position. The upshot of our analysis is that tonal contrast is maintained in the tone mapping unless the markedness constraints on tonal distribution compels neutralization. The constraint on tonal contrast favors the mapping which will not compromise the perceptual distinctiveness of phonological contrastive tones.

Distinctiveness of phonological contrast has been argued to regulate contextual variations, especially vowel coarticulation (Manuel 1990, Flemming 1997), as we noted earlier. Effect of tonal contrast on f0 timing has been identified recently (Myers, to appear). In Mandarin, the four lexical tones on stressed syllables have distinctive f0 patterns in citation form. The f0 patterns are modified in context, resulting in either categorical change (i.e. tone sandhi) or contextual variations (Xu 1994, 1997). These variations could cause dramatic deviations from the f0 patterns in citation form, and even obscure the distinctiveness of two contrasting f0 patterns. For example, a rising f0 transition emerges in a syllable in **H** before a tone with a low f0 ending (L.H or HL.H), which could lead to a potential tonal ambiguity between tone 1 (H) and tone 2 (LH) before L. Similarly, a falling f0 transition appears in a syllable in **L** after a tone with a high f0 ending (H.L or LH.L), which could lead to a potential tonal ambiguity between tone 3 (L) and tone 4 (HL). Nevertheless, crucial acoustic information is still available to preserve the perceptual distinctiveness of tonal contrast despite the contextual modification.

We propose that differences in tonal alignment carry crucial acoustic information to convey the tonal contrast in context. First we show how the contrast is preserved between H and R. The crucial difference lies in where the f0 begins to rise. It has been found that in L.H the f0 rises immediately at the syllable boundary where L is realized. In contrast, the f0 in R does not rise until the middle of syllable rime. The late rise in R makes the dynamic portion of the rising f0 fully approximated in the later portion of a syllable regardless of syllable duration (Xu 1998, 1999). When a rising tone **R (LH)** follows a tone with a high f0 ending, H.LH or LH.LH, there is always a falling f0 transition from H to L.[11] This period of f0 transition is an articulatory necessity since it takes time to reverse the direction of f0 movement (Xu and Sun 2002). When H is anchored at the end of the preceding syllable, the beginning of the rising f0 in the second syllable has to be late. However, the late rise is unexpected when R follows a tone with a low f0 ending, i.e. L.LH,[12] creating a low f0 plateau. In this context, no f0 transition is necessary between two adjacent L targets. So one would expect an immediate rise right after the syllable boundary (in effect, the rising accent in Modern Greek, British English and Dutch). But that is not found in Mandarin. We suggest that the late rise in R renders R perceptually distinct from H by creating a rising f0 curve with abrupt slope. The result is a better, less confusable contrast between H and R. The perceptual motivation of late rise finds support in House (1990). Through a series of psychoacoustic experiments, he shows that rapid spectral changes during the transition from the consonant to the vowel

---

[11]   The rising tone under discussion is not in a prosodically weak position (i.e. being the middle syllable of a three-syllable group, according to Chao 1968), therefore it does not undergo tone 2 sandhi discussed in section 2.2.2.

[12]   The late rise in this tonal context is most conspicuous when the syllable carrying LH is focused, as can be seen in the f0 plots in figure 4 (the panel in the bottom row of the middle column) in Xu (1999).

significantly diminish the hearer's sensitivity to pitch movement. It follows from his theory that a linguistically relevant pitch target avoids the consonantal region. A similar pattern of late rise is also found in Buli, a Gur language spoken in Ghana (Kenstowicz and Akanlig-Pare 2003).

Second, we show how the contrast is preserved between L and F. In this case, the crucial difference lies in where the f0 peak is aligned. The f0 peak in F occurs early in the vowel or near the middle of the syllable regardless of whether the preceding tone has a low or high f0 ending.[13] In contrast to the late fall in F, the falling f0 transition in **LH.L** is early in that the f0 peak occurs before the vowel onset. The perceptual relevance of the f0 peak location in discriminating F and L has been demonstrated in a perception experiment reported by Gårding et al. (1986) in which the location of the f0 peak was systematically shifted and lowered. Results showed that stimuli with the f0 peak appearing after vowel onset were clearly identified as F (tone 4). In other words, when the f0 peak originally associated with LH.L context occurred after the vowel onset in the second syllable, L would be perceived as F, changing the tonal identity from tone 3 to tone 4.

We adopt the family of constraints on tonal contrast preservation that we proposed before, PRESERVECONTRAST(T_i/T_j), according to which T_i and T_j have to be perceptually distinct. For example, PRESCONTRAST(L/HL) penalizes candidates that blur the perceptual boundary between L and HL in the sense we discussed above. In a system with four lexical tones like Mandarin, there are six such constraints (six unordered pairs

---

[13]   When the preceding tone is H or R, there is no high f0 plateau between H/R and F. Rather only one f0 peak shows up in the syllable in F.

out of four tones). Only PRESCONTRAST(L/HL) and PRESCONTRAST(H/LH) are relevant to our analysis. Note that these constraints only apply to the stressed syllables on which four lexical tones are contrasted. There is no tonal contrast on unstressed syllables. Consequently, the constraint on tonal contrast preservation is irrelevant there. This point is crucial in our analysis of the alignment asymmetry of f0 peak in rise.

2.3.4.2 Constraint on Contour Duration

Although the duration of a rising or falling f0 contour varies with context, as we have illustrated in Mandarin, there is a canonical duration or target duration that a contour tone will approximate for a certain pitch excursion. We assume that the canonical duration is approximated when there is no time pressure, as in citation form, or when followed by an unstressed syllable. We also assume that it varies with pitch excursion size, i.e. larger pitch excursion requires longer canonical duration (cf. Xu and Sun 2002). In context, it is not always reached as a result of contextual factors (for example, Modern Greek as reported in Arvaniti, Ladd and Mennen 1998, Buli as reported in Kenstowicz and Akanlig-Pare 2003).

Despite small cross-linguistic variations, the duration of a rise or fall at a specific speech rate is largely comparable. For example, the rising accent takes about 200 ms in Dutch (Ladd, Mennen and Schepman 2000) and about 190 ms in British English (Ladd, Faulkner, Faulkner and Schepman 1999) when there is no time pressure. These values are comparable to the citation duration of the rising part in tone 2 in Mandarin, about 200 ms, as estimated from figure 1. The rising f0 transition in L.H context is about 180 ms, obtained from our experiment (shown in the table in 30).

In addition to canonical duration, a contour tone is subject to an articulatory constraint regarding how fast a pitch shift can be possibly made. In other words, for a given pitch excursion, there is a minimal amount of duration required to implement a rise or a fall. Xu and Sun (2002), following up on Ohala and Ewan (1973) and Sundberg (1979), assessed the maximum speech of pitch change by asking native speakers of English and Mandarin Chinese to imitate synthesized pitch undulation patterns. They found that the maximal speech of pitch change is employed more often in speech than previously thought by comparing their experimental results with Caspers and van Heuven (1993), Ladd, Faulkner, Faulkner and Schepman (1999), Xu (1999), and Ladd, Mennen and Schepman (2000).

In light of the above discussions on contour duration, we suppose that for a given pitch excursion, the canonical duration is what is being aimed for in the production of a contour tone. The actualized duration normally falls between the canonical duration and the minimal duration, dictated by the articulatory constraint in f0 production. The actual situation is more complex in that under time pressure both duration and pitch excursion can be manipulated. To simplify the picture, we assume that the pitch shift for a contour does not vary much at a specific speech rate when the focus condition is held constant. We define a family of contour markedness constraints which evaluates the closeness of the actual duration of a contour to the canonical duration:

(32) a. $\Delta d_i = d_c$ (canonical duration) $- d_i$ (actual duration)

     b. MAXIMIZEDURATION(CONTOUR): the contour duration must be maximized such that it approaches the canonical duration of the contour.

     c. d_m better satisfies MAXIMIZEDURATION(CONTOUR) than d_n if $\Delta d_m < \Delta d_n$

Rise and fall can be referred to by this constraint: MAXDUR(RISE)[14] and MAXDUR(FALL).

Given a candidate x, MAXDUR(CONTOUR) does not assign a list of violation marks to x.

Rather, the constraint asserts that y in the same candidate set is more harmonic than x (i.e.

$y \rangle x$). The canonical duration for rise is taken to be 180 ms, which is the duration of

rising f0 transition in the L.H context. It is the rising f0 obtained under no obvious time

pressure in our experiment. The canonical duration for fall is irrelevant in our analysis,

but it can be estimated from the figure 1. It is smaller than that for rise.

Sometimes the rise or fall is seemingly longer than the canonical duration in the

sense defined above. For example, in our experiment when R or H is followed by two or

three neutral tone syllables, there is a long falling f0 from the f0 peak in R and H till the

f0 minimum at the end of the last neutral tone syllable (from 200 ms to 400 ms). This is

because the falling f0 spans two or three syllables. We stipulate that the candidate with

duration longer than the canonical duration is equally harmonic to the candidate with

canonical duration.

Another class of constraints on contour duration defines the minimal duration

required to implement a contour with given pitch excursion. As we mentioned earlier,

how fast a pitch change can be made is subject to physiological limitations of the human

vocal apparatus. Therefore, for a given pitch excursion, a threshold duration has to be

allocated. Xu and Sun (1992) is the most recent study which provides empirical

estimation of the threshold values. For our purposes, we assume that for a rise or fall, the

---

[14]   Abbreviation for MAXIMIZEDURATION(RISE).

minimal duration requirement has to be satisfied. Such a requirement can be modeled by the MINIMALDURATION(CONTOUR) constraint, defined below:

(33) a. MINIMALDURATION(CONTOUR): the contour duration may not be smaller than the minimal duration required to implement the contour for a given pitch excursion size.

b. Given $\Delta d_i = d_m$ (minimal duration) $- d_i$ (actual duration), d_i satisfies MINIMALDURATION(CONTOUR) when $\Delta d_i \leq 0$.

Like MAXIMIZEDURATION(CONTOUR), MINIMALDURATION(CONTOUR) is relative to rise and fall. The minimal duration for a fall is taken to be 125 Hz and that for a rise 130 Hz. They are estimated from the table in (30). The MINIMALDURATION(CONTOUR) is unviolated in our analysis. It is noted that a more realistic implementation of these two classes of constraints should incorporate pitch excursion as an argument. To simplify the analysis, we assume that the pitch shift does not undulate when the focus condition is held constant. They will be ranked with other constraints in the grammar to derive the observed tonal alignment patterns.

2.3.4.3 Other Constraints

We identify the other constraints that will be used in our analysis. We observed that the f0 peak in R is delayed into the following vowel when there is more than one unstressed syllable, but it never goes into the second unstressed syllable even when there are three unstressed syllables following R. It seems that the f0 peak in R has to maintain a local relation with the syllable that it belongs to. But given the MAXCONTOURDUR constraint, the longer rise is always the winner when tonal contrast is not involved. We suggest that

if that syllable the bears R and the syllable into which the f0 peak in R migrates are not contiguous, then the structural affiliation of the f0 peak would be obscured. We propose that the relevant constraint that bans the excessive rightward migration of the f0 peak is CONTIGUITY(SYL-TONE). The constraint requires that the tonal target is realized either in its host syllable or in the immediately following syllable.

The other constraints align L and H tones with the left and right edge of the syllable: ALIGN-LOW(σ, LEFT) and ALIGN-HIGH(σ, RIGHT). They are defined as follows.

(34) a. CONTIGUITY(SYL-TONE): No syllable intervenes between the syllable that bears the tone in the input and the syllable that realizes the tone in the output. (after McCarthy and Prince 1995)

b. ALIGN-LOW(σ, LEFT): L tone is aligned with the left edge of the syllable. (after McCarthy and Prince 1993)

c. ALIGN-HIGH(σ, RIGHT): H tone is aligned with the right edge of the syllable. (after McCarthy and Prince 1993)

2.3.4.4 Constraint Rankings

By presenting the ranking arguments for the constraints we discussed above, we illustrate that the detailed alignment differences can be modeled in the framework of the Optimality Theory. Furthermore, phonological and phonetic constraints both are involved in phonetic implementation.

The f0 peak in R is likely to delay into the following segment because of the joint effect of PRESCONTRAST(H/LH) and MAXDUR(RISE). The former constraint demands a late rise in the syllable and the latter requires a longer rise. The rising f0 is not completed

by the end of the syllable as MAXDUR(RISE) favors a rise as close to the canonical duration as possible. If the rise starts in the middle of a syllable, the rise has to terminate *after* the syllable because half of the syllable duration (about 120 ms, as can be seen from (22) and (23)) is much smaller than the canonical duration for a rise, which is taken to be about 180 ms in our case. Therefore, PRESCONTRAST(H/LH) and MAXDUR(RISE) have to outrank ALIGN-LOW(σ, LEFT) and ALIGN-HIGH(σ, RIGHT). In the following tableau, the candidate that better satisfies MAXDUR(RISE) but violates PRES(H/LH) loses to the candidate that satisfies PRES(H/LH), motivating the ranking of MAXDUR(RISE) over PRES(H/LH).

In (35), we consider three candidates for the input R / L__L: (a) L: syllable onset, H: syllable offset; (b) L: middle of syllable, H: syllable offset; and (c) L: middle of syllable, H: following onset.

(35) PRES(H/LH) >> MAXDUR(RISE) >> ALIGN-LOW(σ, LEFT), ALIGN-HIGH(σ, RIGHT)

	R / L L	PRES(H/LH)	MAXDUR(RISE)	ALIGN-LO	ALIGN-HI
a.	L: σ onset H: σ offset	*! (T1/T2)			
b.	L: mid of σ H: σ offset		c ⟩ b	*	
c. ☞	L: mid of σ H: following ons			*	*

In (a), the beginning of the rise is aligned with the onset of its host syllable, and the f0 peak is aligned with the offset of the syllable. In order words, the rise is spanning over the whole syllable. In (b), the beginning of the rise is aligned with the middle of its host syllable, and the f0 peak is aligned with the syllable boundary. This is a very short rise.

In (c), the beginning of the rise is the same as in (b), but the f0 peak is delayed into the following syllable onset. The three candidates are schematically shown below:

(36) Schematic representation of three possible implementations of a rising tone

a. L: σ onset; H: σ offset   b. L: mid of σ; H: σ offset   c. L: mid of σ; H: following ons
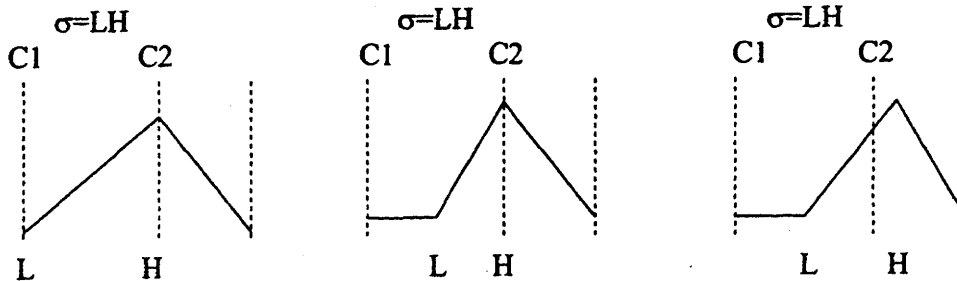
The syllable in rising tone is marked by C1 and C2. It is phonologically associated with a rise. Note that the R under discussion occurs in L__L context. The preceding L is aligned with the end of its own host syllable. Therefore, we do not consider candidates that fail to satisfy such an alignment pattern. It is conceivable that a rise could satisfy MAXDUR(RISE) by moving its beginning into the preceding syllable rather than delaying its f0 peak into the following syllable. This candidate fails to align L with the right edge of its host syllable, and it is excluded from consideration.

The optimal candidate is (c), which only violates the lower-ranked alignment constraints. (a) is ruled out by PRES(H/LH) despite the fact that its duration is the closest to the canonical duration. It fatally neutralizes the contrast between H (tone 1) and LH (tone 2) in the context of L__L because (a) is what H will be realized as in this particular context. MAXDUR(RISE) establishes the harmonic ordering relation between the remaining (b) and (c). Since (c) has longer rise duration than (b), it is picked as the winner.

The above analysis does not explain why R cannot be delayed further into the following syllable rime. When R is followed by L (i.e. a stressed syllable in tone 3), another constraint on contrast preservation PRES(L/HL) is invoked to prevent the f0 peak in R from occurring in the following syllable rime in order to preserve the perceptual distinctiveness of L and HL on the following syllable. Therefore, the f0 peak occurs in the following onset. The mean rise duration is 130 ms in this context. Like PRES(H/LH), it is ranked higher than MAXDUR(RISE). In the tableau below, the input is R / L__L, and the beginning of the rising f0 is in the middle of its host syllable in all candidates. For the sake of space, the alignment of L is not shown in the tableau. The f0 peak is in the middle of the following syllable (a), at the C-V boundary of the following syllable (b), and in the following onset (c). The tonal contrast preservation constraint rules out (a) and (b) since they compromise the perceptual distinctiveness of tone 3 (L) and tone 4 (HL) on the following syllable in the output.

(37) PRES(L/HL) >> MAXDUR(RISE)

R / L__L		PRES(L/HL)	MAXDUR(RISE)
a.	H: mid of following σ	!* (T3/T4)	
b.	H: following C-V boundary	!* (T3/T4)	
c. ☞	H: following onset		*

The f0 peak in H is not delayed because the rising f0 transition starts from the beginning of the syllable. Consequently PRESCONTRAST(H/LH) and MAXDUR(RISE) are both satisfied. The alignment of f0 peak in H is regulated by ALIGN-HIGH(σ, RIGHT). The analysis is given in (38). The input is H / L__L. The three candidates we are considering for the implementation of H in the context of L__L are the same as for the implementation of R / L__L. The only difference lies in the target tone in the input.

(38) PRES(H/LH) >> MAXDUR(RISE) >> ALIGN-LOW(σ, LEFT), ALIGN-HIGH(σ, RIGHT)

	H / L L	PRES(H/LH)	MAXDUR(RISE)	ALIGN-LO	ALIGN-HI
a. ☞	L: σ onset H: σ offset		a $\rangle$ c $\rangle$ b		
b.	L: mid of σ H: σ offset	*! (T1/T2)		*	
c.	L: mid of σ H: following ons	*! (T1/T2)		*	*

Candidates (b) and (c) both violate PRES(H/LH) because realizing a H like (b) and
(c) would neutralize the contrast between tone 1 and tone 2 in the output. (a) is the only
choice. It is also the closest to the canonical duration.

When R is followed by an unstressed syllable in neutral tone, PRES(L/HL)
becomes irrelevant because there is no tonal contrast on unstressed syllables. This is
illustrated in the following tableau.

(39) a. MINIMALDUR(FALL) >> MAXDUR(RISE)

	R / L l	MINIMALDUR(FALL)	MAXDUR(RISE)
a.	H: mid of following σ (R=165/F=100)	*!	
b. ☞	H: following C-V bnd. (R=150/F=128)		a $\rangle$ b
c.	H: following onset (R=130/F=125)		

The input is R / L__l, i.e. R followed by one unstressed syllable carrying a short L target.
The three candidates are the same as in (35), with durations of rising f0 and falling f0
specified on the basis of the data in (30). PRES(L/HL) is irrelevant and not shown.
Candidate (a) has a longer rise, but it loses to candidate (b) due to a crucial violation of
MINIMALDUR(FALL). However, a longer rise is still preferred because of MAXDUR(RISE)
when MINDUR(FALL) is satisfied. So (b) wins over (c).

74

In the above example, we found the f0 peak at the C-V boundary in the following syllable, with a mean rise duration of 150 ms. It is not quite a comfortable rise, though. We could not get a longer rise to make it more harmonic with respect to MAXDUR(RISE) because a high-ranking constraint, MINIMALDUR(FALL) has to be satisfied. As we discussed earlier, there is a short L target on the neutral tone syllable after R. The falling f0 curve from the f0 peak in R to this short L target has to satisfy MINIMALDUR(FALL). Moving the f0 peak into the following neutral tone syllable does not violate PRES(L/HL), but it crucially violates MINIMALDUR(FALL), for the duration of the unstressed syllable is not long enough for the falling f0 to complete its movement. MINIMALDUR(FALL) is undominated in Mandarin.

Another context in which MINIMALDUR(FALL) is satisfied is when there is more than one neutral tone syllable. The established ranking predicts that the candidate with longer rise should be the winner. The analysis is given in following tableau.

(40) MINIMALDUR(FALL) >> MAXDUR(RISE)

R / L N N l	MINIMALDUR(FALL)	MAXDUR(RISE)
a. ☞ H: mid of following σ (R=165/F=290)		a ⟩ b
b. H: following C-V bnd. (R=150/F>290)		

As shown above, when R is followed by more than one unstressed syllable, the f0 peak in R occurs in the middle of the following vowel, with a mean rise duration of 165 ms. Crucially, the L target is realized on the last neutral tone syllable, therefore there is always a long falling f0 after the f0 peak in R. This means MINIMALDUR(FALL) is satisfied, and PRES(L/HL) is still irrelevant.

Now the question is what prevents the rise from getting even longer, for example, longer than the canonical duration. Recall that when the actual duration is longer than the canonical duration, the candidate is more harmonic than the candidate with shorter duration. In our analysis, CONTIGUITY-SYL/TONE ensures that the f0 peak of R does not migrate two syllables away. Having a rise longer than 180 ms in this context will push the f0 peak into the second unstressed syllable, in violation of CONTIGUITY-SYL/TONE. The effect of CONTIGUITY-SYL/TONE is shown in (41).

(41) CONTIGUITY-SYL/TONE >> MAXDUR(RISE)

R / L N N 1		CONTIGUITY	MAXDUR(RISE)
a. ☞	H: mid of following σ (R=165)		a 〉b
b.	H: second unstressed σ (R=190)	*!	

The input is R followed by three syllables in neutral tone. In (b), the f0 peak is aligned with the second neutral tone syllable, and as a result, R has a longer duration. However, (a) wins because it satisfies the high-ranking CONTIGUITY-SYL/TONE.

Before concluding our analysis, we consider a context in which the f0 peak in R can actually occur in the following rime even when followed by a stressed syllable. In our theory, PRES(L/HL) preserves the perceptual distinctiveness of L vs. HL in the following syllable in L. When R is followed by F or H, the two H targets (i.e. LH.HL) merge into one due to OCP, and a single f0 peak shows up in the early part of the following vowel.[15] As a result, the rising f0 terminates in the rime of the following syllable. The mean f0 curves from Speaker 3 are displayed in (42a). It is noted that

---

[15]   As observed in Xu (1999), the f0 peak in this context is associated with F rather than R.

PRES(L/HL) is *not* violated since the f0 peak is in the right place for F in the following syllable. In our analysis given below, PRES(L/HL) and MINIMALDUR(FALL) are satisfied by all candidates. Therefore, candidate (a) with the longest rise wins.

(42) a. Tone 2 + Tone 4 (LH + HL) (top), Tone 2 + Tone 1 (LH + H) (bottom)
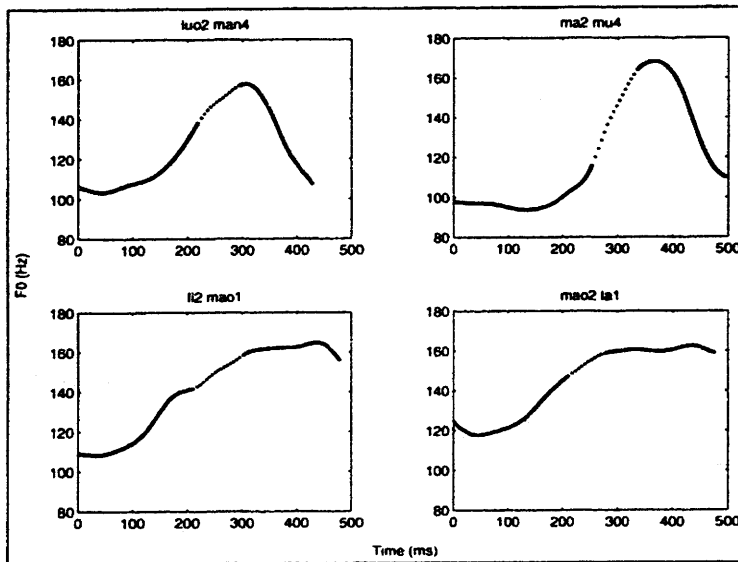
b. MINIMALDUR(FALL) >> MAXDUR(RISE)

R / L___F		MINIMALDUR (FALL)	MAXDUR (RISE)
a. ☞	H: mid of following σ (R=160/F=120)		a 〉b 〉c
b.	H: following C-V bnd. (R=140/F=130)		
c.	H: following onset (R=130/F=140)		

The rankings are summarized in (43).

(43) a. MINIMALDUR(FALL), PRES(T_i/T_j), CONTIGUITY-SYL/TONE >> MAXDUR(RISE)

>> ALIGN-LOW(σ, LEFT), ALIGN-HIGH(σ, RIGHT)

2.3.5 Summary

We have shown that the alignment of f0 peak in R varies with the following metrical and tonal context. One important factor in limiting the alignment of f0 peak is the constraint that preserves the perceptual distinctiveness of tonal contrast, e.g. L vs. HL in the current context. When PRES(L/HL) is satisfied, a rise closer to canonical duration is preferred over a shorter rise because of MAXDUR(RISE). We propose that the detailed alignment differences can be modeled in the framework of the Optimality Theory through the interaction of phonological and phonetic constraints on tonal contrast, contour duration and tonal alignment. This approach to the problem suggests that there is strong parallelism between phonology and phonetic implementation in that both domains can be modeled as systems of constraint interaction. Further research could be pursued along the following two lines. First, languages with different tonal contrast patterns could be studied to cross-examine the effect of tonal contrast on tonal implementation. Second, the OT approach to tonal implementation could be extended to account for the cross-linguistic difference in tonal alignment.

2.4   Concluding Remarks

In this chapter, we provide further evidence to justify the move to incorporate constraints on phonological contrast in the phonology. We have pursued the issue from both phonological allotonic selection and phonetic timing of f0 targets, drawing on evidence from Mandarin Chinese. We have demonstrated that the two processes are both subject to constraints on preserving tonal contrast. In an effort to model phonetic implementation

in the framework of OT, we have showed that both phonological and phonetic constraints are needed in the phonetic implementation component of the grammar. More importantly, phonological constraints such as constraints on tonal contrast have been shown to operate in both phonology and phonetics.

In the following chapters, we will explore other aspects of tone mapping. Some of the findings and constraints in the chapter will be invoked to explain other tonal processes in a number of Chinese dialects.

Chapter 3 Neutral Tone and Boundary Intonation

3.1 Introduction

The primary goal of this chapter is to explore the phonetic and phonological properties of neutral tone in Mandarin Chinese, with particular reference to its relation to boundary intonation. After laying out the basic properties of neutral tone, we provide a succinct review of previous analyses and an overview of the present analysis. Then we report results pertaining to pitch targets of neutral tone from the phonetic experiment described in chapter 2. We focus on the effect of boundary intonation on neutral tone. Finally we present an analysis for the alternation of full tone and neutral tone on monosyllabic function words in Mandarin. Our major proposals are the following. Functional elements (pronouns and prepositions) are de-stressed, and consequently lose their underlying tones, when they occur in the non-initial position of a prosodic word. The neutral tone syllables are toneless (either underlyingly toneless or via tone deletion), and their surface pitch patterns are determined by boundary intonation.

3.2 Basic Properties of Neutral Tone in Mandarin

3.2.1 Distribution of Neutral Tone

Mandarin Chinese distinguishes full and weak syllables. A full syllable is stressed and carries one of the four lexical tones. In contrast, a weak syllable is unstressed. The

contrast of four lexical tones is lost on a weak syllable. Hence it is said in neutral tone (e.g. Chao 1968). A weak syllable is also called a neutral tone syllable or a toneless syllable.

There are different types of neutral tone syllables, depending on whether they are always toneless or derived via deletion of their lexically associated tones. A very small number of morphemes, such as perfective suffixes, question markers and sentence particles, are always unstressed, and said in neutral tone. There is no way to identify their underlying tones. When said in isolation, they become stressed and take on tone 1 (high level). Other neutral tone syllables originate from stressed syllables, but are said in neutral tone when used in certain grammatical or prosodic contexts. For example, the directional verbal ending *lai* is used as a verb in its stressed, toned form (tone 2), meaning 'to come'. Some functional elements, for example pronouns and prepositions, alternate between stressed, toned and unstressed, toneless forms, depending on the prosodic contexts in which they occur. Some examples are given below. Most of them are taken from Dong (1958). The syllables in neutral tone are not marked.

(1)   a. suffixes: zi, tou, men 'plurality marker', mo

      examples: yĭ zi 'chair', mù tou 'wood', wŏ men 'we'

    b. particles: de 'possesive marker', le 'aspect marker (perfective) and sentence

      particle', guo 'aspect marker (perfective)', zhe 'aspect marker (progressive)'

      examples: wŏ de 'mine', lái le 'come+aspect marker'

    c. localizers: li 'inside', shang 'above'

      examples: wū li 'inside the house', tiān shang 'in the sky'

    d. pronouns as objects: wo 'I', ni 'you', ta 's/he'

examples: zhǎo ni 'look for you', jiào wo 'call me'

e. verbs reduplicated as cognate objects:

examples: kàn kan 'have a look', shūo shuo 'say it', xiǎng xiang 'think it over'

f. directional verbal endings: lai 'come', chu qu 'go out',

examples: ná lai 'bring here', zǒu chu qu 'walk out'

In a few disyllabic words, whether the second syllable is stressed or unstressed makes a lexical contrast. For these words, the metrical status of the second syllable is an idiosyncratic property of the word. In (2), each pair is distinguished by the second syllable: it is stressed in the first column and in neutral tone in the second.

(2) mǎi mài 'buying and selling' ----- mǎi mai 'business'

   xíng lǐ 'salute' ----- xíng li 'luggage'

   dōng xī 'east and west' ----- dōng xi 'thing'

   bào chóu 'revenge' ----- bào chou 'payment'

   láng tóu 'wolf head' ----- láng tou 'hammer'

   dà yì 'outline' ----- dà yi 'careless'

In a few other disyllabic words, the second syllable is always unstressed and said in neutral tone, such as *pián yi* 'cheap', *páng xie* 'crab', *yá men* 'government office in feudal China' and *zhuó mo* 'to ponder'. There are a couple of disyllabic words in which the second syllables can be optionally stressed or unstressed with no meaning difference, such as lǎo hǔ/hu 'tiger' and bō lí/li 'glass'.

As seen from the preceding examples, neutral tone syllables are always attached to stressed syllables, forming a trochaic pattern. In disyllabic words, if one syllable is in neutral tone, it must be the second one. In more general terms, neutral tone syllables are

prohibited from occurring in the initial position of a prosodic domain, leaving undefined what it is for now.

3.2.2   Phonetic Pitch of Neutral Tone

Neutral tone syllables are generally perceived as short and lax. Unlike stressed syllables which carry one of the four lexical tones, they do not have independent pitch values. It is generally assumed in traditional analyses that the phonetic pitch of neutral tone is determined by the preceding lexical tone, but not the underlying tone of the neutral tone syllable itself (e.g. Chao 1948, 1968, Dong 1958, T. Lin 1962). The phonetic pitch of neutral tone following four lexical tones is summarized below, using Chao's tone digits.

(3)   Phonetic pitch of neutral tone

context	pitch	example	gloss
after tone 1 (55)	2	tā de	'his'
after tone 2 (35)	3	huáng de	'yellow one'
after tone 3 (214)	4	nǐ de	'yours'
after tone 4 (51)	1	dà de	'big one'

In practice, it is sufficient to distinguish two pitch heights: high after tone 3 and low after the other tones (Chao 1968:36, Cheng 1973). This interpretation is confirmed in previous phonetic studies of neutral tone in Mandarin (Lin and Yan 1980, Cao 1986).

3.2.3   Acoustic Properties of Neutral Tone Syllables

Acoustic properties of neutral tone syllables in disyllabic contexts have been intensitively studied in recent experimental reports (Lin and Yan 1980, Cao 1986). On average, neutral tone syllables are about half as long as the corresponding stressed syllables. A

perception experiment also reveals that duration plays a vital part in the correct

identification of neutral tone syllables (T. Lin 1985). Intensity is not a consistent acoustic

cue for the contrast of full and weak syllables. The f0 patterns of neutral tone syllables

are found to be consistent with the above interpretation: they are high after tone 3 and

low after the other tones.

3.3   Previous Analyses and Overview of Present Analysis

3.3.1   Previous Analyses of Pitch Target of Neutral Tone

Given that neutral tone is realized as L or H pitch targets after the four lexical tones, the

question of interest is what conditions the choice. Consider the data in (4) first.

(4)   Pitch targets of neutral tone

      a. fei55 + le → 55 + L   'fly + ASP'

      b. lai35 + le → 35 + L   'come + ASP'

      c. mai214 + le → 21 + H   'buy + ASP'

      d. lei51 + le → 53 + L   'tire + ASP'

The aspectual marker *le* is said in neutral tone. Most analyses assume that the neutral

tone after tone 1, tone 2 and tone 5 takes the default pitch L (Yip 1980, J. Wang 1997,

Duanmu 1999). There are a variety of explanations for H after tone 3, depending on what

is taken to be the underlying form for tone 3. As we discussed in chapter 2, tone 3 is

realized as 214 in citation form and 21 before neutral tone and any lexical tone except

tone 3. Some take the citation form to be basic (Chao 1968, Milliken 1989); others take

the non-final form to be basic (Yip 1980, J. Wang 1997, Duanmu 1999). Milliken (1989) proposes that tone 3 is underlying L followed by a floating H, i.e. L(H). The floating H associates with the toneless neutral tone syllable, resulting in L-H sequence, as in (c). Yip (1980) and J. Wang (1997) propose a rule of H insertion after L. Duanmu (1999) adopts the idea of tone polarity, which requires a tone to be followed by an opposite tone. Therefore, neutral tone surfaces as L after H as in (4a), and H after L as in (4c). After LH (35) and HL(51), which already satisfy polarity, the neutral tone takes the default L.

3.3.2   Overview of Present Analysis

We argue against both parts of the previous analyses, i.e. default L plus H-insertion or tone polarity. Our proposals are laid out as follows. Instead of invoking H-insertion or tone polarity, we propose that the H target on neutral tone after tone 3 results from re-association of the complex contour tone in Mandarin, assuming the underlying form of tone 3 to be its citation form. Results from our phonetic experiment on multiple neutral tone syllables, to be reported next, contravene the idea of default L. When there is more than one neutral tone syllable, if the preceding lexical tone is not tone 3, the default pitch approach predicts that default L will be inserted on *all* neutral tone syllables, potentially giving rise to a sequence of L targets. F0 tracks corresponding to multiple neutral tone syllables clearly show only one L target aligned with the *last* neutral tone syllable. We attribute the L target to boundary intonation in Mandarin and propose that a L boundary tone is inserted on the last neutral tone syllable. The non-final neutral tone syllables get transitional pitch by interpolating from the preceding lexical tone to the L target on the last neutral tone syllable.

3.4   Deriving Pitch Targets of Neutral Tone

3.4.1   Experimental Setup

The experiment reported in chapter 2 also allows us to examine the phonetic pitch of multiple neutral tone syllables, and compare the idea of default L with boundary low insertion. The speech data were constructed such that the four lexical tones are followed by one, two or three neutral tone syllables in utterance-final and non-utterance-final positions. The other aspects of the experiment stay the same as described in chapter 2.

3.4.2   Deriving Pitch Targets of Neutral Tone

The most striking fact about f0 curves over multiple neutral tone syllables is the gradual f0 declination till the last neutral tone syllable. In the figure below, four lexical tones are followed by three neutral tone syllables in utterance-final position. The f0 curves were averaged over four repetitions for each tone sequence, produced by a male speaker recruited in the experiment. H, L, R and F label the relevant portion of f0 curves corresponding to four tones. N refers to neutral tone syllable. The initial rising f0 in H and F is due to the influence of preceding tonal context. It is observed that f0 curves over the three neutral tone syllables fall almost linearly to the end, indicating the presence of a low pitch target aligned with the last neutral tone syllable. The neutral tone immediately following tone 3 (L) is H, as we discussed before. The f0 curve begins to approach the low target from there. Notice that R has a much delayed f0 peak when followed by three neutral tone syllables.

(5) Mean f0 curves of four tones followed by three neutral tone syllables (speaker 3)
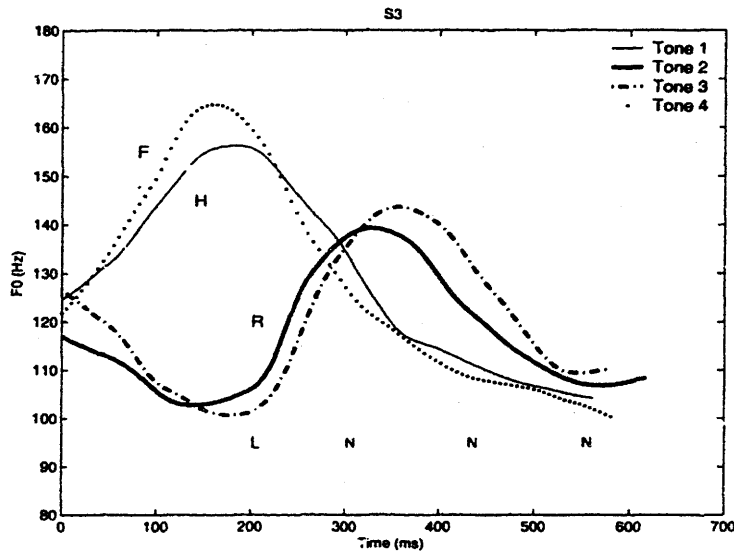
The observed f0 patterns for multiple neutral tone syllables can be derived quite straightforwardly by assuming that neutral tone syllables are toneless and boundary intonation is responsible for inserting a low boundary tone (L-) on the last neutral tone syllable. The falling f0 curves are simply f0 transitions from the ending f0 of the preceding lexical tone to the low target at the end. After tone 3, the f0 transition starts from the H target on the first neutral tone syllable. This idea is spelled out in (6). The non-final neutral tone syllables (N_1 and N_2) remain toneless in the output in (a). An additional step is required to re-associate the complex contour tone to the first two syllables in (b). As we will argue in chapter 6, the re-association is triggered by one of the contour licensing constraints, which excludes non-final complex contour tones.

(6) a. Deriving pitch targets of neutral tone after tone 2

$$\sigma \quad N_1 \quad N_2 \quad N_3 \quad \rightarrow \quad \sigma \quad N_1 \quad N_2 \quad N_3$$

b. Deriving pitch targets of neutral tone after tone 3

$$\sigma \quad N_1 \quad N_2 \quad N_3 \quad \rightarrow \quad \sigma \quad N_1 \quad N_2 \quad N_3 \quad \rightarrow \quad \sigma \quad N_1 \quad N_2 \quad N_3$$

L L H $\qquad\qquad$ L L H $\qquad\qquad$ L L H $\quad$ L-
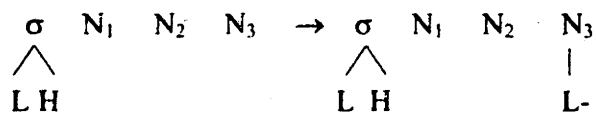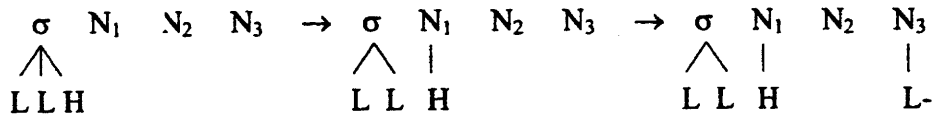
3.4.3 Evidence from F0 Peak Delay

A crucial difference between default L and boundary L analysis is that the former assigns a L target to each neutral tone while the latter only assigns a L target to the last neutral tone. Evidence from f0 peak delay in Mandarin supports the idea that the L target is on the last neutral tone syllable, but not on all neutral tone syllables. Recall that f0 peak delay in they Mandarin rise exhibits systematic variation as a function of the following context. When a rise is followed by one neutral tone syllable, its f0 peak is realized near the beginning of the following syllable rime. When it is followed by two or three neutral tone syllables, its f0 peak is found in the middle of the following syllable rime. Our explanation for peak delay is based on the proposal that all but the last neutral tone syllables are toneless, i.e. they have no pitch target. Therefore, when there is more than one neutral tone following a rise, the f0 peak in the rise is able to migrate into the immediately following syllable without creating a sharp fall.

If all neutral tone syllables are assigned a L target, the context-dependent variation of f0 peak delay in Mandarin would be inexplicable.

3.4.4 Boundary Intonation and Prosodic Domains

We have proposed that a low boundary tone is aligned with the last neutral tone syllable, but have not defined what prosodic boundary is involved. In the above figure, there is at

least an intonational phrase boundary after the last neutral tone since the three neutral tones are in utterance-final position. But apparently the relevant domain could be smaller than an intonational phrase. When neutral tone syllables occur in non-final position, we observed similar f0 patterns, namely, a L target on the last neutral tone syllable, accompanied by a f0 transition from the preceding lexical tone. The L target in non-final position is not scaled as low as that in final-position probably because of the final lowering effect (Liberman and Pierrehumbert 1984). This can be seen from figures of f0 tracks in section 2.3.3.2, chapter 2, in which tone 1 and tone 2 are followed by two or three neutral tones respectively in final and non-final positions. All the non-final target words in our data are followed by nouns in V-NP-V serial verb constructions.

Following the end-based theory of prosodic phrasing (Selkirk and Shen 1990, Selkirk 1995), we propose that the relevant prosodic domain is a prosodic word. According to Selkirk and Shen (1990), function words lose their lexically associated tones within a prosodic word. A prosodic word boundary is projected from the left edge of a lexical X^0. As we have shown, neutral tone syllables in Mandarin are always cliticized to the preceding full syllable in a lexical word. Our analysis of neutral tone alternation of function words in the next section shows that a prosodic word is the domain in which function words are de-stressed and then said in neutral tone.

3.5   Prosodic Word and De-stressing

3.5.1   Neutral Tone Alternation of Function Words

Some monosyllabic function words (pronouns and prepositions) are said in neutral tone
in some contexts, but in their underlying tones in others. Consider the examples below:

(7)   Subject /object pronouns

      a. **Ta** xihuan Zhangsan.   (tone 1)
         He likes Zhangsan

      b. Zhangsan xihuan **ta**.   (neutral tone)
         'Zhangsan likes him.'

(8)   Preverbal / postverbal PPs

      a. Zhangsan **zai** Beijing zhu.   (tone 4)
         Zhangsan in   Beijing live
         'Zhangsan lives in Beijing.'

      b. Zhangsan zhu **zai** Beijing.   (neutral tone)
         Zhangsan live in   Beijing
         'Zhangsan lives in Beijing'

The pronouns in (7) and prepositions in (8) have their own underlying tones, but we find

the following alternations: the pronoun *ta* is stressed in the subject position (7a), hence

said in its underlying tone, but is unstressed in the object position (7b), hence said in

neutral tone. By the same token, the preposition is stressed in preverbal PP (8a), but is

unstressed in postverbal PP (8b).

      The task of this section is to determine the prosodic domain within which function

words are de-stressed. Assuming that unstressed syllables lost their underlying tones in

the process of tone deletion (Cheng 1973), the phonetic pitch of unstressed function words will be determined by boundary intonation.

3.5.2 Prosodic Domain of De-stressing

The general approach we take to explain the neutral tone alternation is to figure out the domain within which a monosyllabic function word is de-stressed and hence said in neutral tone. There is a positional constraint on neutral tone syllables, i.e. they are not allowed in the initial position of a prosodic domain. That is why we always find the second syllable in neutral tone in disyllabic words. The stress pattern is trochaic for those disyllabic words. Consequently, when a stressed syllable occurs in the initial position of a prosodic domain, it cannot be de-stressed; in non-initial position, it can be de-stressed, and hence said in neutral tone.

Following Selkirk and Shen's (1990) theory of prosodic domains, we propose that the domain of de-stressing and neutral tone alternation is the prosodic word, defined as follows:

(9)   Prosodic word in Mandarin Chinese

{LEFT, X^0}, where X^0 stands for a lexical item.

We also formulate a de-stressing process, stated as follows:

(10)   De-stressing

Monosyllabic function words (i.e. pronouns and prepositions) are de-stressed

in non-initial position of a prosodic word.

Now we can explain the subject / object asymmetry. The syntactic structure and the corresponding prosodic words are shown.

(11) a. **Ta** [$_{VP}$ [$_V$ xihuan] [$_{NP}$[$_N$ Zhangsan]]] syntax
                (　　　　　(　　　　　　　　projection: {LEFT, X^0}
    ( ) (　　　) (　　　　) 　　　　　　prosodic word

b. [$_{NP}$[$_N$ Zhangsan]] [$_{VP}$ [$_V$ xihuan **ta**]] syntax
            (　　　　　　(　　　　　　projection: {LEFT, X^0}
        (　　　) (　　　) 　　　　　　prosodic word

The left boundary of a prosodic word is projected from the syntactic parsing first, and then right boundaries are filled in accordingly. There is no prosodic word boundary projected from the left edge of a pronoun in (a) and (b) because pronoun is not a lexical category. Note that the subject pronoun *ta* in (a) is parsed as a prosodic word by assuming exhaustive parsing of prosodic domains (Selkirk 1995). Since the object pronoun in (b) is the in the non-initial position, it is subject to the de-stressing process.

The asymmetry with respect to preverbal and postverbal PPs can be explained similarly with an additional assumption.

(12) a. [$_{VP}$[$_{PP}$ [$_P$ **zai**] [$_{NP}$ [$_N$ Beijing]]] [$_{VP}$ [$_V$ zhu]]] syntax
                (　　　　　　　(　　　　projection: {LEFT, X^0}
            (　　　) 　(　) 　　　　　prosodic word

b. [$_{VP}$ [$_V$ tiao] [$_{PP}$ [$_P$ **zai**] [$_{NP}$ [$_N$ Beijing]]]] syntax
        (　　　　　　　(　　　　　　projection: {LEFT, X^0}
    (　　　　　) (　　) 　　　　　　prosodic word

Consider (b) first. There is no prosodic boundary projected from the left edge of a preposition *zai* because it is not a lexical category. After automatic filling in the right boundaries, the preposition turns out to be in the non-initial position in a prosodic word, hence de-stressed and said in neutral tone. The preposition *zai* is unparsed in (a) after mapping the prosodic word boundary. Since it is not in the sentence-initial position, it will be parsed into a prosodic word with a preceding lexical item, e.g. the subject.

However, as Selkirk and Shen (1990) point out, a prosodic word boundary is still projected to the left of the pronoun in (a) due to the top-down effect imposed by prosodic structure wellformedness constraints. Assuming that a phonological phrase boundary is projected from the left edge of a lexical XP in Mandarin, there will be a phonological phrase boundary projected from the first VP in (a). A prosodic word boundary, although not projected from a lexical item, is imposed by the phonological phrase boundary, illustrated below.

(13) [vp[pp [p zai] [NP [N Beijing]]] [vp [v zhu]]]   syntax
      {   ((projection: {LEFT, X^0}
      (   )   (   )   (   )   prosodic word

The projection of XP is indicated by a curly bracket. Since the preverbal preverbal is now parsed as a prosodic word, it does not undergo the de-stressing process.

The prosodic words formulated above correctly predict the neutral tone alternation in sentences in which the pronoun is used as the possessor and the double object constructions in which the pronoun is the indirect object.

(14) Possessor pronoun

      a. [DP Ta [NP [v gege]]] [vp [v xihuan] [NP [N Zhangsan]]]   (tone 1)
          (   )   (   )   (   )   (   )
          his   brother   like   Zhangsan
          'His brother likes Zhangsan.'

      b. [NP [N Zhangsan]] [vp [v xihuan] [DP ta [NP [N gege]]]]   (neutral tone)
          (   )   (   )   (   )
          Zhangsan   like   his   brother
          'Zhangsan likes his brother.'

The possessor pronoun is parsed similarly as it occurs in subject and object position. It is worth mentioning that the possessor pronoun is the syntactic head of a DP, which is a

94

functional projection. Therefore, no prosodic word boundary will be projected from its left edge.

In double object constructions, when the indirect object is a pronoun, it is said in neutral tone. This is also predicted by our theory.

(15) Double object constructions

[NP [N Zhangsan]] [VP [V song] ta [NP [N qian]]]]   (neutral)
(   )   (   ) (   )
Zhangsan   give him   money
'Zhangsan gave him money.'

The pronoun *ta* is a function word. It is parsed into a prosodic word with the preceding verb.

3.6   Concluding Remarks

We have proposed that the pitch target of neutral tone syllables in Mandarin is derived by inserting a low boundary tone on the last neutral tone syllable. We also have provided evidence based on neutral tone alternation of monosyllabic function words that the relevant prosodic domain is a prosodic word. We argue against the default tone approach on the basis of f0 target alignment patterns observed in our phonetic experiment on neutral tone and f0 peak delay in Mandarin rise.

It is worth noting that many Mandarin dialects have neutral tone syllables; boundary tone insertion might not be the only choice for assigning a pitch target to neutral tone syllables. For example, as we will discuss in chapter 6, the pitch target of neutral tone syllables in Hai'an dialect is assigned either from contour re-association or spreading from the preceding level tone. Tone spreading also seems to be found in

neutral tone syllables in Wenzhou, a northern Wu dialect. In Xiamen, object pronouns are unstressed in the domain-final position. Their citation tones are supplanted by a default mid pitch or an immediately preceding tone (Hsiao 2002).

Nonetheless, boundary tone insertion is seen in a number of Sino-Tibetan languages and Chinese dialects, for example, Lhasa Tibetan (Meredith 1990, Duanmu 1992), and Tianjin (Jiang 1994). It offers a new perspective to look at some tone sandhi phenomena in Wu dialects as well, as we will demonstrate in chapter 4.

Chapter 4 Positional Prominence in Tone Mapping

4.1   Introduction

Certain positions are singled out by the phonology as being privileged. One of the diagnostics of positional privilege is the licensing of phonological contrasts in those positions, which are neutralized elsewhere (Beckman 1998). There have been two major approaches to licensing in Optimality Theory. Positional faithfulness theory (Selkirk 1994, Alderete 1995, Steriade 1995, Beckman 1998) demands identity in phonological mapping in prominent positions, thereby protecting those positions from alternation. Positional markedness theory (Lombardi 1995, Zoll 1996, 1998, Steriade 1997, Zhang 2001, among others) associates phonological markedness with a particular position (strong or weak). With respect to the main factor in licensing and neutralization of phonological contrasts, there are also two major approaches: licensing by cue (Steriade 1997) and licensing by prosody (Ito 1986, 1989, Lombardi 1991, 1995, 1999). In the licensing-by-cue approach, licensing or neutralization of a phonological contrast is controlled by the relative strength of the contrast-specific phonetic cues in a particular position (Steriade 1997, Zhang 2001), whereas in the licensing-by-prosody approach, it is controlled by the prosodic positions in which the relevant contrast is supposed to occur (Ito 1986, 1989, Lombardi 1991, 1995, 1999).

This chapter investigates positional prominence within the domain of tone. The major task is to identify prominent positions in tone mapping with particular reference to

Chinese dialects. A typology of tone-prominence interaction will be discussed. Specifically, we argue for the following points in light of a wide range of tone sandhi phenomena in Chinese dialects.

(a) Both head position and peripheral positions in a prosodic domain of tone sandhi have to be referred to by the phonology;

(b) Phonological patterns of tone retention and realization emerge from the interaction of different kinds of positional prominence.

4.2   Positional Prominence in Phonology

4.2.1   Prominent and Non-Prominent Positions in Phonology

Beckman (1998) distinguishes privileged positions and non-privileged positions right at the outset:

(1)   <u>Privileged positions</u>   <u>Non-privileged positions</u>

- Root-initial syllables	vs.	- Non-initial syllables
- Stressed syllables	vs.	- Unstressed syllables
- Syllable onsets	vs.	- Syllables codas
- Roots	vs.	- Affixes, clitics, function words
- Long vowels	vs.	- Short vowels

The privileged positions enjoy some perceptual advantage in the processing system, via either psycholinguistic or phonetic prominence, over the complement set of non-privileged positions (also Smith 2002, cf. Barnes 2001). In addition to the functional unity to those prominent positions, they also distinguish themselves from the non-prominent positions in the following closely related patterns of phonological asymmetry:

98

(2)   Positional asymmetries in phonology (Beckman 1998:1-2)

    a. Preservation of phonological contrasts in prominent positions, which are

       neutralized in non-prominent positions;

    b. Positional triggering of phonological processes in prominent positions;

    c. Positional resistance to processes in prominent positions, which apply in non-

       prominent positions.

We are primarily concerned with positional neutralization within the domain of tone. It

will be shown in this chapter that positional asymmetries also obtain in tone mapping, by

which tonal contrasts are preserved in certain positions, but neutralized elsewhere.

4.2.2   Multiple Prominence in Tone Mapping

Zoll (1997) argues for a three-way prominence typology, replacing the notion of uniform

accent.   Specifically, she distinguishes three kinds of prominent positions that are

relevant in tone mapping:

(3)   Prominent positions in tone mapping (Zoll 1997:98)

    a. *imposed*   METRICAL   penult, ante-penult, etc.
    b. *inherent*   PERIPHERAL   initial and final syllable in a domain
    c. *inherent*   ORGANIC   long vowels, vowels with high sonority, etc.

The proposal is motivated by several considerations. For example, there exist cases in

which the grammar crucially refers to the distinction between inherent and metrical

prominence, as exemplified by Safwa. In this Bantu language, peripheral and metrical

positions must be distinguished in order to explain the non-uniform behavior of different

noun classes with respect to tone mapping. It is true, as noted by Zoll (1997, fn 2), that

the three kinds of prominence are not always completely independent. Metrical prominence may coincide with inherent prominence such that a foot head may be attracted to a long vowel, thereby disturbing the binary iterative footing (Hayes 1995).

We will show below that there is evidence in Chinese tonal systems as well that single prominence is inadequate in explaining a wide range of tone sandhi phenomena; it even fails to capture some well-known facts of tone sandhi such as Shanghai. Instead, multiple prominent positions have to be admitted into the construction of grammar. In addition, phonological patterns of tone retention and realization emerge from the interaction of metrical prominence and peripheral prominence.

4.2.3   Approaches to Positional Prominence

Positional prominence, in association with positional neutralization, has received a great deal of attention in current debate concerning the interaction of phonetics and phonology in the grammar. It is widely acknowledged that certain positions enjoy some perceptual and phonetic advantages in the processing system. The current debate centers on the extent to which the phonetic information is necessary in the phonology. In some approaches, the positional prominence is largely stipulated as such (e.g. Beckman 1998, Zoll 1996). Although the perceptual and phonetic advantages are considered, they are not directly built into the phonology. In some other approaches, the perceptibility effect and phonetic detail are directly included in the phonology (e.g. Flemming 1995, Steriade 1997, 2001, Zhang 2001). With respect to contour tone distribution, Zhang (2001) shows that duration and sonority can be referred to directly in defining the contour licensing constraints, which are unified in their reference to positions of canonically longer

duration and higher sonority. For example, a syllable is a better licensor of a contour tone if it satisfies one of the following conditions: (1) it is stressed, (2) has a long rime, (3) in the final position of a prosodic domain, or (4) belongs to a short word. In terms of duration, final position patterns with the other cases because it is often subject to final lengthening (Oller 1973, Klatt 1975, 1978, Beckman and Edwards 1990, Edwards et al. 1990, Wightman et al. 1992). In contrast, domain-initial positions are not privileged for contour tone licensing since there is no clear evidence of initial lengthening in the literature (Zhang 2001:13).

We propose that although phonetic implementation factors could influence how phonology is shaped, as shown convincingly by Steriade (1997), the utilization of those factors is subject to language-specific phonologization. An example of this sort is the effect of f0 timing on contour re-association in many Wu dialects. We presented the phonetic timing patterns of the f0 target in Mandarin rising tone in chapter 2. Despite the fact that the durational patterns of disyllabic words are similar, the f0 peak delay found in Mandarin Chinese is a phonetic effect, whereas in Shanghai it is phonologized as contour re-association (cf. Duanmu 1999). Further discussion of this issue will be given in chapter 6.

4.3   Dual Prominence in Tone Sandhi: Stress and Edge

4.3.1   Prosodic Edges in Phonology and Phonetics

Prosodic edges (i.e. initial and final positions) in a prosodic domain have been shown to enjoy some kind of prominence in phonology and phonetics. For example, in metrical

systems, main stress is usually attracted toward some edge of a prosodic domain (Halle and Vergnaud 1987, Hayes 1995). In terms of segmental processes, vowel harmony for example, the syllable in root-initial position licenses more phonological contrast than the syllable in non-initial position and it also resists phonological alternation (Beckman 1998). In addition, initial syllable lengthening has been observed cross-linguistically as a specific case of a more general process known as initial strengthening (Barnes 2001, Byrd 2000, Dilley, Shattuck-Hufnagel and Osterndorf 1996, Fougeron 1999, Fougeron and Keating 1996, Keating, Cho, Fougeron, and Hsu 1999, Pierrehumbert and Talkin 1992, inter alia). It has been found that domain-initial consonants are realized with more linguopalatal contact and longer durations of VOT or consonantal closures in a number of languages (e.g. Korean, English, French and Taiwanese). The strengthened articulation is also sensitive to the level in the prosodic hierarchy: the higher the prosodic domain, the greater the effect of strengthening. Extending Clark's (1983) observation on contour tone distribution in African languages, Zhang (2001) reveals the extensive initial-final asymmetry in contour tone licensing in his survey of over 180 different tone languages. According to him, contour tone is more likely to be licensed on the final syllable than the initial syllable as a result of a cross-linguistic tendency of lengthening domain-final syllables. The final position enjoys the durational advantage to license a contour tone.

4.3.2　Single Prominence or Dual Prominence

Stress or metrical prominence[1] has been accorded a central role in explaining a wide range of tone sandhi phenomena in many Chinese dialects, especially tone preservation and tone deletion in Chinese Wu dialects. For example, it is assumed, explicitly or implicitly, in almost all analyses of Shanghai tones that in a compound the initial stressed syllable keeps its lexically associated tone, and the non-initial unstressed syllables lose their tones (Zee and Maddieson 1979, Yip 1980, 2002, Duanmu 1993, 1995, 1999, Chen 2000, cf. Selkirk and Shen 1990, Zhu 1995). Since the stressed syllable is the only prominent position that is referred to in the theory sketched above, I will call it the "single-prominence theory" hereafter, in contrast with the "dual-prominence theory", to which we turn now.

The dual-prominence theory that we propose is similar to the idea of multiple prominence in tone mapping reviewed earlier. In recognition of the fact that two types of prominent positions – head position and edge positions (i.e. initial and final) have been implicated for prominence-related tonal phenomenon in African languages (e.g. Zoll 1997), we further illustrate that phonological patterns of tone retention and realization emerge from the interaction of metrical prominence (associated with head position) and edge prominence (associated with peripheral positions) in a prosodic domain of tone sandhi. In other words, a tone could originate in one place, but surface in another. Our proposal is primarily based on tone movement or tone migration, a tonal process in which the underlying tone of a syllable at one edge of a tone sandhi domain is preserved, but

---

[1]　Zoll (1997) distinguishes between stress and metrical prominence. For our purpose, they are used interchangeably.

surfaces on the stressed syllable at the other edge in the same domain. Since stress is often attracted toward some edge of a prosodic domain, it is not always transparent as to whether the tone is retained for being stressed or being at the edge. When the tone is retained at one edge, but the word-level stress is at the other, tone movement results. The single-prominence model cannot handle tone movement in any obvious way.

4.3.3 A Tone-Prominence Typology

The dual-prominence theory predicts the interaction of different kinds of prominent positions in tone mapping. A somewhat simplistic typology of tone-prominence interaction can be generated by relating tone retention and tone realization to those positions respectively, as shown below.

(4)   A simplified tone-prominence typology

        a. tone retention and tone realization in different positions → tone movement

            - *initial* tone preservation, realization on *final stress* (Zhenhai)

            - *final* tone preservation, realization on *initial stress* (Wenzhou)

        b. tone retention and tone realization in same position → no tone movement

            - *initial* tone preservation, realization on *initial stress* (Shanghai, Suzhou)

            - *final* tone preservation, realization on *final stress* (Xiamen, Yantai)

        c. irrelevance of positional prominence → tone alternation

            - linear order of tones (Mandarin, Tianjin, Jinan)

Depending on where the tone is retained and where it is realized, tone movement occurs in (a). When there is no tone movement, as in (b), the stressed syllable keeps its lexically

associated tone and the preserved tone is realized on the stressed syllable. How the preserved tone is actually realized on the stressed syllable is further regulated by contour tone licensing constraints, to be discussed in chapter 6.

There exist cases in which tones migrate from stressed syllables to prosodic edges. What is going on in those cases is that stressed syllables are not able to carry their own underlying contour tones in a specific prosodic position (domain-initial for example) due to contour licensing constraints. Then the underlying contour of the stressed syllable seeks out another prosodic prominent position, resulting in tone movement. Such a pattern is instantiated in Shanghai when the initial stressed syllable in a compound is in tone 5 (one of the two checked tones in Shanghai).

There seem to be cases in which where the tone is retained and where it is realized have nothing to do with the position of stress. For example, in Shaoxing, a Wu dialect spoken in Zhejiang province, the initial tone of a tone sandhi domain is preserved and stays put there, but the final syllable is perceived to be stressed (i.e. longer duration) (F. Wang 1959, H. Wang 1999, personal communication). However, the final syllable (i.e. utterance-final in most tone sandhi literature) is subject to final lengthening. Without other independent evidence for final stress, we leave it as a questionable case. We are not aware of an example in which the final tone is preserved and stays put there, but stress is initial.

There also exist cases in which tone sandhi is not sensitive to either stress or prosodic edges. Instead, what matters is the linear order of tones. For two adjacent tones, the first, the second or both can be changed to other tones. Most Mandarin and Jin dialects fall into this category.

In what follows, we first review the major tone sandhi patterns in Chinese dialects and their relation to stress. Then we examine each of the typological patterns discussed earlier, starting with three cases of tone movement in Zhenhai, Shanghai and Wenzhou. We will show that reference to peripheral positions is required in explaining a wide range of tone sandhi phenomena, and that stress cannot be replaced by edge prominence. Instead, both are needed.

4.4   Stress and Edge in Tone Sandhi

4.4.1   Setting the Stage: Tone Sandhi and Stress

Lexical tones in Chinese dialects take certain phonetic shapes in isolation. Tone 1 in Mandarin Chinese is high level (transcribed as 55) and tone 2 is high rising (35). When tones come into contact with one another in connected speech, their phonetic shapes are often altered. This process of tone alternation is referred to as tone sandhi, The best-known example is the Mandarin third tone sandhi, in which a syllable in tone 3 changes to tone 2 before another syllable in tone 3 (Chao 1968).

Some dialects have extensive tone sandhi, as in Min and Wu dialects, while others have very few, as in Yue dialects (e.g. Cantonese). Tone sandhi takes places in a domain determined by metrical, prosodic or syntactic factors (Chen 1987, Duanmu 1995, 2000, J. Lin 1994, Selkirk and Shen 1990, Shih 1986, 1997, among others). In her survey of 83 Chinese dialects, Yue-Hashimoto (1980) classified tone sandhi into two major types.[2]

---

[2]   She also discussed a third type of tone sandhi which she called "local modification", but she noted that this type may be viewed as an elementary form of the last-syllable dominant type (pp. 459-460).

The first type of tone sandhi is called "first-syllable dominant", in which the first syllable in the tone sandhi domain maintains its lexically associated tone, while the following syllables assume different tonal values from their citation forms. Most Wu dialects, especially the northern Wu dialects belong to this type. The second type is called "last-syllable dominant", in which the last syllable in the domain keeps its underlying form, while the preceding syllables assume sandhi forms. Most of the Min dialects, southern Wu dialects and Mandarin dialects display this type of tone sandhi. Yue-Hashimoto also noted that the phenomenon of dominance is conditioned by stress. Therefore, the dialects that display the first-syllable or last-syllable dominance are sometimes referred to as left-prominent or right-prominent in the literature.

Given Yue-Hashimoto's (1980) classification, one of the major tasks confronting us is to unveil the linguistic principles which control the tone preservation and alternation. The prevailing view in the literature has been that tone sandhi is intimately related to stress. Almost all previous studies of Chinese phonology implicitly or explicitly assume that tone retention and formation of tone sandhi domains are tied to the metrical structure (Ao 1992, 1993, Chan and Ren 1989, Chan 1995, Chen 2000, Duanmu 1991, 1993, 1995, 1999, Yue-Hashimoto 1980, Kennedy 1953, Shih 1986, 1997, Wright 1983, Yip 1980, 2002). In this view, the syllable that retains its underlying tone is usually taken to be the head of the metrical constituent. All other syllables, which are taken to be non-heads, lose their lexically associated tones as in many Wu dialects, or take on a different allotone, as in many Mandarin and Min dialects. In addition, languages vary as to whether the relevant metrical constituent is head-initial or head-final. For example, Shanghai is argued to have metrically left-headed compounds (Duanmu 1993, 1995,

1999). When syllables are combined into a compound, the initial syllable tone survives on the surface whereas the non-initial syllable tones are suppressed. Xiamen (also called Amoy) is argued to have metrically right-headed compounds (Duanmu 1995). Consequently, the final syllable tone remains unaltered and non-final syllables use a different allotone.

Despite the importance of stress in tone sandhi, a number of studies have pointed out the elusive nature of stress both acoustically and perceptually in Chinese (Selkirk and Shen 1990, Duanmu 1995, 2000, Chen 2000). It is usually difficult to obtain consistent stress judgments. Often the native speakers simply do not feel stress. In practice, tonal preservation or resistance to tonal alternation/neutralization has been employed as the primary diagnostic of metrical prominence, especially in many Wu dialects (Yip 1980, 1999, 2002, Wright 1983, Duanmu 1993, 1995, 1999, Chen 2000, among others). The situation is aptly described by Yip (1999): "if one and only one syllable surfaces with its citation tone unchanged, while the others either lose or change their tone, the intact syllable will be taken to be the head."

Nevertheless, the existence of metrical structure in Chinese dialects has been strongly argued for on the basis of tone sandhi phenomena, most notably by Duanmu (1995, 2000) (see also Chen 2000: 285ff). Additionally, in recent years an increasing number of phonetic studies on tone and stress have become available and they provide further evidence for the presence of metrical prominence. Phonological and phonetic evidence will be presented as it relates to tone sandhi phenomena of the dialect under discussion.

In summary, what emerges from our review of the previous analyses of tone sandhi in Chinese dialects, particularly Wu dialects is that stress is the only relevant factor in determining tone preservation and alternation/neutralization in previous approaches. As we will demonstrate, any theory built on single prominence cannot handle tone movement in an obvious way. Conversely, it falls out naturally when both metrical prominence and edge prominence are invoked.

4.4.2   Tone Movement in Zhenhai

Zhenhai dialect is a northern Wu dialect spoken in a rural area located 20 km northeast of Ningbo in Zhejiang province. Rose (1990) describes its tonal system and disyllabic lexical tone sandhi on the basis of instrumental acoustic analysis. The trisyllabic tone sandhi is subsequently described in Rose (1994). All data are cited from these two papers. The Zhenhai data will be described in detail since they are not well-known to most phonologists.[3]

Rose (1990) shows that Zhenhai lexical tone sandhi represents an example in which the underlying tonal contour of the unstressed syllable is retained, but it is shifted to the stressed syllable. The shift of the preserved tone is unexpected, and also inexplicable, if tone preservation is only correlated with the position of metrical prominence. We argue that Zhenhai provides a crucial example that leads to the separation of tone retention and tone realization: where a tone is realized is not necessarily where it originates. Specifically, the initial syllable tone is always preserved,

---

[3]   To our knowledge, only Rose (1990) and Chen (2000) provide a phonological analysis for Zhenhai disyllabic tone sandhi based on the phonetic data in Rose (1990).

but it is realized on the syllable that carries metrical prominence, initial or final in disyllabic words, depending partly on tonal combinations.

We present the tone sandhi facts of Zhenhai to illustrate how stress and prosodic edge interact to derive the tone movement. A full OT analysis will be given in chapter 5.

4.4.2.1   Tones in Citation Form

Zhenhai has an inventory of six contrasting tones, summarized in the following table. The tone digits used in Rose are translated into a notation using H, M, and L.

(5)   Citation tones in Zhenhai[4]

		tone	example	tone value	notation	gloss
long tone	high register	1	tɕi	441	HL	"to fill"
		2	tɕiʔ	323	MH	"chicken"
	low register	3	tɕi̤	231	ML	"to ride"
		4	tɕi̤ʔ	213	LM	"he/she/it"
short tone	high register	5	tɕɪʔ	<u>5</u>	Hʔ	"knot"
	low register	6	tɕɛ̤ʔ	<u>23</u>	Lʔ	"straight"

Following Bao (1990, 1999) and Chen (2000), we assume that a tone (t) consists of tone register (r) and contour (c) in the following representation:

(6)   Tonal representation

$$
\begin{array}{c}
t \\
\wedge \\
r \quad c
\end{array}
$$
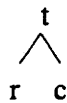
The r node specifies the pitch level of the tone. In all Wu dialects, lexical tones are classified according to high and low registers, corresponding to *yin* and *yang* registers in

---

[4]   The voiceless onset consonants with breathy voice are indicated by two dots underneath.

110

traditional terminology. The c node specifies the f0 movement of the tone and it can be static (high or low) or dynamic (rising or falling). Therefore, the six citation tones can be represented in (7):

(7) Autosegmental representations of Zhenhai citation tones

long tone				short tone	
HL	MH	ML	LM	H?	L?
tone 1 /\ r c | /\ h H L	tone 2 /\ r c | /\ h L H	tone 3 /\ r c | /\ l H L	tone 4 /\ r c | /\ l L H	tone 5 /\ r c | | h H	tone 6 /\ r c | | l L

For the ease of exposition, we will continue to use notations in H, M, and L unless register and contour have to be referred to separately. To avoid any potential confusion, when we talk about the underlying or surface tone of a syllable, we intend it to be interpreted as the union of register and contour.

As is clear from (5) and (7), a tone in Zhenhai can be characterized by three phonetic properties: duration, tone register and contour shape. Tones 1 to 4 co-occur with long syllables (CVV and CVN)[5]. Tones 5 and 6 co-occur with checked tone syllables (CV?), which correspond to MC syllables with consonantal endings. Their duration is much shorter than long syllables. Inspection of the duration of citation tones (equivalent to the duration of syllable rime) from the data reported in Rose (1990) reveals that the mean duration of long tones is on the order of 250 ms to 350 ms, while that of short tones is about 100 ms.

---

[5] Syllables in tones 2 and 4 do not develop from MC checked tone syllables, although they end with glottal stop in Zhenhai. This is a case of irregular correspondence, compared with neighboring Wu dialects.

111

The correlation between tone register and voiced/voiceless contrast in the syllable onset is still transparent in Zhenhai, like other Wu dialects. Each of the four MC tonal categories is split into *yin* and *yang* registers, conditioned by the voicing contrast in the onset. The *yin* register with a voiceless onset has a higher pitch value than the corresponding *yang* register with a voiced onset. The correspondence between the six citation tones in Zhenhai and the MC tonal categories is depicted below, like the one we saw in chapter 2:

(8)

register	tone			
	I *ping* "level"	II *shang* "rising"	III *qu* "departing"	IV *ru* "entering"
a. *yin*	Tone 1 (HL)	Tone 2 (MH)	Tone 1 (HL)	Tone 5 (H?)
b. *yang*	Tone 3 (ML)	Tone 4 (LM)	Tone 4 (LM)	Tone 6 (L?)

This six-tone system reflects two mergers in the development from the Middle Chinese tonal categories: tone 1 corresponds to MC Ia and IIIa, and tone 4 to MC IIb and IIIb. Tones 1, 2 and 5 in Zhenhai are in high register and have higher pitch values than the corresponding tones in low register (i.e. tones 3, 4 and 6). In addition, high-register tones appear in syllables with voiceless onset and low-register tones appear in syllables with voiceless onset with breathy voice. The onset with breathy voice becomes modally voiced when the syllable appears in non-initial position in a word, a typical situation in Wu dialects (Cao and Maddieson 1992).

Tone 1 (HL) and tone 3 (ML) are falling tones; tone 2 (MH) and tone 4 (LM) are rising tones. From the f0 tracks shown in figure 1 in Rose (1990), tone 1 is a high falling tone with its f0 peak positioned right at the beginning of the syllable rime. Tone 3 looks more like a convex tone with its f0 peak positioned in the middle of the syllable rime. It

is transcribed as /231/ in tone digits. We treat it as a low falling tone (ML) phonologically on the grounds of the following two considerations. Phonologically, tone 3 patterns with tone 1 in sandhi context, behaving exactly like a falling tone. Phonetically, the initial rising f0 portion is due to the effect of depressor consonants (Rose 1990, 2002). Tone 2 is phonologically a high rising tone (MH) and tone 4 a low rising tone (LM). It is noted that in both tones the rising f0 occurs in the later half of the syllable rime. The first half is covered by a low dipping f0. This kind of "late rise" in rising tones has been repeatedly observed in tone languages spoken in Asia and Africa, for example, Mandarin Chinese (Xu 1999, 2000) and Buli (Kenstowicz and Akanlig-Pare 2003).

It is sometimes difficult to determine the f0 shape for short tones because of their short duration. The task is further complicated by the glottal stop ending and its concomitant effect of vowel glottalization, which jointly give rise to irregular vocal fold vibration, often realized as an abrupt falling f0 at the end of the vowel. With the effect of glottalization filtered out, tone 5 looks like a short high tone and tone 6 a short low rising tone.[6] The realizations of tones 5 and 6 in sandhi context do not seem to provide unequivocal evidence to determine their phonological contour shape. We follow the tone digits given in Rose and treat tone 5 as Hʔ and tone 6 as Lʔ. The IPA symbol ʔ stands for glottal stop. It is used to indicate that H and L occur in short tones ending with a glottal stop.

---

[6] Rose (Personal Communication) mentioned that one of his speakers in Rose (1982) seems to have free variation with what sounded like a falling tone for tone 5, but it is actually difficult to be sure whether one is hearing a difference between level and falling pitch for the reasons we discussed.

4.4.2.2   Metrical Prominence and Initial-Position Prominence

Rose (1990) is primarily a study of disyllabic lexical tone sandhi, i.e. tone sandhi that applies in lexical items, mostly compounds. The location of stress plays a crucial role in determining the output sandhi patterns. The prominence judgments are confirmed by native speakers' intuition.

Disyllabic expressions fall into two metrical patterns: S-W (strong-weak) and W-S (weak-strong). The immediate question is what determines the metrical patterns. We start with the distribution of the six underlying tones.

The distribution is asymmetric in the disyllabic lexical items. Any lexical tone[7] can appear on the second syllable no matter whether it is labeled as stressed or unstressed, but there are restrictions on which lexical tone can occur on the first syllable.

(9)   Restrictions on underlying tone distribution in disyllabic expressions

S-W		W-S	
σ_1-σ_2		σ_1-σ_2	
tone 1 (Hr)		tone 1 (Hr)	
tone 4 (Lr)	Tone	tone 4 (Lr)	Tone
tone 2 (Hr)		tone 3 (Lr)	
tone 5 (Hr)		tone 6 (Lr)	

(Tone = any lexical tone, Hr = high register, Lr = low register)

The metrical pattern is derivable from the tonal category of the first syllable. A lexical compound will assume the S-W pattern if the first syllable is in tone 2 or tone 5 (both high register tone). Similarly, it will assume the W-S pattern if the first syllable is in tone 3 or tone 6 (both low register tone). When the first syllable carries tone 1 or tone 4,

---

7   Here we are talking about underlying tonal categories (tone 1 or tone 2), rather than their tonal values (H, L, R or F).

which metrical pattern will be used is determined by the historical sources of tone 1 and tone 4. Recall that tone 1 and tone 4 each corresponds to two MC tonal categories. According to Rose (1990), tone 1 derived from MC IIIa (*yin qu*) occurs in S-W pattern whereas tone 1 derived from MC Ia (*yin ping*) occurs in W-S pattern. By the same token, tone 4 derived from MC IIb (*yang shang*) occurs in S-W pattern and tone 4 derived from MC IIIb (*yang qu*) occurs in W-S pattern. Once the metrical pattern is ascertained, the surface sandhi forms of disyllabic compounds are predictable on the basis of the underlying tonal contour of the initial tone.

The lexical metrical patterns can be contravened by syntax. For example, unstressed tones 2 and 5 do not occur on the initial syllable in lexical tone sandhi, but they occur freely in verb-object and auxiliary verb-functional verb structures (Rose 1990:7).

Once the metrical patterns are clarified, we look for the acoustic cues associated with the metrical prominence. At the same time, we seek to decide whether there is evidence for edge prominence. We first examine duration.

Rose (1990) reports the average duration for the first and second syllables under all possible tonal combinations in the disyllabic sandhi contexts. The following four graphs, constructed from the data in Rose's table 4, illustrate systematic durational differences in the two syllables.

(10) Durational differences in long-long and short-short combinations
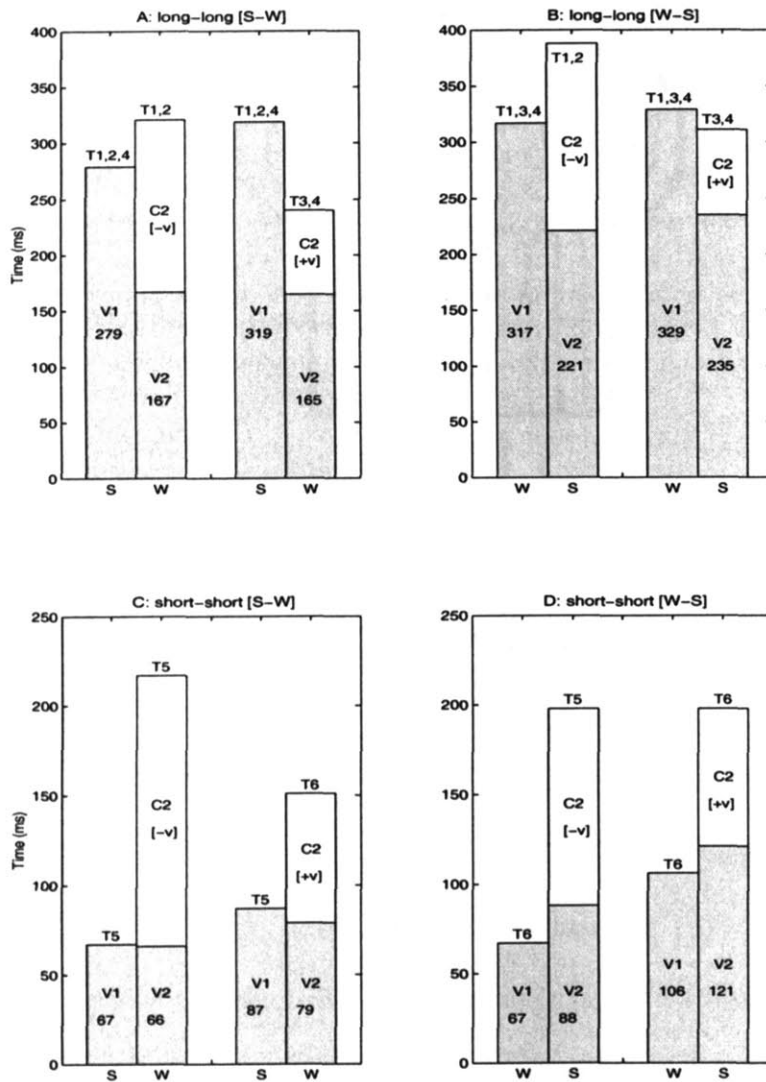
Figure 1: Graphs are plotted on the basis of the mean duration data presented in Rose (1990:10). The original data are re-tabulated and given in the Appendix. V1 and V2 stand for durations of the first-syllable vowel and second-syllable vowel, and C2 for the duration of the second-syllable onset. No duration data on the first-syllable onset is provided in the original data. The voicing of C2 is marked by [±v]: [+v] = "voiced" and [-v] = "voiceless". Voiceless C2 occurs with tones 1, 2, 5, and voiced C2 with tones 3, 4, 6. The numbers are durations (in ms) of V1 and V2.

(A) and (B) are long-long combinations and they differ in the metrical pattern. (C) and (D) are short-short combinations with different metrical patterns.[8] As an example, in graph A the first column represents the average duration (in ms) of the first-syllable vowels (V1) in tones 1, 2 and 4 when the following syllables bear tones 1 and 2. The second stacked column represents the average duration of the second-syllable consonants (C2) and vowels (V2) in tones 1 and 2. C2 is voiceless in this context. The third column represents the average duration of V1 in tones 1, 2 and 4 followed by syllables in tones 3 and 4. The fourth stacked column represents the average duration of C2 and V2 in tones 3 and 4. C2 is voiced in this context. The numbers inside each column are durations of V1 and V2 in different contexts. The metrical pattern is S-W in (A).

The distinction between long and short tones can be readily seen by comparing long-long combinations in (A) and (B) with short-short combinations in (C) and (D). Rose observes that as in citation form, short tones have between one quarter and one third the duration of the corresponding long tones (Rose 1990:21-22). He also finds that duration of the first-syllable vowel (in long and short tones alike) is generally longer before a voiced consonant (Rose 1990:18). The effect of C2 voicing on V1 duration can also be seen by comparing the two V1 columns in each graph.

The most striking fact that emerges from the long-long combinations is the lack of compelling correlation of metrical prominence (i.e. syllable in strong position in S-W or W-S patterns) with longer syllable duration.[9] Rather we observe that V1 is significantly

---

[8]   Long-short and short-long combinations are not plotted since they are unremarkable with respect to the conclusions we draw from long-long and short-short combinations.

[9]   Rose (1990:21) also points out that "stress does not correlate with duration in first-syllable tones".

longer than V2, which is the case in both S-W (A) and W-S (B) patterns. The difference between V1 and V2 ranges from 94 ms to 154 ms.

We propose that the duration difference between V1 and V2 (and accordingly, between the first syllable and the second syllable) is attributable to the fact that V1 occurs in the initial position of a prosodic domain under the assumption that a lexical compound in Zhenhai forms a prosodic word. If this is the right analysis, then we have witnessed two distinct kinds of prominent positions in Zhenhai: the strong positions in the S-W and W-S metrical patterns on the one hand, and the initial position of a prosodic domain on the other. We suggest that the former is associated with metrical prominence and the latter with initial-position prominence.

As we discussed earlier, initial syllable lengthening has been observed cross-linguistically as a specific case initial strengthening. In Zhenhai, the initial position in the disyllabic tone sandhi domain is made prominent by having extra duration. The initial-position prominence can be interpreted as arising from either initial syllable lengthening or non-initial syllable shortening. Both interpretations are compatible with the fact that V1 is prominent in the dimension of duration in comparison with V2, and further, this prominence is primarily associated with prosodic position, rather than the location of stress. As noted before, the mean duration for long tones in citation form is on the order of 250 ms to 350 ms. If we take the tone duration as equivalent to vowel duration, then V1 duration is comparable to the citation form. Assuming that speech sounds are articulated at a slower rate, hence with longer duration, in citation form than in connected speech, the durational difference between V1 and V2 in Zhenhai disyllabic forms is better characterized as V1 lengthening instead of V2 shortening.

The effect of initial syllable lengthening is not observed when the first syllables are short as in (10C) and (10D). This is because duration is used contrastively to distinguish long and short syllables. It is favorable to make long syllables longer in response to positional prominence, but lengthening short syllables would compromise the contrast in the dimension of syllable length, and therefore it is disfavored. We suggest that the initial position prominence is still present, but its effect on duration is overridden by a more important constraint on the contrastive use of syllable length.

There seems to be weaker evidence for stress-related lengthening in both long-long and short-short combinations. In each of the four graphs in (10), the stressed V2 is longer than the corresponding unstressed V2. However, the evidence is not as compelling as that for the initial syllable lengthening. First, this stress-induced lengthening is significantly smaller than the edge-induced lengthening. Second, with the exception of the first two columns in (10D), both V1 and V2 in the W-S pattern become longer than the corresponding V1 and V2 in the S-W pattern. With this co-variance effect filtered out,[10] the accentual lengthening is reduced to about 35 ms on average, whereas the structural lengthening is between 94 ms to 154 ms. Third, we many entertain another source for the lengthening of the stressed V2. The stressed V2 becomes longer probably because it carries a contour tone on the surface. The corresponding unstressed V2 only bears a level tone target. In terms of the phonetics of tone implementation, the lengthening could be induced by the durational properties of the contour tones (Gordon

---

[10] The adjusted accentual lengthening is computed in the following way: $L' = Lv2 - Lv1$, where $Lv2$ is the duration difference between the accented V2 and the corresponding unstressed V2, and $Lv1$ is the duration difference between the unstressed V1 and the corresponding accented V1. Take the first two columns in (6A) and (6B) for example, $Lv2 = 221 - 167 = 64$, $Lv1 = 317 - 279 = 38$, $\therefore L' = Lv2 - Lv1 = 64 - 38 = 26$.

1999, Zhang 2001). The stressed V1 also carries a contour tone on the surface, but its duration is no longer than the corresponding unstressed V1. We hypothesize that it is because the stressed V1 already has long duration to host the contour tones. Therefore, further lengthening is not necessary. Fourth, when both syllables are short, neither will carry contour tones on the surface. In (10C) and (10D), the stress-related lengthening is down to about 20 ms once the co-variance effect is filtered out.[11] Therefore, we conclude that the stress-related lengthening, if present, is very small compared to the initial syllable lengthening.

The initial-position prominence is further justified by initial tone preservation and resistance to register neutralization. As we will show in the next subsection, the initial tone is always preserved independent of where metrical prominence is. In W-S patterns, tone movement results. In addition, the tone register on the stressed syllable is high in W-S patterns. We take the high register as a manifestation of metrical prominence (i.e. stressed syllable realized in higher pitch) and assume that the stressed syllable is required to be in high register. When the stress is initial as in S-W patterns, the register of the initial syllable is faithfully maintained, overriding the effect of metrical prominence on tone register.

Therefore, the two prominent positions are licensed by different cues in Zhenhai. The initial position is cued by longer duration than the non-initial position, initial tone preservation and resistance to register neutralization. The stress position, initial in S-W and final in W-S, is cued by the presence of contrasting tonal contours in high register, as

---

[11] We assume that segmental compositions affect the syllable duration equally with respect to the two types of lengthening.

we will show in the next section. The following chart summarizes the relevant cues in the two prominent positions:

(11)

	S-W		W-S	
	σ_1	σ_2	σ_1	σ_2
initial prominence	yes	no	yes	no
	$d_1 > d_2$		$d_1 > d_2$	
metrical prominence	yes	no	no	yes
	contour	Φ	Φ	contour
register	original	low	original	high

(d = duration, Φ = tone neutralization, σ_1 and σ_2 = long)

We have presented evidence that both metrical prominence and initial-position prominence exist in Zhenhai. They are licensed by different acoustic cues. Short syllables do not exhibit the initial lengthening effect due to the contrastive use of duration to distinguish short from long syllables.

4.4.2.3   Tone Patterns in W-S Tone Sandhi

The W-S tone sandhi patterns are summarized in tabular form in (12), with transcriptions taken from Rose (1990:6). Examples follow in (13).

(12) W-S disyllabic patterns

σ_1	σ_2					
	T1:441	T2:323	T3:231	T4:213	T5:5	T6:23
A T1:441	33-441				33-4	
B T3:231	11-441				11-4	
C T4:213	11-334		11-24		11-4	
D T6:23	1-441	1-35	1-242	1-114	1-4	1-25

(13) A. "spring"   "western calendar"
         tshʏŋ thï   çi lï?
         441-441   441-23   citation tone
         33-441   33-4   sandhi tone

B. "coal mine"   "hair"
   mei khwɑ̃   tœy faʔ
   231-231   231-<u>5</u>   citation tone
   11-441   11-<u>4</u>   sandhi tone

C. "place"   "yesterday"
   ti fɑ̃   sɑ̃ ɲɪʔ
   213-441   213-<u>23</u>   citation tone
   11-334   11-<u>4</u>   sandhi tone

D. "tongue"   "special"
   çɛ tœy   ta pɛʔ
   <u>23</u>-231   <u>23</u>-<u>5</u>   citation tone
   <u>1</u>-242   <u>1</u>-<u>25</u>   sandhi tone

Recall that tones 2 and 5 do not occur in the first position in the input to W-S disyllabic tone sandhi: they show up in S-W tone sandhi, to be discussed later. Therefore, only four lexical tones occur on the first syllable. The second syllable may carry any of the six lexical tones. Two examples are provided for each tone sandhi pattern (marked as A, B, C, D in the table).

The disyllabic tone sandhi presents a case of dramatic positional neutralization in that the unstressed first syllable loses its lexically associated tone and contrasting tonal contours (rising and falling) only appear on the second stressed syllable. The first syllable takes on a level tone target, whose actual scaling is determined by its underlying tone register. For example, tone 1 on the first syllable becomes 33 in pattern A because it is an underlying high-register tone, while tone 3 becomes 11 in pattern B because it is an underlying low-register tone. When the first syllable is phonologically long, as in patterns A, B and C, the second syllable in the prominent position is not realizing its underlying tone target. Instead, as Rose (1990:28) proposes, the underlying tone target of

122

the first syllable is realized on the prominent, second syllable. When the first syllable is phonologically short, as in pattern D, the second syllable retains its underlying tone. When the second syllable is phonologically short, it gets a high tone target in all four patterns.

The tone register on the second syllable requires some comment. What happens in patterns A and B can be readily characterized as register neutralization, i.e. when the underlying tone of the initial syllable is realized on the second syllable, only the tonal contour (falling or rising) is copied, leaving behind whatever register it was originally associated with. The tonal contour is realized in the high register on the prominent, second syllable, regardless of its underlying register. Thus /231-213/ is realized as [11-441] rather than *[11-231]. The same is true for the short second syllable: its underlying specification for tone register is ignored and the tonal contour is realized in the high register.

The f0 tracks corresponding to the patterns A and B are given below. They are plotted on the basis of the data given in Rose (1990:10-14).[12] The second syllables are grouped according to voicing of C2 and syllable length. When C2 is voiced, the f0 curve in C2 is continuous (indicated by dotted lines in 14b and 14d); otherwise, it is broken (as in 14a and 14c). The stressed syllables are in italics.

---

[12]   The same, but differently displayed, f0 tracks can be found in Rose (1990:15-17).

(14) Pattern A: Tone 1 + Ti (i = 1/2, 3/4, 5, 6)

a. Tone1 + Tone1/Tone2
b. Tone1 + Tone3/Tone4
c. Tone1 + Tone5
d. Tone1 + Tone6

(15) Pattern B: Tone 3 + Ti (i = 1/2, 3/4, 5, 6)
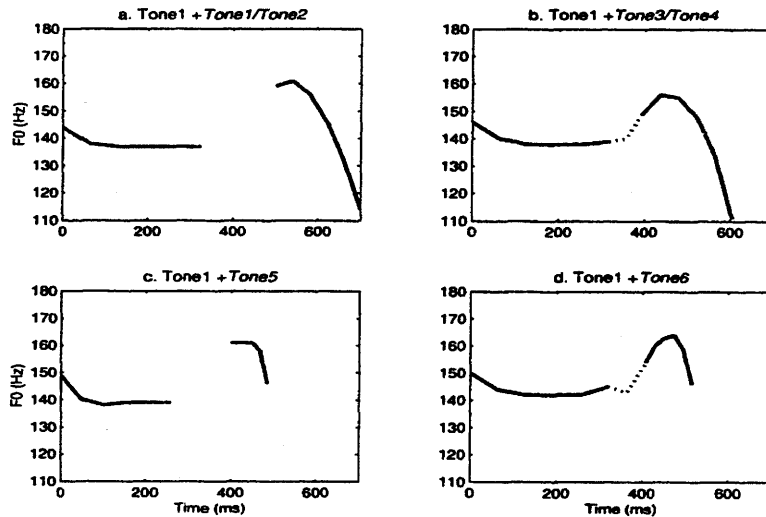
a. Tone3 + Tone1/Tone2
b. Tone3 + Tone3/Tone4
c. Tone3 + Tone5
d. Tone3 + Tone6
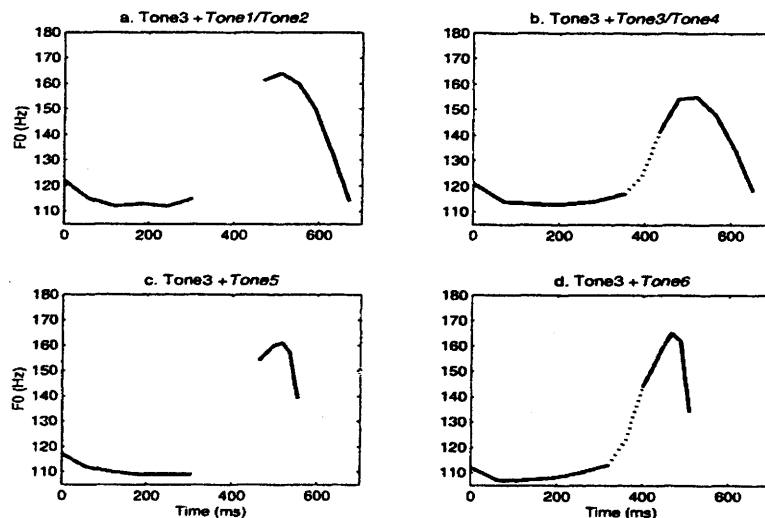
Patterns A and B describe the realization of underlying initial falling tones in high (A) and low registers. The two patterns are almost identical except that the level tone target is M in pattern A and L in pattern B. In both patterns, a high falling tone appears on the second long syllable (a, b). The f0 peak of the falling tone tends to be lowered

after the voiced intervocalic consonant which co-occurs with the low-register tones (i.e. tones 3, 4, and 6) in non-initial position in a lexical word. But this is more like an on-line phonetic effect, rather than a conditioning factor which is used by the phonology to determine how the tone register is specified. Following our analysis of Mandarin neutral tone, we propose that there is a low boundary tone associated with both edges of a prosodic domain and it is phonetically realized when the syllable at the edge is toneless. In (14) and (15), the low boundary tone is inserted on the initial syllable whose underlying tonal contour is realized elsewhere. Its actual realization (M vs. L) reflects the initial syllable's underlying tone register: high in pattern A and low in pattern B.

The f0 tracks corresponding to pattern C are shown in the following graphs.

(16) Pattern C: Tone 4 + Ti (i = 1/2, 3/4, 5, 6)
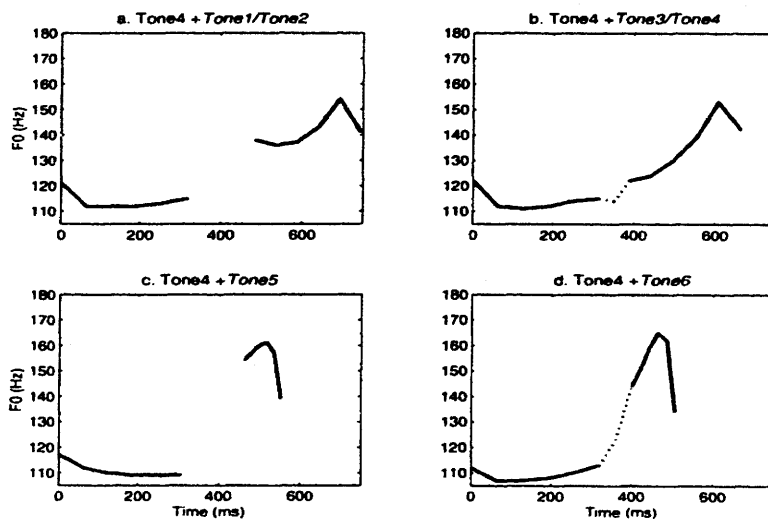
Pattern C describes the realization of underlying initial rising tone in low register. The (a) and (b) f0 tracks show a rise on the second syllable, contrasting with the fall seen in patterns A (14) and B (15). The initial syllable gets a low level tone after a low

boundary tone insertion. It is realized in the low register as a result of the initial syllable's underlying specification for tone register. A rising f0 contour, which originates on the initial syllable, appears on the second long syllable. The question is: what is the register specification for the rising tone on the second syllable? The rising contour is transcribed as [334] when the second syllable has underlying tones 1 and 2 (16a) and as [24] when the second syllable has underlying tones 3 and 4 (16b). Chen (2000:71) takes the different transcriptions as reflecting a real contrast in tone register, high for [334] and low for [24]. After comparing the f0 tracks, we suggest that the contrast is questionable. First, the rising f0 contours corresponding to the two transcriptions has the same f0 peak value. Second, the difference in the onset of the rising f0 is a function of the intervocalic consonant voicing that Rose (1990) discussed extensively: the voiceless consonant tends to raise the following f0 and the voiced consonant to lower the following f0. In Zhenhai, high-register tones (i.e. tones 1, 2 and 5) co-occur with intervocalic voiceless consonant and low-register tones (i.e. tones 3, 4 and 6) with intervocalic voiced consonant. Therefore, we follow Rose (1990)'s analysis and treat the rising tone after the initial tone 4 as a high-register tone on the long syllables.

Finally, we present the f0 tracks corresponding to the pattern D. Pattern D differs from the other three patterns in that the initial syllable is a short syllable. The second syllable retains its underlying tonal contour and is realized as in high register due to the metrical prominence. For example, the underlying tone on the second syllable in (17c) is a low falling tone (ML), but it is actually realized as a high falling tone in the prominent position. The same happens in (17d) where an underlying low rising tone (LM) is

actually realized as a high rising tone, despite the fact that its f0 scaling is lower than an underlying high rising tone (17b).

(17) Pattern D: Tone 6 + *Ti* (i = 1/2, 3/4, 5, 6)

a. Tone6 +*Tone1*   b. Tone6 +*Tone2*
c. Tone6 +*Tone3*   d. Tone6 +*Tone4*
e. Tone6 +*Tone5*   f. Tone6 +*Tone6*

Time (ms)   Time (ms)
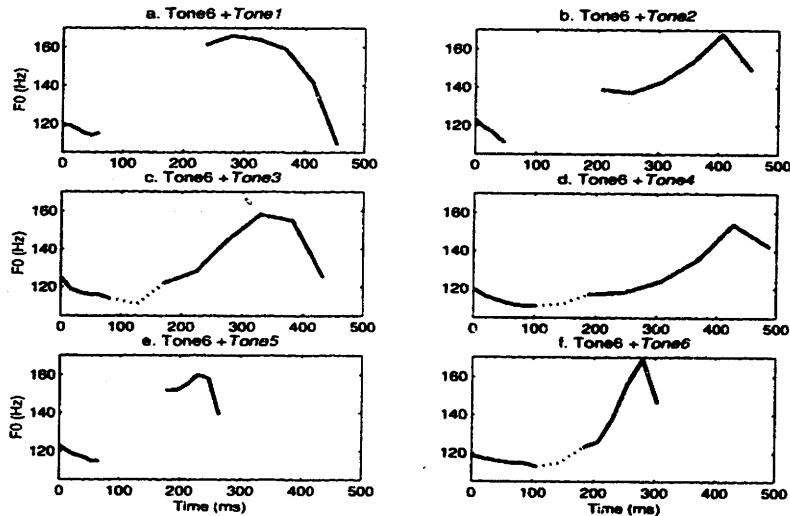
A high tone target appears on the short second syllable in all four patterns, with the final falling f0 discounted as arising from glottalization.[13] The short falling tails in (16a), (16b), (17b), and (17d) are also due to glottalization.

The tone mapping in the W-S disyllabic patterns is summarized in the following table. When the second syllable is underlyingly associated with a checked tone, the tonal output is found on the row marked with "-?".

---

[13] There is a relatively long rising f0 when the second syllable is in tone 6. We treat it as transitional here. However, we also note that the citation form of tone 6 is low rising. Admittedly, the evidence is not clear as to whether the rising f0 is transitional or tonal within such a short period of time.

(18) Tone mapping in W-S disyllabic tone sandhi

patterns	A: T1 + T	B: T3+ T	C: T4 + T	D: T6 + T
input	HL – T	ML – T	LM – T	L? – T
output	M – HL	L – HL	L – MH	L? – T^h
	-? M – H?	L – H?	L – H?	L – H?

(T = any tone, T^h = tone in high register)

From this table, the following generalizations can be derived:

(a) Contrasting tonal contours (rising and falling) occur on the second syllable when it is phonologically long;

(b) The tonal targets on the metrically prominent syllable are all realized in high register;

(c) Short syllables are neutralized to a high tone in the stress position;

(d) The tonal target of the initial tone is retained and shifted onto the metrically prominent syllable in long-long combinations;

(e) When the initial syllable is short, the second syllable retains its underlying tonal target;

(f) The initial syllable retains its underlying register specification;

(g) The initial syllable gets a low boundary tone and the underlying specification for tone register on the initial syllable determines how the low tone target is actually scaled.

4.4.2.4   Initial-Position Prominence and Final Stress: Tone Movement

There are two kinds of prominent positions co-existing in Zhenhai disyllabic tone sandhi. They are licensed by different phonetic cues: the initial-position prominence by duration, tone preservation and resistance to register neutralization, and metrical prominence by

high-register tonal contours. We have seen that the metrical prominence falls on the second syllable in W-S patterns, where the contrasting tonal contours emerge. The lexically associated tone of the initial syllable is always preserved in the output, but it does not surface in-situ. Its underlying tonal contour is displaced to the prominent, second syllable, where a high register is assigned, as proposed by Rose (1990). The underlying tone of the second syllable is completely lost, both contour and register. A boundary low tone is inserted on the initial syllable. The whole process is construed in autosegmental terms below:

(19)
$$\sigma_1 \quad \sigma_2 \;\to\; \sigma_1 \quad \sigma_2 \;\to\; \sigma_1 \quad \sigma_2$$

```
(19)   σ₁        σ₂    →     σ₁        σ₂    →      σ₁        σ₂
       /\        /\          /\        /\           /\        /\
      r₁ c₁    r₂ c₂        r₁ c₁ r₂ c₂           r₁  c     r  c₁
                                |____↑               |   |
                                                     L-   h
```
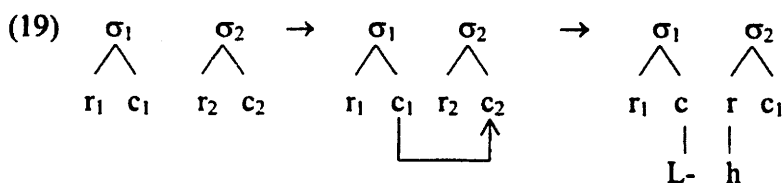
The underlying register r_1 is retained on σ_1 and it combines with the low boundary tone (L-) to give rise to the surface tone pattern: *L-* when r_1 is L, and *M-* when r_1 is H. σ_2 loses its underlying tone completely. The contour part of σ_1 (c_1) moves to σ_2 and it acquires a high register there.

The fact of initial tone preservation and its realization on the final stressed syllable calls for the two prominent positions to be referred to by the phonology. In the single-prominence theory, tone preservation and realization are determined by one parameter – stress: the stressed syllable retains its underlying tone and the retained tone is realized on the stressed syllable. Tone deletion is triggered by lack of stress. This theory predicts that there is no tonal system in which the underlying tones of unstressed syllables are retained. But this is exactly what we observe in Zhenhai.

In the dual-prominence theory, which tone is retained and where it is realized are determined by two different parameters: metrical prominence and edge prominence. Tone movement is predicted by the dual-prominence theory wh⌐n tone preservation and word-level stress occur in two different places. The single-prominence theory cannot handle tone movement in any obvious way.

4.4.2.5   Tone Patterns in S-W Tone Sandhi

The S-W tone sandhi patterns are summarized below, with transcriptions taken from Rose (1990:6):

(20) S-W disyllabic patterns

σ_1	σ_2					
	T1:441	T2:323	T3:231	T4:213	T5:$\underline{5}$	T6:$\underline{23}$
A T1:441	44-31				44-$\underline{32}$	
B T2:323	334-51				334-$\underline{5}$	
C T4:213	114-51				114-$\underline{5}$	
D T5:$\underline{5}$	$\underline{5}$-51				$\underline{5}$-$\underline{5}$	

Tones 3 and 6 do not occur on the first syllable as input to the S-W disyllabic tone sandhi. We end up with four patterns, marked as A, B, C, and D in the table. The second syllable may carry any of the six citation tones. Examples follow:

(21) A. "lake Tai"   "envelop"
         tha wu   çıŋ khɒʔ
         441-231   441-$\underline{5}$ citation tone
         44-31   44-$\underline{32}$ sandhi tone

      B. "scissors"   "arm"
         tçɪ tɑ   çju kwaʔ
         323-441   323-$\underline{5}$ citation tone
         334-31   334-$\underline{5}$ sandhi tone

      C. "wine"   "America"

130

```
         lɑ tɕju            mei kɒʔ
         213-323            213-5             citation tone
         114-51             114-5             sandhi tone


D.  "blackboard"           "to pick up"
         ha pɛ              tsha dzɒʔ
         5-323              5-23              citation tone
         5-51               5-5               sandhi tone
```

Like the W-S tone sandhi patterns, there is a dramatic positional neutralization of tones in that the lexically associated tones of the second syllable are completely gone, and the surface tone patterns are predictable from the first-syllable tone. Further, contrasting tonal contours (rising and falling) appear on the prominent initial syllable.

Before presenting a formal analysis, we have to determine the phonological forms we are dealing with, i.e. how the phonetic (or acoustic phonetic) data are interpreted phonologically. Interpreting phonetic data is difficult, since the mapping from underlying phonological structures to surface phonetic forms is far from straightforward. The observable phonetic events are always subject to constraints in speech production and perception. For example, observing a rising f0 curve in a syllable does not necessarily entail the presence of an underlying rising tone or pitch accent. A high tone or pitch accent may give rise to similar rising f0 curve when it is preceded by a low tone or pitch accent. Therefore it is necessary to distinguish between the f0 contour arising from one contour tone (e.g. rising or falling) in one syllable and the f0 transition between two different tones in adjacent syllables.

With the above digression in place, we now consider the surface disyllabic tone patterns, with the aid of f0 tracks plotted on the basis of the data given in Rose (1990:10-

14).[14] The f0 tracks corresponding to the underlying tonal sequences /Tone 1 + T/ are shown in (22). The second syllables are grouped according to voicing of C2 and syllable length. The stressed syllables are indicated in italics.

(22) *Tone 1* + Ti (i = 1/2, 3/4, 5, 6)
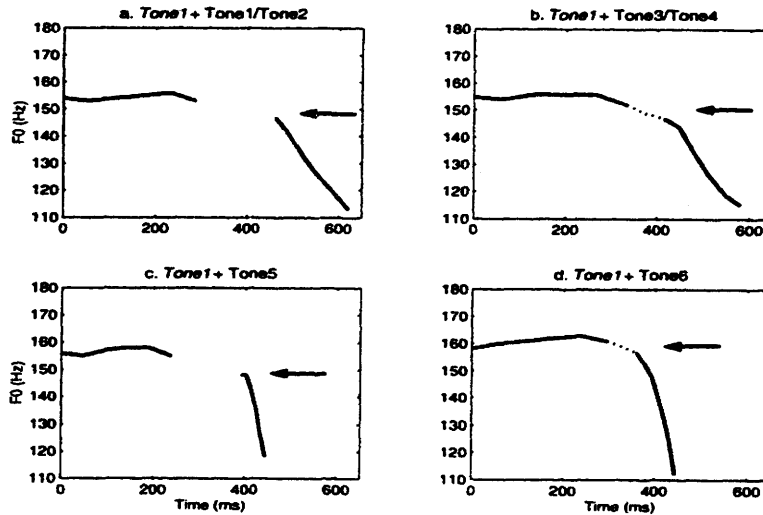
In each case, the first syllable carries a high level tone target, which is transcribed as [44]. Now the question is whether the second syllable carries a falling tone target or low tone target. From recent understandings of tonal target implementation (Xu 1997, 1999, Xu and Wang 2001, Xu and Sun 2002), the f0 peak in falling tone is anchored near the beginning of the syllable rime. When a falling tone is preceded by a high level tone (as in Mandarin tone 1 + tone 4), the f0 peak in the fall is not clearly identifiable, but it is clear that the falling part of f0 begins after the syllable onset. When a low tone is preceded by a high level tone (as in Mandarin tone 1 + tone 3), the high tone target and the low tone target are aligned with the end of their respective syllables, and therefore,

---

[14] Same f0 tracks can also be found in Rose (1990:15-17). They are plotted separately here.

there will be a falling f0 transition in the second syllable. However, the falling f0 transition actually starts before the end of the first syllable, where the high tone target is fully approximated. From the implementation perspective, the onset of the f0 curves on V2 in (21) (indicated by arrows) is lower than the ending f0 on V1. This indicates that the f0 there continues to move toward the low tone target at the end of V2. Consequently, there is no evidence that an f0 peak is planned at the beginning of V2, as there would be in the implementation of a falling tone. The falling f0 in (21) is thus the f0 transition from the high level tone target on the first syllable to the low tone target on the second syllable. We observed almost identical surface f0 patterns in Mandarin when a syllable in tone 1 (high level) is followed by an atonic syllable (i.e. neutral tone syllable). Therefore, we conclude that the underlying tonal sequence /HL + T/ in Zhenhai is mapped to /H + L/ in the wake of complete tone neutralization on the second syllable.

The same reasoning can be applied to the tonal sequences headed by tone 5 (H?). The f0 tracks follow in (23). Note that the low tone target in short syllables (23c, 23d) is not as low as that in long syllables (23a, 23b) because of the shorter durations of short syllables (i.e. checked-tone syllables).

(23) *Tone 5* + Ti (i = 1/2, 3/4, 5, 6)

a. *Tone5* + Tone1/Tone2

b. *Tone5* + Tone3/Tone4

c. *Tone5* + Tone5

d. *Tone5* + Tone6
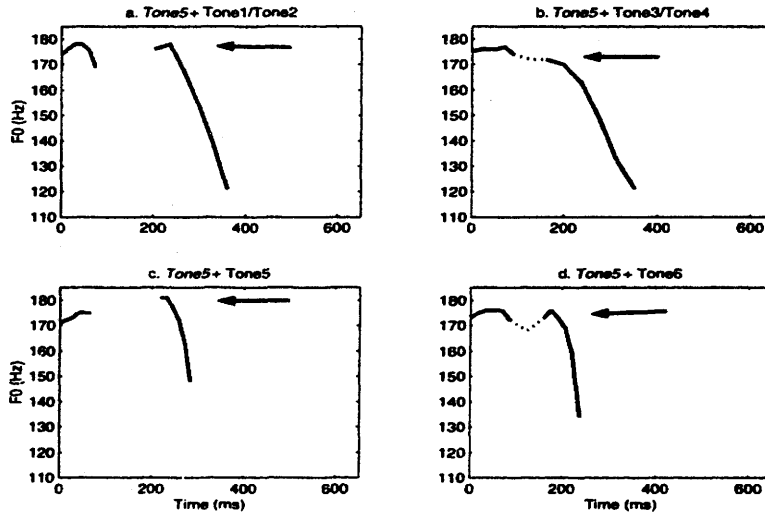
The tonal sequences which start with tone 2 and tone 4 give rise to slightly different f0 patterns on the second syllables. The corresponding f0 tracks are shown below:

(24) *Tone 2* + Ti (i = 1/2, 3/4, 5, 6)

a. *Tone2* + Tone1/Tone2

b. *Tone2* + Tone3/Tone4

c. *Tone2* + Tone5

d. *Tone2* + Tone6

134
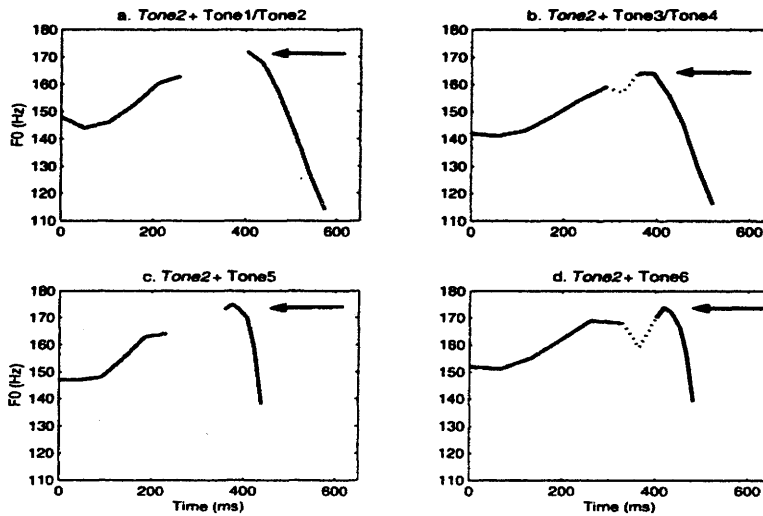
(25) *Tone 4* + Ti (i = 1/2, 3/4, 5, 6)
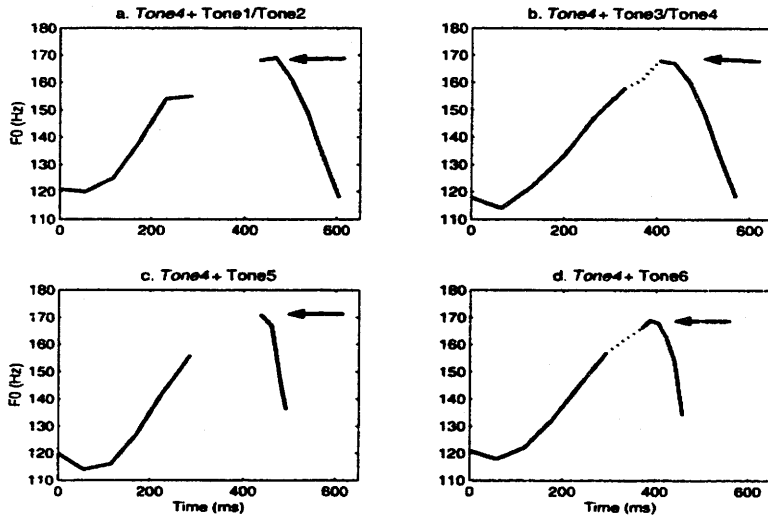
Comparison of the f0 tracks in (24) and (25) reveals that they are almost identical

except that tone 4 has a much lower f0 beginning than tone 2. The difference is expected

because tone 4 is an underlying low-register rising tone (i.e. LM). However, although the

tone register distinction is still preserved, the f0 patterns corresponding to tone 2 and tone

4 are both shifted up, i.e. realized at higher f0 values than their citation forms, in response

to the metrical prominence in the first syllable.[15] The f0 patterns on the second syllable

are different from those we saw in (22). They look more like implementation of a falling

tone because the f0 hits its maximum at the beginning of V2 (indicated by arrows).

However, since the preceding tone is a rising tone, the f0 peak here is very likely to be a

delayed f0 peak belonging to the preceding rise. In Mandarin Chinese, the f0 peak delay

is far more likely to happen in a rising tone than a level high tone (Xu 2001) and there is

more delay when the following context is an unstressed syllable instead of a stressed

---

[15] The difference is on the order of 10 to 20 Hz, as judged from the Figure 5 in Rose (1990:21).

syllable, as we have shown in chapter 2. The resemblance is striking between the f0 patterns corresponding to the underlying S-W tonal sequences /323 + T/ and /213 + T/ in Zhenhai and those corresponding to /tone 2 + neutral tone/ (/35 + N/) in Mandarin Chinese where a low tone target is realized on the neutral tone syllable. Positing a low tone target in the second syllable, everything else will follow from tonal implementation. When the second syllables are short, the falling f0 transition is truncated (as in 22c, 22d), resulting in a higher low tone target.

The mapping from the underlying tone sequences to the surface tone patterns in disyllabic tone sandhi is summarized in the following table, where T stands for any of the six citation tones:

(26) Tone mapping in disyllabic tone sandhi

patterns	A: T1 + T	B: T2+ T	C: T4 + T	D: T5 + T
input	HL – T	MH – T	LM – T	H? – T
output	H – L	MH – L	LM – L	H? – L

Except for pattern A, the initial tone is faithfully preserved in the output. The second syllable loses its underlying tone completely. It takes on a low boundary tone, as we proposed for the W-S patterns. The underlying specification for tone register is preserved on the first syllable; it is neutralized to low register on the second syllable.

4.4.2.6   Initial-Position Prominence and Initial Stress: S-W Patterns

In the S-W patterns, the initial-position prominence and metrical prominence coincide. As we mentioned earlier, the initial syllable tone is attracted to the stressed syllable in Zhenhai. When stress is initial, the initial tone surfaces on the initial syllable. The

single-prominence theory is able to explain this pattern by saying that the stressed syllable (i.e. initial in this case) keeps its lexically associated tone. However, as we already illustrate with the W-S patterns, the initial tone preservation and tone attraction to the stressed syllable have to be separated.

The S-W tone sandhi further illustrate the effects of initial-position prominence by preserving the underlying tone register of the initial syllable, overriding neutralization to high register under stress. The tonal target on the non-initial unstressed syllable is neutralized to low register because of lack of stress.

4.4.2.7   Trisyllabic Lexical Tone Sandhi

The trisyllabic lexical tone sandhi patterns in Zhenhai provide further support for the interaction of prosodic edge and stress. We will present them in the next chapter as we give a full OT analysis. To preview the results, as described in Rose (1990) and (1994), the initial tone is always preserved, but it surfaces on initial, middle or final syllable, depending on where stress falls.

4.4.2.8   Summary

The lexical tone sandhi in Zhenhai exhibits the interaction of initial-position prominence and metrical prominence in a number of ways. We found that the initial syllable has longer duration than the non-initial syllable even when it carries stress. The initial tone is always preserved, independent of where stress falls. The initial syllable also preserves its underlying register specification, overriding neutralization to high register under stress

137

and to low register under no stress. The stressed syllable attracts the preserved initial

tone and neutralizes its register to high if the stress is not initial.

4.4.3   Tone Preservation and Initial Stress in Shanghai

4.4.3.1   Citation Tones in Shanghai

Shanghai is probably the best-known Wu dialect in the phonological literature. It has two

major varieties: Old Shanghai and New Shanghai (also called Mainstream Shanghai).

Old Shanghai has more citation tones and more complex tone sandhi patterns; it is spoken

by a small number of old people (Chao 1928, Shen 1981a, 1981b, 1982, Xu et al. 1981,

Xu et al. 1988). New Shanghai is the variety spoken by the majority of people in

Shanghai. Since our focus is New Shanghai, it is referred to as Shanghai hereafter.

Shanghai has five lexical tones in citation form, shown below:

(27) Citation tones in Shanghai

		tone	example	tone value[16]	notation	gloss
long tone	high register	1	sa	52	HL	"to sift"
		2	sa	34	MH	"what"
	low register	3	za	13	LM	"firewood"
short tone	high register	4	sa?	5	H?	"to kill"
	low register	5	za?	23	LM?	"stone"

Like Zhenhai, Shanghai exhibits the correlation between tone register and

voiced/voiceless contrast in the onset: voiced onsets in low register and voiceless onsets

---

[16] The transcriptions of the tones in Shanghai differ slightly in different sources. For example, tone 1 is
transcribed as 51 or 52 by different authors. Tone 2 is transcribed as 34 or 24. Nonetheless, the same
contour shape can be inferred. For a discussion of some limitations regarding the use of Chao's tone
letters, the reader is referred to Duanmu (1990:103ff).

in high registers. In terms of contour shape, there is one falling tone and one rising tone. The rising tone occurs on four syllable types, differing in register and length. Therefore, its citation realization in terms of register and length is predictable from the host syllable. Tone 4 is realized as a short high level tone. It is considered a phonological rise because it patterns with other underlying rising tones in tone sandhi, to which we turn next.

4.4.3.2   Tone Sandhi in Shanghai

Tone sandhi in Shanghai is typical of the left-dominant tone systems in Yue-Hashimoto's taxonomy. Only the initial tone matters in predicting the surface tone pattern of the whole compound. The underlying tones of the non-initial syllables have no effect at all and they are completely neutralized. The following table, adapted from N. Qian (1992:619), summarizes the tone patterns of compounds consisting of two to five syllables. The initial syllable carries one of the five lexical tones. Checked tones are underlined. As can be seen from this table, the underlying register specifications of non-initial syllables have no effect on the surface tone patterns.

(28)   Complete tone sandhi patterns in Shanghai[17]

initial tone	σσ	σσσ	σσσσ	σσσσσ
tone 1: 52	55-31	55-33-31	55-33-33-31	55-33-33-33-31
tone 2: 34	33-44	33-55-31	33-55-33-31	33-55-33-33-31
tone 3: 13	22-44	22-55-31	22-55-33-31	22-55-33-33-31
tone 4: <u>55</u>	<u>33</u>-44	<u>33</u>-55-31	<u>33</u>-55-33-31	<u>33</u>-55-33-33-31
tone 5: <u>23</u>	<u>22</u>-23	<u>22</u>-22-23	<u>22</u>-55-33-31 <u>22</u>-22-22-23	<u>22</u>-55-33-33-31

---

[17]   N. Qian (1992) transcribes tone 2 as 334 and tone 2 as 113. We change them to 34 and 13 to be consistent with other transcriptions, for example, Xu et al. (1981) and Xu et al. (1988).

It has been claimed that tone sandhi in Shanghai and also in other northern Wu

dialects behaves like African languages in terms of tone mapping (Yip 1980, 1989,

Duanmu 1993, 1994, among others). The retained lexical tone of the initial syllable is

mapped to the segmental string from left to right in one-to-one fashion (cf. Goldsmith

1976, Pulleyblank 1986), of which the following examples are illustrative.

(29) a. Tone 1 + T (T=any lexical tone)
           HL + T → H + L
           "weather"   "thin"
           thi tshi   tæ bu?
           HL-MH   HL-LM?   citation tone
           H-L   H-L   sandhi tone

      b. Tone 2 + T
           MH + T → M + H
           "anger"   "to requite"
           hut shi   pø ta?   citation tone
           MH-MH   MH-LM?   sandhi tone
           M-H   M-H

In (a), the falling tone of the initial syllable spreads over the two syllables of the

compound, H on the first and L on the second. Similarly, the high rising tone splits into

M and H over the two syllables in (b). Any underlying tone can occur in the second

syllable. However, it is completely neutralized, and the second syllable appears with L or

H, depending on the underlying tone of the initial syllable. Facts like these strongly

support the view that a contour tone is composed of a sequence of level tones (Woo 1969,

Leben 1973, Goldsmith 1976, Yip 1980, among others).

One might note that in the table above, the tonal combination of /tone 1 + T/ is

transcribed as 55-31. The question is whether the second tone is a simple L target or a

low fall target. This situation is similar to what we observed in Zhenhai. Based on what

140

we know about tonal co-articulation and tonal target alignment, it is very likely that the second tone only has a L target. A L target following a H target will give rise to a falling f0. Such an interpretation is compatible with our auditory impression and also the fact that /tone 1 + T/ is transcribed as 55-21 by other authors (Xu et al. 1981, Xu et al. 1988). The f0 tracks below, plotted on the basis of four speakers' (three males and one female) data reported in Zhu (1995), clearly exhibit the f0 patterns corresponding to H target on σ_1 and L target on σ_2, with a falling f0 transition in between. The f0 values for /tone 1 + tone 1/ and /tone 1 + tone 2/ are pooled together in the original data in Zhu (1995). The two syllables' duration (only rime duration is shown) is normalized by a 2:1 ratio so that the first syllable rime is twice as long as the second. The reason will become clear when we discuss the phonetic evidence of stress in Shanghai. Putting aside short syllables, Zhu (1995) shows that the first syllable rime is on the order of 200 ms, whereas the second is about 100 ms. The two syllables are separated by a silence period of 50 ms in the f0 plot, which has nothing to do with the actual duration of the second syllable onset.

(30) Tone 1 + Tone1/2

Tone1 + Tone1/2
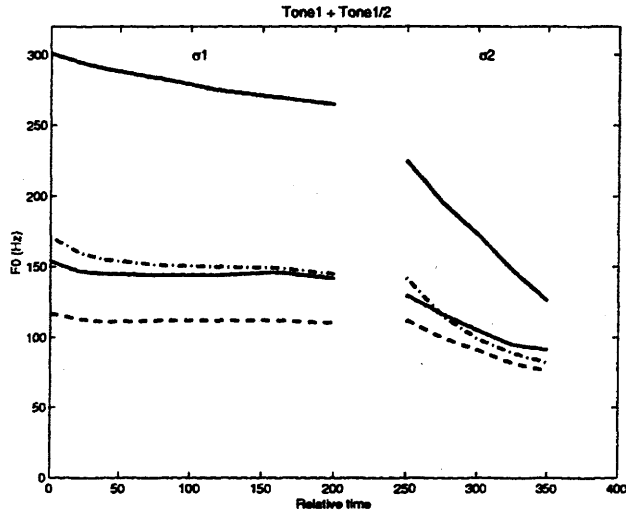
In longer compounds, the initial syllable tone is mapped to the first two syllables in the way described above. The remaining syllables are said to get a default L tone in most analyses (for example, Selkirk and Shen 1990). The mapping is illustrated below:

(31) MH + T + T → M + H + L
      "accordion"
      sou fon dzin
      MH-HL-LM   citation tone
      M-H-o   sandhi tone
      M-H-L   default-L insertion

Following our analysis of Mandarin neutral tone, we propose that the last syllable gets a boundary L- tone in a compound longer than two syllables. Other syllables outside the first two syllable window are simply toneless and their pitch is transitional. Therefore, we end up with the following mapping:

(32) σ_1 σ_2 σ_3 σ_4 σ_1 σ_2 σ_3 σ_4 σ_1 σ_2 σ_3 σ_4

      $\wedge$ $\wedge$ $\wedge$ $\wedge$ → | | | | → | | | |
      M H M H M H M H   M   H   M   H   L-

142

The analysis sketched so far does not seem to work for the compounds in which the initial syllable is in tone 5. Tone 5 is low rising (LM) in citation form. It co-occurs with short checked-tone syllables. In disyllabic compounds, redistributing its tonal contour would yield L on the first syllable and M on the second if the mapping proceeds from left to right in one-to-one fashion. However, this is incorrect. The tone patterns involving initial tone 5 are repeated below:

(33) Tone 5 as the initial tone

initial tone	σσ	σσσ	σσσσ	σσσσσ
tone 5: 23	22-23	22-22-23	22-55-33-31 22-22-22-23	22-55-33-33-31

In other transcriptions of this pattern, tone 5 is described as 13 in isolation, and the disyllabic pattern is 11-23 (Xu et al. 1981, Xu et al. 1988). What happens in this less frequently discussed pattern is similar to Zhenhai in that the initial tonal contour (LM) moves to the final syllable in disyllabic or trisyllabic compounds. The initial syllable ends up with a L tone. Quadrisyllabic and quintisyllabic compounds are rare. In the former, two options are available: spread LM onto the first two syllables or moving LM as a whole to the final syllable. In the former the initial tone redistributes to the first two syllables.

One might ask whether the second syllable in this pattern actually has a M target instead of LM by attributing the rise to the artifact of an f0 transition. The answer is no given the following three reasons. First, it is perceptually a distinct rise. Second, acoustically the f0 curve shows a delayed rise, rather than an immediate rise which is characteristic of the f0 transition between L and H pitch targets. Third, the variations

observed in the second syllable duration in different tonal contexts are compatible with a rising tone. Zhu (1995), in a systematic phonetic study on Shanghai tones using multiple speakers, reports that in a disyllabic word or compound, the longest rime duration is found when the second syllable is after a syllable in tone 5. The following table, taken from Zhu (1995), summarizes the rime duration of syllables in tone 1/2 (pooled together in the original data) and tone 3 in different tonal contexts. The first number in each cell is the rime duration of the first syllable (in tones 1, 2, 3, 4, and 5). The second number (shaded) is the rime duration of the second syllable (in tone 1/2 and tone 3 respectively). The tonal pattern for each combination is also provided.

(34) Rime duration of syllables in tone 1/2 and tone 3 (in ms)

σ_2	T1-σ_2	T2-σ_2	T3-σ_2	T4-σ_2	T5-σ_2
	H-L	M-H	L-M	M-H	L-LM
T1/2	158-73	174-121	172-122	55-139	57-169
T3	187-81	191-135	202-129	77-161	65-193

This table shows the syllable after tone 5 has the longest duration, and the syllable after tone 1 has the shortest duration. The longer duration is expected when the syllable after tone 5 carries a rising tone. Studies on the phonetic properties of contour tone (Ohaha and Ewan 1973, Sundberg 1979, Xu and Sun 2002) show that a rising tone takes longer time to implement than a falling tone. Recent studies on contour tone distribution also reveal that a rising tone is more marked than a falling tone and it tends to occur in a syllable with longer duration and high sonority (Gordon 1999, Zhang 2001).

The tone sandhi in Shanghai can be summarized as follows. First, the initial syllable tone is always preserved while the non-initial syllables delete their lexically associated tones. Second, the initial tone redistributes over the first two syllables when it

is not tone 5. Third, when tone 5 is initial, its contour moves to the final syllable in a disyllabic or trisyllabic compound. In Zhenhai, we argued that tone movement is triggered by attraction of the retained tone to the stressed syllable. Does the tone movement in Shanghai also result from stress attraction? Before we answer these questions, we want to understand whether the initial tone preservation in Shanghai is due to initial stress or initial-position prominence unrelated to stress. Therefore, we review phonological and phonetic evidence of stress in Shanghai.

4.4.3.3 Phonological Evidence of Stress

The tone preservation on the initial syllable and tone deletion on non-initial syllables have been argued to correlate with initial stress or left-headed metrical structure in Shanghai in many current analyses, most notably Duanmu (1993, 1995, 1999). Under this view, only a stressed syllable can retain its lexically associated tone whereas tones of the unstressed syllables do not surface.

In addition to tonal distribution, Duanmu (1995) argues for left-headed stress in Shanghai by drawing on evidence from word length restrictions and focus. He notes that in a trisyllabic compound consisting of two lexical words (W1 and W2), either one or two tone sandhi domains can be formed, depending on the length of W1 and W2. In 2+1 structure (i.e. W1 disyllabic, W2 monosyllabic), both options are permissible. In 1+2 structure (i.e. W1 monosyllabic, W2 disyllabic), only one domain can be formed. The following examples are adapted from Duanmu (1995:23):

(35) a. 1+2

      $[\sigma]_{w1}[\sigma\sigma]_{w2}$ word structure

      $(\sigma\sigma\sigma)$ but $*(\sigma)(\sigma\sigma)$ metrical structure

b. 2+1

$[\sigma\sigma]_{w1}[\sigma]_{w2}$ word structure

$(\sigma\sigma\sigma)$ or $(\sigma\sigma)(\sigma)$ metrical structure

The tone sandhi follows from the mapping process discussed above. One further detail is that when the monosyllabic domain is phrase-final, the underlying contour tone surfaces as such.

The word length effect on domain formation follows from Duanmu's metrical analysis by assuming that word stress and compound stress are both left-headed in Shanghai. The unattested metrical structure in (a) is ruled out because it creates a stress clash, resolved by removing the weaker stress. There is no stress clash in (b), and deletion of weaker stress is optional.

(36)   a. stress clash

```
        x           x
      x   x         x
    *(σ)(σ σ) → (σ σ σ)
```

b. no stress clash

```
      x             x
    x   x           x
   (σ σ)(σ)   or   (σ σ σ)
```

The Shanghai tone sandhi is also sensitive to contrastive focus, whose primary effect is to insert a prosodic boundary to the left of the focused constituent and optionally delete the post-focus tones (Selkirk and Shen 1990, Duanmu 1995), pointing to left-headed stress. When a stress clash results from contrastive focus, it is resolved by removing the stress on the right, as in (a).

The word length restrictions and contrastive focus furnish independent evidence for the existence of metrical structure in general, and the left-headed stress in Shanghai in particular.

4.4.3.4   Phonetic Evidence of Stress

The nature of phonetic stress in Chinese is generally elusive (Chen 2000). Shanghai seems to be no exception. Although tonal distribution and other phonological evidence point to left-headed stress in Shanghai, Selkirk and Shen (1990) note that "native speaker intuition does not support the contention that leftmost syllable in compounds are rhythmically more prominent". However, recent phonetic studies did reveal effects of stress on syllable duration. We have already cited Zhu's (1995) detailed study of Shanghai tones. According to his measurements, in a disyllabic word or compound the rime duration of the first syllable is much longer than that of the second syllable. Ignoring syllables in short tones (tones 4 and 5), the average rime duration of the first syllable is about 200 ms and that of the second syllable is about 100 ms (see the table in 34). Such a disyllabic duration pattern is pretty close to the duration pattern of a stressed syllable followed by an unstressed syllable in Mandarin Chinese, as reported in Lin and Yan (1980, 1988). Although the disyllabic duration pattern in Shanghai seems to support the left-headed metrical analysis, it poses another problem. As pointed out by Duanmu (1999), if Mandarin and Shanghai have identical duration patterns, why is there contour redistribution in Shanghai, but not in Mandarin? We will address this issue in chapter 6, where we compare Duanmu's solution with ours.

147

4.4.3.5 Tone Movement: Effect of Stress or Contour Licensing Restriction?

There are two conceivable approaches to tone movement in Shanghai. The first approach, adopted by Zhu (1995), attributes tone movement to final stress in compounds with initial tone 5. The contour movement process is exemplified derivationally as follows, where stress is indicated by '.

(37) tone 5 + tone1 → tone 5 + tone 1 → tone 5 + tone 1 → tone 5 + tone 1

$$
\begin{array}{cccccccc}
\sigma & \text{'}\sigma & \sigma & \text{'}\sigma & \sigma & \text{'}\sigma & \sigma & \text{'}\sigma \\
\wedge & \wedge & \wedge & & & \wedge & | & \wedge \\
\text{L M} & \text{H L} & \text{L M} & & & \text{L M} & \text{L} & \text{L M}
\end{array}
$$

The first step is to delete all non-initial tones. The second step moves the retained initial tone to the stressed syllable, whose underlying tone is deleted earlier. The third step inserts a L target to the initial syllable. The process is very similar to Zhenhai, except that the register is not neutralized to high under stress.

Based on the duration data and his own intuition as a native speaker of Shanghai, Zhu (1995) suggested that there are three distinct stress patterns in disyllabic compounds. When the first syllable is in tone 1, it is stressed and the second syllable is unstressed. When the first syllable is in tone 2 or tone 3, it is stressed and the second syllable receives secondary stress. When the first syllable is in tone 5, it is unstressed and the second syllable receives stress. When the first syllable is in tone 4, he admits that it is ambiguous.

His classification of stress patterns seems to correlate with his duration data. The table in (34) is repeated here for the sake of easy reference. In general, the first syllable rime is much longer than the second when both are long (i.e. not checked-tone syllables,

first three columns). The syllable after tone 1 has the shortest duration, corresponding to lack of stress. Tone 4 and tone 5 are short checked tones. The syllable after tone 5 has the longest duration, corresponding to stress.

(38) Rime duration of syllables in tone 1/2 and tone 3 (in ms)

σ_2	T1-σ_2 H-L	T2-σ_2 M-H	T3-σ_2 L-M	T4-σ_2 M-H	T5-σ_2 L-LM
T1/2	158-73	174-121	172-122	55-139	57-169
T3	187-81	191-135	202-129	77-161	65-193

The gist of the first approach is that initial tone is preserved, but stress occurs in different positions (final when the initial tone is tone 5; initial elsewhere), and the preserved tone is attracted to the stressed syllable.

There is an alternative explanation to duration variations, which does not need to involve stress. First, as we have discussed earlier, a rising f0 requires longer time to be implemented than a falling f0. This will explain away why the syllable after tone 1 is the shortest. Second, the rising tone requires longer time than a rising f0 transition because the former is often a delayed rise while the latter is often an immediate rise relative to the beginning of the host syllable. This will explain away why the syllable after tone 5 is the longest: it carries a rising tone.

The second approach assumes a uniform stress pattern (i.e. stress is initial) in all five patterns, but attributes the contour movement to contour licensing restrictions, extensively discussed in Zoll (1996, 2003), Gordon (1999) and Zhang (2001). The idea is that certain contour types are not allowed in certain prosodic positions. In Shanghai, we saw that contour tones are banned in domain-initial position, and they redistribute over the first two syllables in most cases. However, contour re-association is not the only

option. Contour movement is another option. When an underlying tone 5 occurs in the domain-initial position, contour movement is used to alleviate the problem of having a contour tone on a short syllable. The contour redistribution is also used when the compounds are longer than three syllables. As we mentioned earlier, this is a case in which the preserved tone moves from a stressed syllable to another prominent prosodic position: domain-final syllable in disyllabic and trisyllabic compounds. In addition, the uniform stress approach is compatible with the facts about word length restrictions and contrastive focus. A more complex story has to be told about those facts if one assumes that there are two different stress patterns co-existing in Shanghai, although such a scenario is not unattested in other dialects like Zhenhai.

Given what we have said about tone movement in Shanghai, we are still facing the problem of why the initial contour of tone 5 moves as a unit, but that of tone 4 re-associates to the first two syllables. In other words, if both tone 4 and tone 5 are short checked tones, why should they behave differently? To this question, we do not have a good answer and have to leave it open.

We conclude that Shanghai represents a typological pattern in which stressed syllables (initial) retain their underlying tones, and the realization of the retained tones are subject to tonal markedness constraints, to be discussed in chapter 6.

4.4.4   Final Tone Preservation and Initial Stress: Tone Movement in Wenzhou

Wenzhou is a coastal city in Zhejiang province. The variety spoken in this area belongs to southern Wu dialects. Its tone sandhi is characterized as being right-dominant (e.g. Yue-Hashimoto 1980), in the sense that the tones on syllables toward the end of a tone

sandhi domain determine the overall tone pattern, while tonal contrasts on syllables toward the beginning of the relevant domain tend to be neutralized (Ballard 1988).

Both auditory descriptions and phonetic studies exist of Wenzhou tones and tone sandhi. For example, Zhengzhang (1964a, 1964b, 1980) provide the most extensive and detailed descriptions of Wenzhou phonology, including tones and tone sandhi. Acoustic analyses of Wenzhou tones and tone sandhi can be found in a number of studies by Rose (1994, 2000, 2002). Chen (2000: 475ff) also gives a detailed analysis of Wenzhou tone sandhi, based on his own collection of data. Phonological discussions of Wenzhou tone sandhi are also available. In addition to Chen (2000: 475ff), they can also be found in Bao (1990, 1999), Duanmu (1990) and Yip (1995, 1999).

We first describe the citation tones and then lay out the disyllabic tone sandhi patterns, which form the basis for longer words. After that we will focus our discussion on a specific tone sandhi involving final tone movement. We will argue that in this specific tone sandhi, the final tone is preserved, but shifts to the initial stressed syllable.

4.4.4.1   Citation Tones in Wenzhou

According to Zhengzhang (1964a, 1964b, 1980), Wenzhou has an inventory of eight tones, which correspond to the four Middle Chinese tonal categories evenly split into two registers.

(39)   Citation tones in Wenzhou (Zhengzhang 1980)

register		tone			
		I ping "level"	II shang "rising"	III qu "departing"	IV ru "entering"
a.	yin	33 (M)	35 (MH)	42 (HM)	313 (L?)
b.	yang	31 (ML)	24 (MH)	11 (M)	212 (L?)

Chao's tone digits on a five-point scale are reduced to three distinct pitch levels, namely H, M, L. Both transcriptions are given in the table above. Zhengzhang (1980) mentioned that both IIa and IIb are read as 35, but the distinction is still maintained because of the onset voicing contrast. Similarly, IVa and IVb both are read as 13. Chen (2000:476) also noted that in his informant's speech the difference between the two rising (IIa and IIb) and two dipping tones (IVa and IVb) is barely perceptible. It can be regarded as intrinsic pitch variations correlated with the voiced (low register) and voiceless (high register) onsets. Therefore, the two rising tones and two dipping tones are labeled as MH and L? respectively.

The following graph, adapted from Rose(2002:1), displays the mean f0 shapes of the eight tones in Wenzhou. The graph is modified in two ways: first, time is changed from csec. to msec. (ms.); second, tone digits (from Zhengzhang, slightly different from Rose's transcriptions) are added after each tone to enhance readability.

(40)   Mean f0 shapes of the eight tones in Wenzhou (adapted from Rose 2002:1)
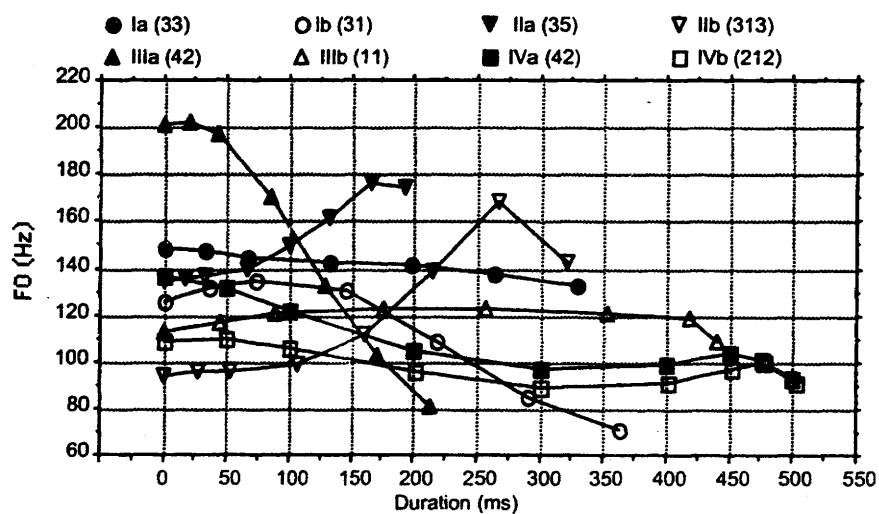
The f0 shapes are very close to the impressionistic transcriptions. It is worth

mentioning that the two checked tones (IVa and IVb), diachronically associated with

checked syllables ending in obstruent codas, surface synchronically as long tones. They

are marked by squares (empty and filled) in the graph. The two rising tones, marked by

inverted triangles (empty and filled), have the shortest duration in isolation.

4.4.4.2 Disyllabic Tone Sandhi

Wenzhou probably represents one of the most intricate tone sandhi patterns in all Chinese

dialects. For the two abutting tones, both may undergo change, the sandhi output being

determined by the tonal identity of both input tones. Out of sixty four combinatorial

possibilities (8 x 8), eight tonal melodies emerge, summarized in tabular form below.

They are referred to as patterns A, B, C (and A', B', C' ... when the first tone is checked

tone) in the published reports by Zhengzhang (1964a, 1964b, 1980).

(41) Disyllabic Tone Sandhi in Wenzhou (from Chen 2000:478)[18]

σ_1	σ_2					
	Ia M	Ib ML	IIIa HM	IIIb L	II MH	IV Lq
Ia M	A: M-M	C: L-L	E: MLM-HM		D: HM-MH	H: HM-Lq
Ib ML	A: L-M					
IIIa HM	B: HM-M	F: HM-ML		G: HM-L		
IIIb L						
II MH						
IV Lq	B': Lq-M	C: L-L	F': Lq-HM	G': Lq-L	D': Lq-MH	H': Lq-Lq

In the above table, neither IIa and IIb nor IVa and IVb are distinguished, for the reasons

mentioned above. We also give Zhengzhang's transcription below as a reference since it

---

[18] q stands for checked tone.

153

is the most often cited source of Wenzhou tones and tone sandhi. The two transcriptions agree with each other most of the time. Differences are probably due to a subdialectal difference between the pronunciations of informants. Note that the two tables are arranged in a slightly different manner with respect to the order of first syllable tone: I, III, II, IV in Chen, but I, II, III, IV in Zhengzhang.

(42) Zhengzhang's (1980) transcription

σ_1	σ_2					
	Ia 33	Ib 31	IIIa 42	IIIb 11	II 35	IV 313/212
Ia 33	A:11-33	C:11 13	E:13-53		D:53-35	H:53-13
Ib 31						
II 35	B: 42-33	F:42-1	F:42-1	G: 42-11		
IIIb 11						
IIIa 42		C:11-13				
IV 212/313	B':1 33		F':1-42	G': 1-11	D': 1-35	H':1-13

Some of the tone sandhi processes do not seem to have phonological motivation. For example, it is unclear why 35-33 and 11-33 both map to 42-33. This process is more like a paradigmatic substitution in the sense that one lexical tone changes to another tone in a specific tonal context. The well-known tone circle in Min dialects (Xiamen as an example) features such a process.

What we are concerned with are two sandhi processes which are related to final tone preservation and stress-conditioned tone shift.

4.4.4.3   Final Tone Preservation and Stress-Conditioned Tone Shift

We consider patterns F and F' and show that a phonological explanation can be motivated. Our analysis is based on the relevant discussions in Bao (1990, 1999) and Chen (2000) but differs from both of them.

Patterns F and F' are exemplified below. The examples are adapted from Bao (1999:90), transcribed in pinyin. The input tone values are modified to make them conform to Zhengzhang (1980).

(43) Patterns F and F'

F	output tone	input tone		example	gloss
a.		IIa-Ib	35-31	kou liang	"food ration"
b.		IIb-Ib	24-31	lao po	"wife"
c.		IIIa-Ib	42-31	zheng ming	"prove"
d.	42-1	IIIb-Ib	11-31	wai po	"grandmother"
e.		IIa-IIIa	35-42	hao huo	"good product"
f.		IIb-IIIa	24-42	yan jing	"eye glasses"
g.		IIIa-IIIa	42-42	chang pian	"record"
h.		IIIb-IIIa	11-42	yun qi	"luck"
F'	output tone	input tone		example	gloss
a.	1-42	IVa-IIIa	323-42	hei bu	"black cloth"
b.		IVb-IIIa	212-42	bai bu	"white cloth"

The combination of IIIa-Ib (c. in the above table) is shaded because it actually gives rise to L-L in Zhengzhang and Chen's descriptions although it is included in Bao (1999).

A nontrivial difference between the table above and Bao's is the output tonal shapes. In Bao (1999), it is taken to be 42-21, a high fall followed by a low fall. However, the original transcription is given as 42-1 in Zhengzhang, who further notes that the second syllable in pattern F is normally said in neutral tone (i.e. unstressed), and that 42-21 only appears in slow speech. The reason that Bao would like to stay with the slow speech version seems to be that he considers 21 to be a weakened 31, which is a low falling tone. Having made this assumption, he proposed the following analysis. First, he observes that the second tone is either a low fall /31/ or a high fall /42/. Second, according to Zhengzhang, the stress falls on the first syllable in disyllabic words, except when the first syllable is a checked syllable. Third, he proposes that the second tone

spreads to the first syllable, replacing the original tone of the first syllable. Finally, the stressed syllable is assigned a high register, thus changing a low fall to a high fall, and the unstressed syllable is assigned a low register. The whole derivation is replicated in autosegmental terms below.

(44) Contour spreading in pattern F

```
 'σ₁     σ₂     →    'σ₁    σ₂     →    'σ₁     σ₂
  /\     /\           \    /\          /\      /\
 r₁ c₁  c₂ r₂         c₂  r₂         r  c₂    c₂  r
     /\ |             /\ |           | /\    /\ |
    H L h/l          H L h/l         h H L   H L l
```
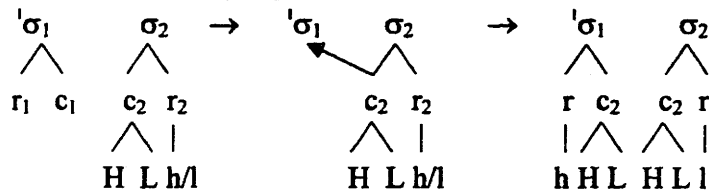
Bao's analysis featuring contour spreading is criticized by Duanmu (1990:148) and Yip (1995). They argue that there is insufficient evidence that 21 is HL rather than L. We add that there is no motivation to use slow speech version as the norm. They also point out that Bao's analysis cannot extend to the remaining disyllabic tone sandhi patterns in Wenzhou. Yip (1995) further argues that at the phrasal level, the tonal shapes of longer phrases are determined by the underlying tone of the last, and sometimes also the penultimate, syllable (see also Yip 1999, Chen 2000).

We note that Wenzhou is probably not the only Chinese dialect whose tone sandhi behavior is sensitive to specific tonal combinations. We have seen that in Zhenhai, the tone movement is conditioned the position of stress, which is in turn determined partly by tonal combinations. In Shaoxing (F. Wang 1959), the general pattern is that the initial tone is retained and then spread to the whole domain. However, when the initial tone belongs to MC tonal categories I and IV, the resulting tonal patterns alternates between two shapes, depending on whether the second tone belongs to MC tonal categories I and

IV or II and III. Examples like these are abundant in Chinese dialects (see Chen 2000 for more examples).

The analysis we will propose adopts Bao's insights on stress and tone sandhi in Wenzhou. However, instead of contour spreading, we propose a tone movement analysis in which the final underlying tone is preserved, but it shifts to the stressed syllable. In pattern F, the stress is initial, therefore, the preserved final tone moves to the initial stressed syllable, leaving behind an unstressed syllable, whose phonetic pitch is determined either by a default L or boundary L insertion. In pattern F', the stress is final, and the preserved final tone stays put. Such an interpretation is consistent with Zhengzhang's original transcription of pattern F as 42-1. The slow speech version, 42-21, is actually an artifact of slowing down of speech. In addition, when it moves to the initial stressed syllable, the preserved tone is realized as a high-register falling tone regardless of its underlying register specification. The whole process is illustrated below.

(45) Final tone movement in pattern F

$$\text{'}\sigma_1 \quad \sigma_2 \quad \rightarrow \quad \text{'}\sigma_1 \quad \sigma_2 \quad \rightarrow \quad \text{'}\sigma_1 \quad \sigma_2$$

$$
\begin{array}{llll}
r_1\ c_1 & r_2\ c_2 & r_1\ c_1\ r_2\ c_2 & r\ c_2\quad r\ c \\
& \wedge & \uparrow \qquad | & |\ \wedge\ |\ \ | \\
& H\ L & & h\ H\ L\ \ \ L\text{-}
\end{array}
$$
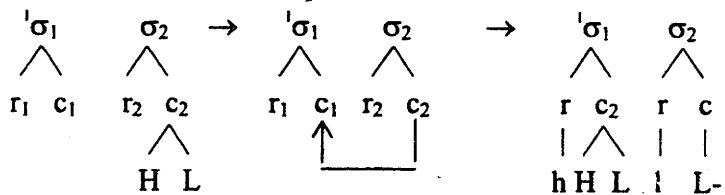
Our movement analysis differs from Bao's spreading analysis in that the underlying tone of the final syllable is detached from its host syllable, and moves to the initial stressed syllable. Bao's problem stems from the fact that in order to do contour spreading, he has to assume that the surface tone (i.e. output tone) on the final syllable spreads to the initial syllable. Since what we observe is a fall on the initial syllable, supplanting its underlying tone, he is forced to say that the surface tone on the final

syllable is also a fall. Otherwise, contour spreading would not work. In our analysis, once the input tone of the final syllable moves to another position, it has nothing to do what which tone will show up on the final syllable. Either default L or boundary L will be inserted. In this sense, pattern F in Wenzhou is the mirror image of the W-S tone sandhi pattern in Zhenhai. In the former, the final tone is preserved, but moves to the initial stressed syllable; in the latter, the initial tone is preserved, but moves to the final stressed syllable. In pattern F', the preserved final tone stays put because the stress is also final.

On spans of three, four or five syllables, if the last two syllables are subject to pattern F, then the overall tonal shape will be determined by the position of stress. Zhengzhang (1980) reports the following tonal patterns.

(46) Last two syllables belonging to pattern F

a.	σσσ	42-1-1, 1-42-1
b.	σσσσ	42-1-31-1
c.	σσσσσ	42-1-31-1-1

The trisyllabic words have alternating patterns, depending on where the main stress is. What happens is that the final falling tone moves to the stressed syllable, and is realized in high register. The unstressed syllables lose their underlying tones, being assigned a boundary L tone. In longer spans a low fall /31/ breaks the strings of 1s, imposing alternating stress on a toneless span (Yip 1995:481). This is predicted from our analysis by simply assuming that the initial syllable has the main stress and the penult or antepenult has the secondary stress. Both attract the retained final syllable tone.

Before we conclude this subsection, it is necessary to mention a tonal process dubbed "HM-shift" by Chen (2000). According to him, when the ditonic pattern HM-ML (corresponding to pattern F) is mapped to longer span of syllables, the high-falling HM is moved to the first syllable of the compound, skipping over any number of syllables in between. In the process, the medial syllables are linked with mid-level pitch. The HM-shift is stated below. Chen is not explicit about why HM moves to the first syllable of the compound. In our analysis, the high falling HM originates on the final syllable and moves to the syllable where stress falls.

(47) HM-shift (Chen 2000:486)

$$[\sigma \ldots \sigma \quad \sigma]_{\text{word}} \quad \rightarrow \quad [\sigma \ldots \sigma \quad \sigma]_{\text{word}}$$
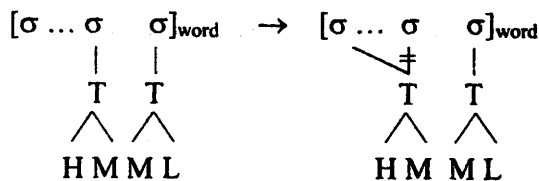
In summary, the pattern F in Wenzhou represents a case in which the final syllable tone is preserved, but is realized on the stressed syllable in non-final position.

4.4.5   Final Tone Preservation and Final Stress: Yantai and Xiamen

4.4.5.1   Final Tone Preservation in Yantai

Yantai is a northern Mandarin dialect spoken in Shangdong province. This dialect has only three citation tones: /31[19], 214, 55/. Its disyllabic tone sandhi patterns are given below. All the data are from Chen (2000), which cites Z. Qian (1981).

---

[19]   The tone 31 splits into 31a and 31b in sandhi context. They derive from different historical sources, namely *yin ping* and *yang ping*, respectively. In citation form, they merge completely.

(48) Disyllabic tone sandhi

    a. 31a → 35 / ___ 31

    b. 31b → 55 / ___ 31

    c. 55 → 31 / ___ 55

    d. 214 → 35 / ___ 31

         55 / ___ 214

The disyllabic tone sandhi in Yantai is no different from other Mandarin dialects in that the first tone changes as a result of the following tonal context. The change could be motivated by phonological considerations, like assimilation and dissimilation, or just paradigmatic substitution.

What is remarkable about Yantai is its trisyllabic tone sandhi, which features a sweeping positional neutralization. Only three surface tonal melodies are found in trisyllabic tonal combinations.

(49) Trisyllabic tone sandhi (T=any tone)

    a. T-T-31 → 31-35-31

    b. T-T-214 → 55-55-214

    c. T-T-55 → 33-21-55

What happens here is that only the final syllable is able to preserve its underlying tone; the non-final syllables lose their underlying tones completely. This can be seen by plugging in the three citation tones in the first two positions in (a), and the resulting shape is still /31-35-31/. No mentioned is made about stress in Chen (2000). Therefore it is impossible to tell whether the final tone preservation is due to final position prominence

or final stress. In Xiamen, there is ample phonological and phonetic evidence which points to the conclusion that it is domain-final stress that is responsible for preserving the final tone.

4.4.5.2 Tones and Tone Sandhi in Xiamen

Xiamen (also Amoy) is representative of southern Min dialects. Its tone sandhi and tone sandhi domain formation have been extensively explored in the literature (Wang 1967, Yip 1980, Wright 1983, Chen 1987, 1992, J. Lin 1994, Duanmu 1995, among others). We focus on the relation between final tone preservation and prominence (final or metrical). Some tonal preamble is given first.

Xiamen distinguishes seven tones in citation, five long tones and two short tones. Long tones occur on smooth syllables and short tones on checked syllables. According to Chen (1987, 2000), in a tone sandhi domain only the domain-final tone preserves its citation form; other tones (domain-initial and domain-internal) are replaced by their corresponding sandhi tones. The pattern of substitution of one tone by another has been referred to as the "southern Min tone circle" (for long tones). The citation and sandhi tones in Xiamen are given below.

(50) a. Citation and sandhi tones in Xiamen

	long tones / smooth syllables					short tones	
citation	44 (H)	24 (MH)	22 (M)	21 (ML)	53 (HM)	4? (H)	32? (ML)
sandhi	22	22	21	53	44	21?	4?, 53?

b. Smooth syllable tone sandhi (tone circle) (Chen 2000:432)

$$24$$
$$\downarrow$$
$$22$$
$$44 \nearrow \quad \searrow 21$$
$$\nwarrow 53 \quad \swarrow$$
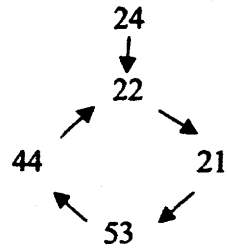
c. checked syllable tone sandhi

4ʔ → 21

32ʔ → 4ʔ for syllables ending in p, t, k

53ʔ for syllables ending in ʔ

4.4.5.3 Final Tone Preservation and Stress

We have left undefined what a tone sandhi domain is in Xiamen. Chen (1987) provides an algorithm which sees a direction relation between syntax and prosody. Basically, a tone sandhi domain can be formed by "marking the right edge of every XP with #, except where XP is an adjunct c-commanding its head" (Chen 1987:131). Later refinement of his proposal can be found in J. Lin (1994).

The crux of the problem is whether the final tone preservation happens because it is final in a tone sandhi domain, or because it is hosted by the domain stress which happens to be final. If it is simply final tone preservation, then we have to worry about where the stress is and whether stress plays any role is determining how the preserved tone is realized. If it is the final stressed syllable that preserves its underlying tone, then Xiamen becomes a mirror image of Shanghai and Suzhou which preserve the underlying tone of the initial stressed syllable. We will show that it is the domain-final stress that induces final tone preservation by drawing on evidence from word length restriction,

focus involving lexical elements (Duanmu 1995), and focus involving function words (e.g. pronouns) (Hsiao 2002).

Duanmu (1995) argues for right-headed stress in Taiwanese (same dialect as Xiamen). He used the same word length test as in Shanghai. He found that like Shanghai, for 1+2 and 2+1 trisyllabic word structures, either one or two sandhi domains can be formed in Taiwanese. However, only 1+2 is parsed into two sandhi domains in Taiwanese. Recall that in Shanghai only 2+1 is parsed into two domains.

(51) a. 1+2
$[\sigma]_{w1}[\sigma\sigma]_{w2}$ word structure
$(\sigma\sigma\sigma)$ or $(\sigma)(\sigma\sigma)$ metrical structure

b. 2+1
$[\sigma\sigma]_{w1}[\sigma]_{w2}$ word structure
$(\sigma\sigma\sigma)$ but $*(\sigma\sigma)(\sigma)$ metrical structure

Following his analysis for Shanghai, Duanmu (1995) argues that such an asymmetry falls out naturally in a metrical analysis assuming right-headed stress for Taiwanese.

(51) a. no stress clash

```
        x                    x
    x   x                    x
 (σ)(σ σ)      or      (σ σ σ)
```

b. stress clash

```
        x                    x
    x   x                    x
 *(σ σ)(σ)      →      (σ σ σ)
```

In (a) there is no stress clash and it is optional to delete the initial stress. In (b) the stress clash has to be resolved by removing the weaker, initial stress.

163

The right-headed stress is further supported by its effect on tone sandhi domain formation. As Duanmu (1995:24) shows, a modifier-noun compound forms one domain, but focusing on the modifier breaks it into two.

(52) a. neutral, no contrastive focus
      (bi-kok   sio-tsia)   one domain
      America Miss   'Miss America'

    b. contrastive focus on modifier
      (bi-kok)   (sio-tsia) two domains
      AMERICA Miss   'Miss AMERICA'

The choice of citation tone or sandhi tone depends on how the tone sandhi domain is formed.

So far we have reviewed evidence for right-headed stress in Taiwanese. It is yet to be shown that it is the property of being stressed rather than being final that preserves the tone. The evidence for the former is furnished by focus involving pronouns (Hsiao 2002).

Hsiao (2002) examines the tonal behavior of subject and object pronouns under focus in Taiwanese. Since the subject pronouns under focus have no effect on tone sandhi patterns, we only focus on her discussion on object pronouns. She demonstrates the following patterns, exemplified by examples from Hsiao (2002).

(53) Tonal behavior of object pronouns under focus

object pronouns	neutral	focus
non-domain-final	sandhi tone	sandhi tone
domain-final	default or from preceding tone	citation tone

When the object pronoun is in non-domain-final position, it takes on its sandhi tone no matter whether it is focused or not. When the object pronouns are in domain-final

position, they get a default pitch or the preceding tone in neutral context. The preceding lexical element keeps its citation tone. The object pronouns behave like their counterparts in Mandarin, which are de-stressed and de-toned, and are said in neutral tone. When the object pronouns are focused in domain-final position, their own citation tones surface, and the preceding lexical element surfaces with its sandhi tone.

(54) a. object pronouns in non-domain-final position (neutral/focus)

```
a-  gong ga QUA gong go-su
33 44   33 53   53   31 33        citation tone
(31 44)(31 44   44   53 33)       surface tone
grandpa to ME tell   story        'Grandpa told me a story.'
```

b. object pronouns in domain-final position (neutral)

```
ma-ma  hua qua
53 31  31  53                     citation tone
(44 31)(31  31)                   surface tone
mother ask  me                    'Mother asked me (something).'
```

c. object pronouns in domain-final position (focus)

```
ma-ma  hua QUA
53 31  31  53                     citation tone
(44 31)(53  53)                   surface tone
mother ask  ME                    'Mother asked me (something).'
```

Note that in (b), the object pronoun takes the tone of the preceding lexical item when it is neutral in domain-final position. Its sandhi tone would be 44 (from citation 53).

Hsiao's (2002) study on focus in Taiwanese tone sandhi clearly reveals the relation between tone preservation and stress. In neutral context, the domain-final lexical element receives the domain-final stress, therefore its citation tone is preserved. Function words do not carry domain-final stress because they are unstressed in final position.

When the domain-final pronoun is contrastively focused, it carries the domain-final stress, which makes its citation tone surface.

We conclude that Xiamen (Taiwanese) tone sandhi represents a pattern in which the syllable carrying domain-final stress preserves its underlying tone. All the syllables preceding the domain-final stress take their corresponding sandhi tones.

4.4.6   Tone Sandhi as Contextual Modification

We have shown that in order to account for a wide range of tone sandhi phenomena, both edge prominence and metrical prominence have to be invoked by the grammar, and that they interact to derive the surface tonal shapes. However, as we noted earlier, there are tonal systems in Chinese dialects whose tone sandhi is more appropriately described as contextual modification. In a disyllabic combination, the first, second or both tones can be modified because of the nature of the two abutting tones. What really matters is the nature and the linear order of tones. Stress does not play any role in determining which tone is preserved and which tone is modified. Sometimes, stress or focus can influence the formation of tone sandhi domain formation, but it does not change the way a tone sandhi process applies. Tone sandhi in many Mandarin dialects exhibits this property, for example, Mandarin Chinese (Chao 1968), Tianjin (Li and Liu 1985), Jinan (Z. Qian 1995), Canzhou (Wang 1995). In what follows, we will illustrate with Mandarin third tone sandhi by first presenting Duanmu's (2000) metrical theory of stress in Mandarin Chinese.

In Duanmu's theory, stress in Mandarin is determined in two ways. First, a simple compound forms disyllabic trochees from left to right, and the first trochee has

more stress than others. Second, compounds and phrases are subject to Nonhead stress, by which syntactic nonheads must have stress relative to syntactic head (Duanmu 1990, Cinque 1993). For example, in a verb-object construction, the object gets stress while in a modifier-noun construction, the modifier gets stress.[20]

The third tone sandhi applies to two full syllables in tone 3 and changes the first one into tone 2. It applies within a prosodic domain. Following Shih (1986, 1997), Chen (1996), and Duanmu (2000), the tone sandhi applies cyclically starting from each foot, and optionally between two cyclic branches. The main point to be made here is that when the third tone sandhi applies, it always changes the first tone 3 no matter if it is the head or non-head of the foot. The head position within a foot is initial, but could be final when the initial stress is overridden by the Nonhead stress.[21] Consider the following two examples:

(55) a.   x
      [xiu gai]   "to revise"
      (214 214)   compound stress on *xiu*
      (<u>35</u> 214)

   b.   x
      [hao jiu]   "good wine"
      (214 214)   Nonhead stress on *hao*
      (<u>35</u> 214)

   c.   x
      [mai jiu]   "buy wine"
      (214 214)   Nonhead stress on *jiu*
      (<u>35</u> 214)

---

[20] See Duanmu (2000:130ff) for details about how Nonhead stress is assigned.
[21] Duanmu (2000) maintains that the foot is always trochaic. When the Nonhead stress assigns stress to the final syllable, it forms a binary trochaic foot with an empty beat.

The compound stress (trochaic) and Nonhead stress assign either initial or final stress, but the third tone sandhi always changes the tone 3 to the left.

Emphasis affects foot structure, but does not protect tone 3 from being changed. The following examples are similar to the ones discussed in Duanmu (2000). When the main verb *mai* is focused, it becomes the head of trisyllabic foot (b). Nevertheless, the focused tone 3 becomes tone 2 by the third tone sandhi. The two third tones in the middle in (a) are not subject to the tone sandhi because they belong to two cyclic branches.

(56) a.   x   x
      [xiang [mai [hao jiu]]]   "want to buy good wine"
      (xiang mai)(hao jiu)   foot structure
      (214   214)(214 214)
      (<u>35</u>   214)(<u>35</u> 214)

    b.   x
      [xiang [MAI [hao jiu]]]   "want to BUY good wine"
      xiang (MAI   hao jiu)   foot structure
      214   (214   214 214)
      214   (<u>35</u>   <u>35</u> 214)

Similar arguments can be made for other Mandarin dialects with respect to tone-prominence interaction. Cases like Mandarin do not undermine the dual-prominence theory. As we will show later, an OT typology of tone-prominence interaction predicts the typological pattern in which reference to both prominent positions is suppressed when the prominence-related constraints are low ranking.

4.5 Concluding Remarks

We have argued that both peripheral positions and metrically strong position in a prosodic domain have to be admitted into the grammar in order to explain a wide range of tone sandhi phenomena in Chinese dialects. Furthermore, we have shown that different prominent positions interact with each other to give rise to the surface tonal patterns. For example, although tones are preserved in either initial or final position, they are attracted to where metrical prominence resides. When tone retention and stress occur in different positions, tone movement results, as we have demonstrated in Zhenhai and Wenzhou. A single-prominence theory is not able to account for tone movement in an obvious way.

In the next chapter, we will lay out an OT-based formal theory of tone movement, using Zhenhai lexical tone sandhi as an example.

Chapter 5 A Formal Theory of Tone Movement

5.1   Introduction

In the last chapter, we presented evidence to illustrate that peripheral positions in a prosodic domain, together with the metrically prominent position, have to be referred to in order to account for a wide range of tone mapping processes involving tone retention and tone deletion. They are singled out as privileged positions by the phonology. We have shown that tone movement occurs when a tone is retained in one of the peripheral positions (initial or final), but may be attracted to another syllable carrying metrical prominence. Zhenhai lexical tone sandhi exhibits initial tone preservation and rightward movement of the retained tone, while Wenzhou pattern F exhibits final tone preservation and leftward movement of the retained tone. The aim of this chapter is to develop a formal theory of tone movement in Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1995) which makes reference to metrical prominence and edge prominence. It will be shown that an adequate account of tone movement relies on positional faithfulness constraints and positional markedness constraints, arguing for the necessity of admitting both in the grammar. The theoretical apparatus of our approach will be reviewed, and then applied to the lexical tone sandhi in Zhenhai. The implications of Zhenahi tone movement for positional faithfulness theory (Beckman 1998) and phonetically-driven contour tone licensing theory (Zhang 2001) will also be discussed.

5.2 Positional Faithfulness and Positional Markedness

5.2.1 Positional Faithfulness Constraints

In the Correspondence Theory developed by McCarthy and Prince (1995), faithfulness constraints regulate the identity of correspondent strings (S) in the input-output or base-reduplicant relation. Three of the faithfulness constraint families are shown in (1) (from McCarthy and Prince 1995):

(1)　a.　MAX: Every segment of S_1 has a correspondent in S_2. ('Do not delete.')

　　　b.　DEP: Every segment of S_2 has a correspondent in S_1. ('Do not insert.')

　　　c.　IDENT(F): Correspondent segments in S_1 and S_2 have identical values for some feature [F]. ('Do not change a feature value'.)

Compared to the context-free faithfulness constraints, positional faithfulness constraints are defined relative to a specific phonological context. In this approach, positional phonological asymmetries can be modeled by the following ranking schema (Beckman 1998).

(2)　positional faithfulness >> markedness >> context-free faithfulness

Positional faithfulness constraints, when ranked over markedness constraints, protect the phonological structure in prominent positions from being modified by the markedness constraints. Phonological structure in non-prominent positions is subject to modifications induced by markedness constraints.

We first entertain the following context-free faithfulness constraints with reference to tone, defined below:

172

(3)   a. MAX(TONE): If T is a tone in the input, then T has a correspondent in the output.

('Do not delete a tone.') (after McCarthy and Prince 1995)

   b. DEP(TONE): If T is a tone in the output, then T has a correspondent in the input.

('Do not insert a tone.') (after McCarthy and Prince 1995)

   c. IDENT(TONE): If a tone bearing unit has a tonal specification in the input, then

its correspondent has identical tonal specification in the output. ('Do not change

a tonal specification.') (after McCarthy and Prince 1995)

We propose that the positional faithfulness constraints, MAX(TONE)/INITIAL,

MAX(TONE)/FINAL, MAX(TONE)/σ́, and IDENT(TONE)/σ́, are responsible for preserving

tones in different prominent prosodic positions. They are defined as follows.

(4)   a. MAX(TONE)/INITIAL: if T is a tone in the domain-initial position in the input,

then T has a correspondent in the output. ('Do not delete a tone in the domain-

initial position.')

   b. MAX(TONE)/FINAL: If T is a tone in the domain-final position in the input, then

T has a correspondent in the output. ('Do not delete a tone in the domain-final

position.')

   c. MAX(TONE)/σ́: If T is a tone on the stressed syllable in the input, then T has a

correspondent in the output. ('Do not delete a tone on the stressed syllable.')

   d. IDENT(TONE)/σ́: If a stressed syllable has a tonal specification in the input, then

its correspondent has identical specification in the output. ('Do not change the

tonal specification of a stressed syllable.')

When MAX(TONE)/INITIAL and MAX(TONE)/FINAL are high ranking, they have the effect of protecting the domain-initial and domain-final tones from being expunged in the output. Similarly, MAX(TONE)/σ́ preserves input tones of the stressed syllables. IDENT(TONE)/σ́ has the effect of ensuring that stressed syllables maintain the identical tonal specifications in tone mapping. For example, in Xiamen the syllable carrying domain-final stress preserves its underlying tonal specification. It is violated when the input tonal specification of a stressed syllable is changed in the output. Contour tone re-association in Shanghai and other Wu dialects in general incurs a violation of IDENT(TONE)/σ́ if the left-headed stress in Shanghai is responsible for initial tone preservation. In other words, although the underlying contour tone of the initial stressed syllable is preserved, only the first tone segment of the contour tone is eventually realized on the stressed syllable. As originally conceived by McCarthy and Prince (1995), input-output featural faithfulness (tonal specifications in our case) is mediated through a segment.[1] That is why IDENT(TONE)/σ́ is violated when an underlying contour tone of the stressed syllable is retained, but spreads over into the following syllable on the surface. MAX(TONE)/σ́ assesses the tonal faithfulness of the correspondent input and output without going through a segmental mediator. Therefore, contour tone re-association passes through MAX(TONE)/σ́, but not IDENT(TONE)/σ́.

---

[1] Struijke (2002) proposes an existential interpretation of IDENT, which is able to account for feature movement, but maintains segmentally-based implementation of IDENT. We adopt the original universal interpretation of IDENT.

MAX(TONE)/INITIAL and MAX(TONE)/FINAL make it possible that underlying tones of the domain-initial or -final syllables will surface in the output. Since these two positional faithfulness constraints preserve input tones at prosodic edges regardless of whether their host syllables are underlyingly stressed or not, unstressed syllables can retain their underlying tones in our constraint system. This scenario does not happen in a single-prominence model where only stressed syllables retain their underlying tones. The tableau below illustrates how an initial unstressed syllable retains its underlying tone in our analysis. The general markedness constraint, *TONE, is invoked to delete any tonal specifications of a tone bearing unit.

(5) a. *TONE: Any tone bearing unit with a tone in the output is banned.

   b. MAX(TONE)/INITIAL >> *TONE >> MAX(TONE)/σ́, MAX(TONE)

σ₁ σ́₂ | | T₁ T₂	MAX(TONE)/INITIAL	*TONE	MAX(TONE)/σ́	MAX(TONE)
a. σ₁ σ́₂ | T₂	*!	*		*
b. σ₁ σ́₂ \\\| T₂	*!	**		*
c. ☞ σ₁ σ́₂ | T₁		*	*	*
d. ☞ σ₁ σ́₂ | T₁		*	*	*
e. σ₁ σ́₂ |/ T₁		**!	*	*
f. σ₁ σ́₂ | | T₁ T₂		**!		

175

In this hypothetical example, the input is disyllabic with stress on the second syllable, marked by an acute accent. Candidates (a) and (b) each incur a fatal violation of MAX(TONE)/INITIAL by deleting the initial tone from the input. The initial tone surfaces in candidates (c), (d), (e) and (f), but the next high-ranking constraint *TONE is violated twice in (e) and (f) because both syllables are associated with a tone. Thus the winner is (c) or (d). Both (c) and (d) keep the initial tone, but (d) moves it to the stressed syllable. Two other constraints from the MAX family are also considered in the tableau, MAX(TONE)/ǒ and MAX(TONE). MAX(TONE)/ǒ requires keeping the input tone on a stressed syllable. Given that stress is on the second syllable, when it is outranked by MAX(TONE)/INITIAL, the input tone on the initial unstressed syllable is preserved. The general MAX(TONE) constraint is always violated when some input tones are deleted from the output.

The above example illustrates how the initial unstressed tone can be preserved in the output, but it does not indicate where the retained tone will be realized; both (c) and (d) survive. Further, nothing is said about the surface toneless syllables in (a), (c) and (d). They are dealt with by positional markedness constraints.

5.2.2   Positional Markedness Constraints

Candidates satisfying positional faithfulness constraints are not guaranteed to win. They have to be assessed by positional markedness constraints in the grammar. In our constraint system, wherever an input tone is retained, i.e. initial or final in the prosodic domain, it will be attracted by the stressed syllable. However, whether or not it will

eventually surface on the stressed syllable is further scrutinized by contour licensing constraints. For example, in Zhenhai W-S disyllabic tone sandhi, the initial contour tone is preserved. The preserved tone surfaces on the second syllable (i.e. stressed) when it is phonologically long. When the second syllable is a short checked syllable, it defaults to H. Therefore, surface tonal shapes come out from the interaction of positional faithfulness and positional markedness constraints.

In tonal systems with extensive tone neutralization, tone deletion is triggered by the general markedness constraint, *TONE, which penalizes any tone bearing unit with a tonal specification. No reference is made to the metrical status of the tone bearing unit, though. Alternatively, a positional markedness constraint can be proposed to penalize any tonal specifications on unstressed syllables, *TONE/UNSTRESSED σ. When they are outranked by other positional faithfulness and positional markedness constraints, their effects will be nullified. We will not try to tease them apart and will use *TONE in the following discussions.

(6)   *TONE/UNSTRESSED σ: Any tonal specification on an unstressed syllable is banned.

The constraint system is to be constructed such that prominent positions in a prosodic domain should be able to retain the underlying tones of syllables occurring there. For example, in Zhenhai W-S disyllabic tone sandhi, it is important to observe that the initial syllable preserves its underlying tone, but the second syllable is stressed. The preserved tone moves to the stressed syllable, creating tone movement. This is a case of initial-position prominence and metrical prominence occurring in different positions.

Realization of the retained tone on a stressed syllable can be achieved by contour tone licensing constraints (Zoll 1996, 2003, Zhang 2001). They constitute a subset of

positional markedness constraints that restrict the distribution of tonal contours to a limited set of positions. The following contour licensing constraints will figure prominently in our analysis.

(7)   Contour licensing constraints I: negative constraints

    a. *CONTOUR/UNSTRESSED σ: No contour tone is allowed on an unstressed syllable.

    b. *CONTOUR/σ́: No contour tone is allowed on a stressed syllable.

    c. *CONTOUR/SHORT: No contour tone is allowed on a short syllable.

    d. *CONTOUR/LONG: No contour tone is allowed on a long syllable.

    e. *CONTOUR/INITIAL: No contour tone is allowed on a domain-initial syllable.

    f. *CONTOUR/FINAL: No contour tone is allowed on a domain-final syllable.

Recent studies on the phonetics of contour tone distribution (Gordon 1999, Zhang 2001) point to the conclusion that the contour tone licensing constraints are unified in their reference to positions of longer duration and higher sonority. For example, contour tones are licensed on a long vowel or a stressed syllable, both of which have longer duration than the corresponding short vowel or unstressed syllable. By the same token, final position in a prosodic domain enjoys durational advantage due to a final lengthening effect. Therefore a contour tone is more likely to be licensed in domain-final position than domain-initial position. However, such unification does not seem to exist in Zhenhai. We saw contrasting tonal contours appearing only on a stressed syllable. In the W-S tone sandhi, the stressed syllable occurs in the non-initial position, but it has much shorter duration than the initial, unstressed syllable. Contour tones are disallowed on the

unstressed syllables. If the stressed syllable is phonologically short (in tone 5 and tone 6), it does not carry a contour tone either.

In addition to the negatively defined contour licensing constraints, we also propose the following positively defined contour licensing constraints, following Zoll (1996, 2003). These positional markedness constraints restrict certain contour shapes to different prosodic positions, either strong or weak. Constraints referring to stressed syllables are given below:

(8)   Contour licensing constraints II: positive

   a. COINCIDE(ố, CONTOUR): If a syllable is stressed, it bears a contour tone.

   b. COINCIDE(ố, H): If a syllable is stressed, it bears a H tone.

   c. COINCIDE(ố, L): If a syllable is stressed, it bears a L tone.

What we show above are contour licensing constraints referring to metrically prominent positions (i.e. stressed syllables). In fact, they can refer to other prominent positions as well, like initial and final positions in a prosodic domain, and regulate the contour shapes in those positions. For example, COINCIDE(FINAL, H) and COINCIDE(FINAL, L) will restrict a level H or L to a domain-final position. We will see their effect in an account of non-local migration of tonal targets in Tangxi in chapter 6.

In the forgoing discussions, contour tone licensing constraints are positively as well as negatively stated. Although both positive and negative positional markedness constraints have been used in the literature (for example, positively defined constraints in Zoll 1996, 1997, 2003, and negatively defined constraints in Zhang 2001), they achieve similar results with respect to contour tone distribution in those studies. In her discussion

of de Lacy's (2001) theory of tone-prominence interaction, Yip (2001)[2] suggests that Wuming Zhuang (a Thai language spoken in Guangxi province, China) is a suggestive case for positive positional markedness constraints. However, it can still be handled within de Lacy's theory using negative constraints. As we present our analysis of Zhenhai, it will be seen that both negative and positive licensing constraints are needed to explain the complex lexical tone sandhi of this dialect. Further, contour licensing constraints interact with positional faithfulness constraints like MAX(TONE)/INITIAL to derive surface tone patterns. It is worth mentioning that there are significant redundancies in the contour tone licensing constraints we have proposed. Further research is required to determine if this set can be reduced.

Up to now, we consider constraints referring to TONE as a conjunction of constraints referring to register feature (r) and contour feature (c) respectively. Such a consideration is in line with the model of tone representation that we are assuming (Bao 1990, 1999, Chen 2000). To avoid confusion with contour tone, we use TARGET for the contour node (c) in the representation of a tone. A tonal target can be either level or contour. For example, MAX(TONE)/INITIAL is used as a cover term for MAX(REG)/INITIAL and MAX(TARGET)/INITIAL. When both REG and TARGET are referred to simultaneously, we will simply use TONE.

---

[2]   Yip's (2001) discussion of de Lacy's theory is based on the unpublished version of de Lacy (2001).

5.3   An OT Analysis of Tone Movement in Zhenhai

In this section, we present a complete analysis of Zhenhai lexical tone sandhi, using the theoretical apparatus outlined above. The tone patterns in disyllabic and trisyllabic tone sandhi will be reviewed first before an analysis is given.

5.3.1   W-S Disyllabic Tone Sandhi

As we have discussed in chapter 4, the W-S tone sandhi in Zhenhai is characterized as initial tone preservation and movement of the preserved tone to the second stressed syllable. This is what happens when both syllables are phonological long. When a checked syllable (i.e. phonologically short) is involved, it defaults to H in the stressed position regardless of its underlying tone. Tone movement does not occur when the initial syllable is checked. The underlying tone of the stressed syllable shows up on the surface, with its register specification neutralized to high (pattern D in the following table). Since the initial syllable preserves its underlying tone in some contexts (patterns A, B, C) and the stressed syllable preserves its underlying tone in other contexts (pattern D), conceivably there will be a competition between MAX(TONE)/INITIAL and IDENT(TONE)/σ́. The major task confronting us is to determine the ranking of these two positional faithfulness constraints and their ranking with other markedness constraints.

Tone mapping in W-S disyllabic patterns is summarized in the following table. When the second syllable is underlyingly associated with a checked tone, the tonal output is found on the row marked with "-?". Stress falls on the second syllable in W-S disyllabic tone sandhi.

(9)   Tone mapping in W-S disyllabic tone sandhi

patterns	A: T1 + T	B: T3+ T	C: T4 + T	D: T6 + T
input	HL – T	ML – T	LM – T	L? – T
output	M – HL	L – HL	L – MH	L? – T^h
	-? M – H?	L – H?	L – H?	L – H?

(T = any tone, T^h = tone in high register)

We have also arrived at the following generalizations regarding W-S disyllable tone sandhi.

(a) Contrasting tonal contours (rising and falling) occur on the second syllable when it is phonologically long;

(b) Tonal targets on the metrically prominent syllable are all realized in high register;

(c) A short checked syllable defaults to H when it is stressed;

(d) The tonal target of the initial tone is retained and shifted onto the metrically prominent syllable in long-long combinations;

(e) Given an initial checked syllable, the second syllable retains its underlying tonal target;

(f) The initial syllable retains its underlying register specification;

(g) The initial syllable gets a low boundary tone and the underlying specification for tone register on the initial syllable determines how the low tone target is actually realized,

Tone movement from an unstressed first syllable to a stressed second syllable motivates the ranking of MAX(TONE)/INITIAL over IDENT(TONE)/ó. When the initial tone shifts to the stressed syllable, it supplants the underlying tone of the stressed syllable in

182

violation of IDENT(TONE)/σ́. Both syllables are long in this case, with stress on the

second syllable.

(10) a. HL (tone 1) + MH (tone 2) → L- + HL
      b. MAX(TONE)/INITIAL >> IDENT(TONE)/σ́

tone 1 $\sigma_1\ \acute{\sigma}_2$ + | | tone 2 $F_1\ R_2$	MAX(TONE)/INITIAL	IDENT(TONE)/σ́
a. $\quad\ \sigma_1\ \acute{\sigma}_2$ $\qquad\quad$| $\qquad\quad R_2$	*!	
b. ☞ $\ \sigma_1\ \acute{\sigma}_2$ $\qquad\quad$| $\qquad\quad F_1$		*
c. ☞ $\ \sigma_1\ \acute{\sigma}_2$ $\qquad$| $\qquad F_1$		
d. ☞ $\ \sigma_1\ \acute{\sigma}_2$ $\qquad$| | $\qquad F_1\ R_2$		

R and F are used as shorthand for rising and falling tones. Each candidate deletes an

underlying tone from the input. Deleting the underlying tone of the first syllable (a)

crucially violates MAX(TONE)/INITIAL, although IDENT(TONE)/σ́ is satisfied. The winner

(b) satisfies MAX(TONE)/INITIAL by preserving the initial tone. When the initial tone

moves to the stressed syllable, IDENT(TONE)/σ́ is violated. (c) and (d) are even better

candidates because they satisfy both positional faithfulness constraints. They are shaded

in the tableau because they are excluded by contour licensing constraints. For example,

(c) violates a constraint requiring a stressed syllable to be associated with a contour tone

COINCIDE(σ́, CONTOUR), which is satisfied by the winner (b). (d) has more violations

than the winner of the general tonal markedness constraint *TONE, and it also violates *CONTOUR/UNSTRESSED σ for failing to eliminate a contour tone on the unstressed syllable.

It is observed that the unstressed syllable, although its underlying tone moves to another position in the output, is not toneless on the surface. We proposed in chapter 4 that the low tone target on the initial syllable can be accounted for as a low boundary tone aligned with the edge of a prosodic domain. We provide some justification for this analysis here by comparing it with an alternative. Rose (1990) and Chen (2000) suggested that the low tone target in Zhenhai is due to a default tone. In our theory, a boundary tone is aligned with a prosodic edge. The default tone hypothesis predicts that a default tone is assigned to all toneless syllables, whereas the boundary tone hypothesis predicts that a boundary tone is inserted at the edge of a prosodic domain. We have demonstrated in chapter 3 that in Mandarin Chinese a low tone target appears only at the end of a prosodic word when there are a string of neutral tone syllables. Non-final neutral tone syllables do not get any pitch target. The same argument can be extended to Zhenhai as well. Admittedly, for disyllabic sequences it is impossible to tease apart the two hypotheses for Zhenhai because there is only one unstressed syllable in disyllabic words. When we move to the discussion of trisyllabic words, it will become clear that the analysis adopting the idea of a low boundary tone insertion is more appropriate. Following our analysis for Mandarin neutral tone, we propose the following boundary tone alignment constraint:

(11) BND(L)/EDGE: An unstressed syllable at the edge of a prosodic domain coincides
     with a low tone target.

This constraint basically belongs to the COINCIDE family. It has the effect of aligning a low boundary tone with the unstressed syllable at the edge of a prosodic domain. It does not dictate anything about stressed syllables or unstressed syllables not at the edge.

BND(L)/EDGE is unviolated in Zhenhai. In disyllabic words, when both syllables are long, the initial contour tone moves to the second stressed syllable, leaving behind a low tone target on the initial syllable. Since a low boundary tone is inserted on the unstressed initial syllable, BND(L)/EDGE has to crucially rank higher than DEP(TONE), which penalizes tone insertion: BND(L)/EDGE >> DEP(TONE). This is illustrated in the tableau below. The input is the same as in (10). We do not consider candidates which delete the first tone, but keep the second tone. They are ruled out by MAX(TONE)/INITIAL.

(12) BND(L)/EDGE >> DEP(TONE)

tone 1 σ_1 $\acute{\sigma}_2$ + | | tone 2 F_1 R_2	BND(L)/EDGE	DEPT(TONE)
σ_1 $\acute{\sigma}_2$ a. | F_1	*!	
σ_1 $\acute{\sigma}_2$ b. ☞ | | L- F_1		*

BND(L)/EDGE requires that an unstressed syllable at one prosodic edge is aligned with a low boundary tone (marked by L-). Tone movement occurs in both (a) and (b), but (b) satisfies BND(L)/EDGE at the cost of violating DEP(TONE).

After a low boundary tone is inserted, both syllables have tonal specifications. Recall that deletion of non-initial tones is achieved by MAX(TONE)/INITIAL >> *TONE in the constraint system we have developed. The following tableau seems to suggest that

deletion of the second tone and movement of the first tone are simply driven by BND(L)/EDGE.

(13) BND(L)/EDGE, MAX(TONE)/INITIAL >> *TONE

tone 1 σ_1 $\acute{\sigma}_2$ + tone 2 F_1 R_2	BND(L)/EDGE	MAX(TONE)/INITIAL	*TONE
a. σ_1 $\acute{\sigma}_2$ F_1 R_2	*!		**
b. σ_1 $\acute{\sigma}_2$ L- R_2		*!	**
c. ☞ σ_1 $\acute{\sigma}_2$ L- F_1			**

The most faithful candidate (a) is ruled out by BND(L)/EDGE, and the candidate satisfying IDENT(TONE)/$\acute{\sigma}$ is ruled out by MAX(TONE)/INITIAL. (c) is the most unfaithful mapping with tone movement and L- boundary tone insertion, but it satisfies both MAX(TONE)/INITIAL and BND(L)/EDGE.

Although the correct tonal pattern is derived in the above example, BND(L)/EDGE is not the only driving force behind tone deletion and tone movement. When we get to the trisyllabic combinations, it will become clear that once MAX(TONE)/INITIAL and BND(L)/EDGE are both satisfied, *TONE will be responsible for eliminating tonal specifications on non-peripheral unstressed syllables, which are not protected by MAX(TONE)/INITIAL, nor filled in with a boundary tone. It is conceivable that BND(L)/EDGE has to outrank *TONE so that the inserted boundary tone survives in the output.

The next question is why does the initial tone move to the stressed syllable? To see the question more clearly, we consider a trisyllabic pattern in which the first tone is preserved and the third syllable is stressed (see section 5.3.3 for more details).

(14) W-W-S: LM (tone 4) + T^l + T^l → L- + Φ + MH (T^l=T in low register, Φ=toneless)

	σ_1 σ_2 $\acute{\sigma}_3$ | | | R_1 F_2 F_3	BND(L)/EDGE	MAX(TONE)/INITIAL	*TONE
a.	σ_1 σ_2 $\acute{\sigma}_3$ | | | R_1 F_2 F_3	*!	✓	***
b.	σ_1 σ_2 $\acute{\sigma}_3$ | | | L- F_2 F_3	✓	*!	***
c.	σ_1 σ_2 $\acute{\sigma}_3$ | | | L- F_2 R_1	✓	✓	***!
d. ☞	σ_1 σ_2 $\acute{\sigma}_3$ | | L- R_1	✓	✓	**
e. ☞	σ_1 σ_2 $\acute{\sigma}_3$ | | L- R_1	✓	✓	**

(a) and (b) violate either BND(L)/EDGE or MAX(TONE)/INITIAL. In addition, they both fail to delete any tone from the input, incurring most violations of *TONE. (c), (d) and (e) pass through these two candidates, but (c) loses to (d) and (e) for having more violations of *TONE. *TONE is active in excluding (c). (d) and (e) are both possible outputs: L-boundary tone is inserted on the unstressed syllable at the prosodic edge and the initial tone in the input is successfully preserved in the output. However, there is a crucial difference between (d) and (e): the preserved initial tone is realized on the stressed syllable in (e). Movement of the initial contour tone to the stressed syllable has to be

driven by other constraints, i.e. contour tone licensing constraints. Specifically,

COINCIDE($\acute{\sigma}$, CONTOUR) restricts a contour tone to the stressed syllable. It is violated by

(d) and satisfied by (e). The expected output is (e).

In the examples we have discussed, the initial tone is retained and it moves to the

stressed syllable, which is true for long-long combinations in W-S tone sandhi. When

checked syllables are involved (either initial or final), the initial tone is invariably

suppressed on the surface. We discuss three specific cases: long-short, short-long, and

short-short combinations.

In long-short combinations, the underlying contour tone of the first syllable does

not surface on the second syllable, for example:

(13)   HL (tone 1) + L$\textrm{?}$ (tone 6) → L- + H$\textrm{?}$, *L- + HL
       cf. HL (tone 1) + MH (tone 2) → L- + HL

(14) *CONTOUR/SHORT >> COINCIDE($\acute{\sigma}$, CONTOUR) ($\acute{\sigma}_2$ = short)

tone 1 σ_1 $\acute{\sigma}_2$ + | | tone 6 F_1 L_2	*CONTOUR/SHORT	COINCIDE($\acute{\sigma}$, CONTOUR)

a. σ_1 $\acute{\sigma}_2$
      | |
    L- F_1
	*!	

b. ☞ σ_1 $\acute{\sigma}_2$
      | |
    L- H
		*

In the theory we are developing, the inability of short syllables to bear a contour tone is

due to a contour licensing constraint *CONTOUR/SHORT, which penalizes a contour tone

on a phonologically short syllable. The ranking argument for *CONTOUR/SHORT and

COINCIDE($\acute{\sigma}$, CONTOUR) is furnished by the fact that COINCIDE($\acute{\sigma}$, CONTOUR) is violated

by the winner (b) but satisfied by the loser (a), which violates *CONTOUR/SHORT for realizing a contour tone on a short syllable. Since it is more important to satisfy *CONTOUR/SHORT than to realize a contour tone on a stressed syllable (but phonologically short), *CONTOUR/SHORT dominates COINCIDE($\acute{\sigma}$, CONTOUR). No ranking is established for COINCIDE($\acute{\sigma}$, CONTOUR) and MAX(TONE)/INITIAL.

Movement of the initial tone is blocked by ranking *CONTOUR/SHORT above MAX(TONE)/INITIAL so that the initial tone is suppressed from appearing in the output. In the next tableau, candidate (a) satisfies the positional faithfulness constraint, MAX(TONE)/INITIAL, but it violates *CONTOUR/SHORT for realizing a contour tone on a short syllable. (b) satisfies the high ranking contour licensing constraint at the cost of violating MAX(TONE)/INITIAL.

(15) *CONTOUR/SHORT >> MAX(TONE)/INITIAL ($\acute{\sigma}_2$ = short)

tone 1 σ_1 $\acute{\sigma}_2$ + $\vert$ $\vert$ tone 6 F_1 L_2	*CONTOUR/SHORT	MAX(TONE)/INITIAL
a. σ_1 $\acute{\sigma}_2$ $\vert$ $\vert$ L- F_1	*!	
b. ☞ σ_1 $\acute{\sigma}_2$ $\vert$ $\vert$ L- H		*

Since the stressed syllable defaults to H when it is phonologically short, COINCIDE($\acute{\sigma}$, H) is invoked to ensure that the short stressed syllable has a H level tone when it is not able to bear a contour tone (due to the high-ranking *CONTOUR/SHORT).

This constraint has to outrank IDENT(TONE)/ó́ so that the underlying tone of the short stressed syllable does not surface. The tableau below illustrates this ranking.

(16) COINCIDE(ó́, H) >> IDENT(TONE)/ó́ (ó́$_2$ = short)

tone 1 + tone 6 σ$_1$ ó́$_2$ | | F$_1$ L$_2$	COINCIDE(ó́, H)	IDENT(TONE)/ó́
a. σ$_1$ ó́$_2$ | | L- L$_2$	*!	
b. ☞ σ$_1$ ó́$_2$ | | L- H		*

The last crucial candidate we consider is the one in which the initial tone is preserved and realized on the initial syllable and the stressed short syllable defaults to H. BND(L)/EDGE excludes such a candidates at the cost of violating the context-free IDENT(TONE).

(17) BND(L)/EDGE >> IDENT(TONE)

tone 1 + tone 6 σ$_1$ ó́$_2$ | | F$_1$ L$_2$	BND(L)/EDGE	IDENT(TONE)
a. σ$_1$ ó́$_2$ | | F$_1$ H	*	*
b. ☞ σ$_1$ ó́$_2$ | | L- H		**

The winner (b) violates MAX(TONE)/INITIAL (not shown) for failing to preserve the initial tone F$_1$. It does not violate other contour tone licensing constraints we covered earlier.

Consequently there seems to be a ranking argument for BND(L)/EDGE >> MAX(TONE)/INITIAL. However, since (a) also violates an undominated constraint *CONTOUR/UNSTRESSED σ, this example does not constitute a real ranking argument.

To summarize the long-short combinations, the crucial fact is that the initial tone is suppressed when its eventual host (i.e. the second syllable, phonological short) is not able to carry a contour tone (due to *CONTOUR/SHORT). The stressed short syllable defaults to H. What happens here is that neither the initial tone nor the underlying tone of the stressed syllable appears on the surface.

In short-long combinations in the W-S pattern, the second syllable surfaces with its underlying tone and the first syllable with a short level tone, as in the following example:

(18)   L? (tone 6) + MH (tone 2) → L- + MH, *L- + L

We have established the ranking for MAX(TONE)/INITIAL >> IDENT(TONE)/σ́ on the basis of long-long combinations. However such a ranking would yield the wrong output for short-long combinations. The generalization is that when the first syllable is short, its underlying tone (a level tone) does not move to the second syllable, which is stressed and phonologically long. Recall that COINCIDE(σ́, CONTOUR) is responsible for restricting a contour tone to a stressed syllable. This constraint is only concerned with the surface tonal distributions, indifferent to the source of a contour tone. It has to override MAX(TONE)/INITIAL to allow the underlying contour tone of the stressed syllable to surface in the output. Consider the following tableau.

(19) $\text{COINCIDE}(\acute{\sigma}, \text{CONTOUR}) \gg \text{MAX}(\text{TONE})/\text{INITIAL} \gg \text{IDENT}(\text{TONE})/\acute{\sigma}$ (σ_1 = SHORT)

tone 6 σ_1 $\acute{\sigma}_2$ + tone 2 $\;$L₁ R₂	COINCIDE($\acute{\sigma}$, CONTOUR)	MAX(TONE)/INITIAL	IDENT(TONE)/$\acute{\sigma}$
a. $\quad\sigma_1$ $\acute{\sigma}_2$ $\qquad$|$\quad$| $\qquad$L- L₁	*!		
b. $\quad\sigma_1$ $\acute{\sigma}_2$ $\qquad$|$\quad$| $\qquad$L- F		*	*!
c. ☞ $\;\sigma_1$ $\acute{\sigma}_2$ $\qquad$|$\quad$| $\qquad$L- R₂		*	

Candidate (a) preserves the initial tone, but it fails to supply a contour tone to the stressed syllable. (b) and (c) both fail to perverse the initial tone. (b) loses to (c) because it changes the tonal specifications of the second syllable. (b) also violates DEP(TONE) (not shown) because of the general-purpose tone insertion (except boundary tone insertion) in the output.

We have shown that COINCIDE($\acute{\sigma}$, H) outranks IDENT(TONE)/$\acute{\sigma}$. COINCIDE($\acute{\sigma}$, CONTOUR) has to dominate COINCIDE($\acute{\sigma}$, H) so that a contour tone rather than a level H appears on the stressed syllable if it is long (i.e. not subject to *CONTOUR/SHORT). In the tableau below, (a) satisfies COINCIDE($\acute{\sigma}$, H) by simplifying a contour rise to a level H. It is excluded by the high-ranking COINCIDE($\acute{\sigma}$, CONTOUR). It is worth mentioning that although *CONTOUR/SHORT dominates COINCIDE($\acute{\sigma}$, CONTOUR), it is irrelevant when the stressed syllable is phonologically long, as in short-long combinations.

(20) COINCIDE($\acute{\sigma}$, CONTOUR) >> COINCIDE($\acute{\sigma}$, H) (σ_1 = short)

tone 6 σ_1 $\acute{\sigma}_2$ + | $\wedge$ tone 2 $L_1 L_2 H_2$	COINCIDE ($\acute{\sigma}$, CONTOUR)	COINCIDE($\acute{\sigma}$, H)	IDENT(TONE)/$\acute{\sigma}$
b. σ_1 $\acute{\sigma}_2$ | | L- H_2	*!		*
c. ☞ σ_1 $\acute{\sigma}_2$ | $\wedge$ L- $L_2 H_2$		*	

Before leaving the discussion of tone movement, we consider a possible candidate which satisfies COINCIDE($\acute{\sigma}$, CONTOUR) and MAX(TONE)/INITIAL simultaneously by fusing L_1 and H_2 into a rise on the second syllable, as shown in the following tableau.

(21) COINCIDE($\acute{\sigma}$, CONTOUR) >> MAX(TONE)/INITIAL >> IDENT(TONE)/$\acute{\sigma}$ (σ_1 = short)

tone 6 σ_1 $\acute{\sigma}_2$ + | $\wedge$ tone 2 $L_1 L_2 H_2$	COINCIDE ($\acute{\sigma}$, CONTOUR)	MAX(TONE)/INITIAL	IDENT(TONE)/$\acute{\sigma}$
a. ☜ σ_1 $\acute{\sigma}_2$ | $\wedge$ L- $L_1 H_2$			*!
b. σ_1 $\acute{\sigma}_2$ | $\wedge$ L- $L_2 H_2$		*	

Candidate (a) satisfies COINCIDE($\acute{\sigma}$, CONTOUR) and MAX(TONE)/INITIAL through a fusion process. Given the ranking we have arrived at, (a) will be picked as the optimal candidate.   The actual output is (b).   Phonetically speaking, (a) and (b) are

indistinguishable. However when the second syllable carries an underlying fall, a rising tone could be erroneously generated on the second syllable through the same fusion process, shown below.

(22) COINCIDE($\acute{\sigma}$, CONTOUR) >> MAX(TONE)/INITIAL >> IDENT(TONE)/$\acute{\sigma}$ (σ_1 = short)

tone 6 + tone 1 $\begin{array}{cc}\sigma_1 & \acute{\sigma}_2 \\ | & \wedge \\ L_1 & H_2\ L_2 \end{array}$	COINCIDE ($\acute{\sigma}$, CONTOUR)	MAX(TONE)/INITIAL	IDENT(TONE)/$\acute{\sigma}$
a. ☞ $\begin{array}{cc}\sigma_1 & \acute{\sigma}_2 \\ | & \wedge \\ L\text{-}L_1 & H_2 \end{array}$			*!
b. $\begin{array}{cc}\sigma_1 & \acute{\sigma}_2 \\ | & \wedge \\ L\text{-} & H_2\ L_2 \end{array}$		*	

A solution to this problem lies in that fact that there is a preference for preservation of contour tones when they are present underlyingly, but no way to create them. As we discussed in chapter 4, tone movement in Zhenhai exhibits a process in which the contour node (c_1) of the initial tone moves to the second syllable, leaving behind the register node (r_1), as shown below.

(23) $\quad \sigma_1 \qquad \sigma_2 \quad \rightarrow \quad \sigma_1 \qquad \sigma_2 \quad \rightarrow \quad \sigma_1 \qquad \sigma_2$

$\quad\quad \wedge \qquad \wedge \qquad\quad \wedge \qquad \wedge \qquad\quad\quad \wedge \qquad \wedge$

$\quad\quad r_1\ c_1 \quad r_2\ c_2 \qquad r_1\ c_1\ r_2\ c_2 \qquad\quad r_1\ c \quad r\ c_1$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad | \quad\quad |$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\ L\text{-} \quad h$
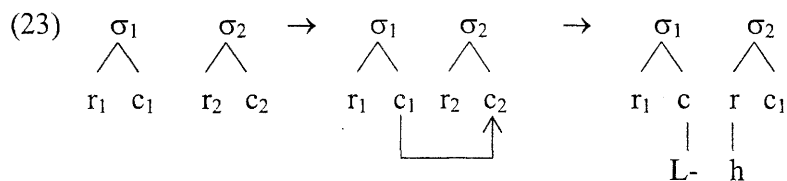
The resulting rise from the fusion of L_1 with H_2 can be either dominated by c_1 or c_2. When it is dominated by c_1, it violates DEP(TARGET)/INITIAL which bans insertion of a terminal tone feature under the contour node of the initial tone. We use TARGET for the contour node (c) in the representation of a tone. When it is dominated by c_2,

194

MAX(TONE)/INITIAL is violated because there is no c_1 in the output. In addition,

IDENT(TONE)/ớ is also violated. The following tableau illustrates.

(24) DEP(TARGET)/INITIAL, MAX(TONE)/INITIAL >> IDENT(TONE)/ớ (σ_1 = short)

tone 6 + tone 1 σ_1 $\acute{\sigma}_2$ | | c_1 c_2 | $\bigwedge$ L_1 H_2 L_2	DEP(TARGET)/ INITIAL	MAX(TONE)/ INITIAL	IDENT(TONE)/ớ
a. σ_1 $\acute{\sigma}_2$ | | c c_1 | $\bigwedge$ L- L_1 H_2	*		*!
b. σ_1 $\acute{\sigma}_2$ | | c c_2 | $\bigwedge$ L- L_1 H_2		*	*!
c. ☞ σ_1 $\acute{\sigma}_2$ | | c c_2 | $\bigwedge$ L- H_2 L_2		*	

To summarize the short-long combinations, although MAX(TONE)/INITIAL requires the initial tone to be preserved in the output, its effect can be overridden when the initial tone does not contribute a contour tone in response to the high-ranking contour licensing constraint COINCIDE(ớ, CONTOUR). When that happens, the stressed syllable, if phonologically long, surfaces with its own underlying tone.

The ranking of COINCIDE(ớ, H) and MAX(TONE)/INITIAL can be determined from the tonal patterns in short-short combinations, for example:

(25)   L? (tone 6) + L? (tone 6) → L- + H, *L- + L

This example demonstrates that when the initial tone is underlyingly a short level tone, the second syllable, stressed but also phonologically short, still defaults to H rather than taking the initial tone. Thus the following ranking ensues:

(26)   COINCIDE($\acute{\sigma}$, H) >> MAX(TONE)/INITIAL (σ_1, $\acute{\sigma}_2$ = short)

tone 6 σ_1 $\acute{\sigma}_2$ + | | tone 6 L_1 L_2	COINCIDE($\acute{\sigma}$, H)	MAX(TONE)/INITIAL

a. σ_1 $\acute{\sigma}_2$
      | |
     L- L_1
	*!	

b. ☞ σ_1 $\acute{\sigma}_2$
      | |
     L- H
		*

Candidate (a) moves the initial tone to the second syllable whereas candidate (b) simply defaults to H.

Since COINCIDE($\acute{\sigma}$, H) has been invoked to insert a H on the phonologically short syllable, it has to dominate DEP(TONE) to allow tone insertion to happen.

(27)   COINCIDE($\acute{\sigma}$, H) >> DEP(TONE) (σ_1, $\acute{\sigma}_2$ = short)

tone 6 σ_1 $\acute{\sigma}_2$ + | | tone 6 L_1 L_2	COINCIDE($\acute{\sigma}$,H)	DEP(TONE)

a. σ_1 $\acute{\sigma}_2$
      | |
     L- L_1
	*!	*(L-)

b. ☞ σ_1 $\acute{\sigma}_2$
      | |
     L- H
		**(L-, H)

The rankings we have established for the most relevant constraints are summarized below:

(28)   Constraints and their rankings (for tonal contour)

*CONTOUR/SHORT
|
COINCIDE(σ, CONTOUR)
|
BND(L)/EDGE   COINCIDE(σ, H)
| _____|
DEP(TONE)   MAX(TONE)/INITIAL
|_____
IDENT(TONE)/σ *TONE
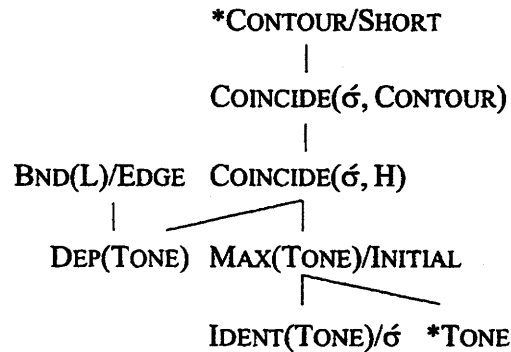
We have discussed how the constraint system determines which tone is retained and where it surfaces eventually. Now we discuss how the tone registers on the two syllables are specified in our theory. We saw that the initial syllable, though unstressed, still retains its underlying specification for register. This can be seen clearly by comparing the f0 tracks in (29) and (30), repeated from chapter 4. The inserted low boundary tone is realized as a mid level tone when the underlying register of the initial syllable is high (29), and as a low level tone when the register is low (30).

(29)   Pattern A: Tone 1 + *Ti* (i = 1/2, 3/4, 5, 6)
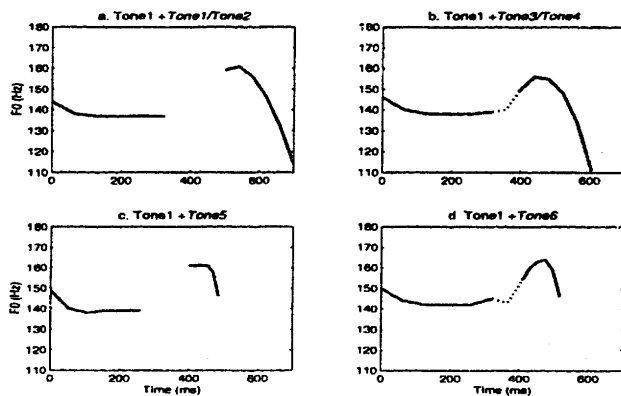
(30)   Pattern B: Tone 3 + *Ti* (i = 1/2, 3/4, 5, 6)
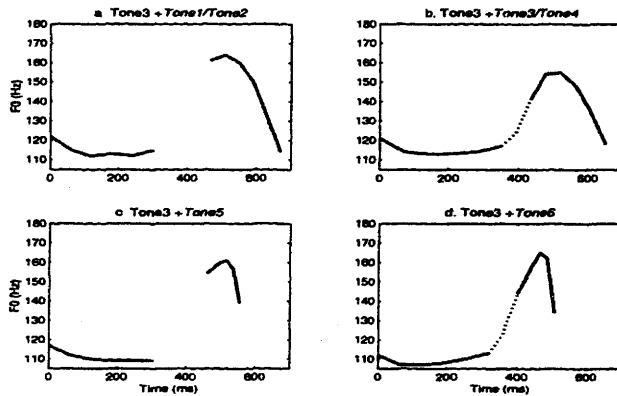
We attribute this difference to another positional faithfulness constraint, IDENT(REG)/INITIAL. In addition, the tonal contour realized on the prominent, second syllable is always in high register, due to a positional markedness constraint, REG(H)/ớ, which is grounded in the cross-linguistic preference for the prominent syllable to be realized in higher pitch (cf. de Lacy 2002). The relevant constraints are defined as follows:

(31) a. IDENT(REG)/INITIAL: if α is a tone-bearing unit in the domain-initial position in the input and β is a correspondent of α in the output, then the tone register specification of α must be identical to the tonal specification of β.

   b. REG(H)/ớ: the tone on the accented syllable must be in high register.

   c. IDENT(REG)/ớ: if α is a tone-bearing unit of the stressed syllable in the input and β is a correspondent of α in the output, then the tone register specification of α must be identical to the tonal specification of β.

d. IDENT(REG): if α is a tone-bearing unit in the input and β is a correspondent of α in the output, then the tone register specification of α must be identical to the tonal specification of β.

As the next tableau shows, REG(H)/σ́ dominates IDENT(REG)/σ́ to neutralize the underlying register specification of the stressed syllable to high.

(32) HL (tone 1) + LM (tone 4) → M- + MH, *L- + LM

(33) IDENT(REG)/INITIAL, REG(H)/σ́ >> IDENT(REG)/σ́

σ_1 $\sigmá_2$ Tone 1 /\ /\ + r c r c Tone 4 | | | | h F$_1$ l R$_2$	IDENT(REG)/INITIAL	REG(H)/σ́	IDENT(REG)/σ́
a. ☞ σ_1 $\sigmá_2$ /\ /\ r c r c | | | | h L- h F$_1$			*
b. σ_1 $\sigmá_2$ /\ /\ r c r c | | | | h L- l F$_1$		*!	
c. σ_1 $\sigmá_2$ /\ /\ r c r c | | | | l L- h F$_1$	!*		*

Although no crucial ranking argument emerges for IDENT(REG)/INITIAL and REG(H)/σ́ in this example, there is evidence, to be discussed in the next subsection, that IDENT(REG)/INITIAL actually dominates REG(H)/σ́. When stress falls on the initial

syllable as in the S-W tone sandhi patterns, the surface tonal contour on the initial syllable is faithful to the underlying register specification. For the moment, we leave these two constraints unranked.

As before, R and F in the tableau are used as shorthand for rising and falling tones. The register specifications are correctly derived on both syllables. IDENT(REG)/INITIAL rules out (c), which does not keep the underlying high register of the initial syllable. The underlying contour of the initial tone, once realized on the stressed syllable, is in high register. Thus (b) has a fatal violation of REG(H)/ó, although it is most faithful to the underlying register specifications in the input. (a) only violates the low-ranking IDENT(REG)/ó and it is the winner. The inserted low boundary tone in (a) surfaces as M- (i.e. high-register L-) because the underlying register of the initial syllable is high.

The next example further establishes the effect of register neutralization on the stressed syllable. When the initial tone is originally in low register, it will be realized in high register after being shifted to the stressed syllable, as the following tableau shows. Both input syllables are in low register, but when the initial tone is realized on the stressed syllable, it takes on a high register, as compelled by REG(H)/ó. Therefore, a low falling tone on the first syllable in (a) is mapped to a high falling tone on the second syllable. (a) is the output. The competing candidate (b) is faithful to the register specifications of the two syllables in the input, thereby violating the high-ranking REG(H)/ó.

(34)   ML (tone 3) + LM (tone 4) → L- + HL, *L- + ML

(35)   IDENT(REG)/INITIAL, REG(H)/σ́ >> IDENT(REG)/σ́

Tone 3 + Tone 4 σ_1 σ_2 $\wedge$ $\wedge$ r c r c | | | | 1 F_1 1 R_2	IDENT(REG)/INITIAL	REG(H)/σ́	IDENT(REG)/σ́
a. ☞ σ_1 σ_2 $\wedge$ $\wedge$ r c r c | | | | 1 L- h F_1			*
b. σ_1 σ_2 $\wedge$ $\wedge$ r c r c | | | | 1 L- 1 F_1		*!	

We assume that the low boundary tone has a low register specification in light of the following positional markedness constraint. This constraint specifies the unmarked register for a tone.

(36)   REG(L)/EDGE: A tone must be in low register at the edge of a prosodic domain.

In our analysis, REG(L)/EDGE is dominated by IDENT(REG)/INITIAL so that the inserted low boundary tone will get the underlying register of the initial syllable. It is also outranked by REG(H)/σ́: the stressed syllable at the edge of a prosodic domain will get a high register.

(37) IDENT(REG)/INITIAL, REG(H)/σ́ >> REG(L)/EDGE

Tone 1 + Tone 4: σ₁(r c: h F₁) σ₂(r c: l R₂)	IDENT(REG)/INITIAL	REG(H)/σ́	REG(L)/EDGE
a. ☞ σ₁(r c: h L-) σ₂(r c: h F₁)			**
b. σ₁(r c: h L-) σ₂(r c: l F₁)		*!	*
c. σ₁(r c: l L-) σ₂(r c: h F₁)	*!		*

The winner (a) satisfies the two high-ranking constraints, but it violates REG(L)/EDGE twice because in a disyllabic word both syllables are at a prosodic edge and both are in high register. The other two candidates are the same as in (33). REG(L)/EDGE assesses a violation when the syllable at a prosodic edge is not in low register.

The other tone combinations (i.e. long + short, short + long, and short + short) will have their tone register determined in the same way.

The rankings are summarized in (38). They reflect the following generalizations. When the second syllable is stressed, it requires its register to be high so that the retained contour tone is realized in high register. The first syllable keeps its underlying register specification no matter whether it is stressed or not.

(38) Constraints and their ranking (for tone register)

$$\text{IDENT(REG)/INITIAL}$$

$$|$$

$$\text{REG(H)/}\acute{\sigma}$$

$$\text{REG(L)/EDGE} \qquad \text{IDENT(REG)/}\acute{\sigma}$$
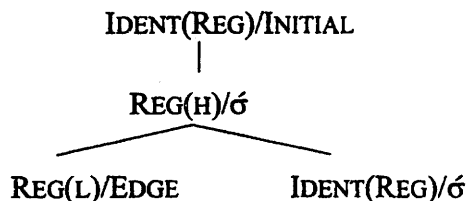
The ranking argument for IDENT(REG)/INITIAL and REG(H)/$\acute{\sigma}$ will be provided in the next subsection.

To summarize, in the dual-prominence theory surface tone patterns are determined by the interaction of positional faithfulness and positional markedness constraints. The main observation we have made is that which tone is retained and where it will eventually surface are not determined by the position of stress alone. Rather they result from the interaction of two prominent positions (i.e. peripheral and metrically prominent positions) in a prosodic domain, both being referred to by positional faithfulness and positional markedness constraints. Zhenhai W-S tone sandhi patterns provide such a case in point. In Zhenhai, we have identified two kinds of prominence: initial-position prominence and metrical prominence. The initial position preserves its underlying tone and also has longer vowel duration than the non-initial position. It maintains its underlying register specification. The stressed syllable is associated with the contrasting tonal contours, realized in high register. These two prominent positions are both referred to by the constraints we have proposed. The positional faithfulness constraint, MAX(TONE)/INITIAL, keeps the initial tone in the input, while the contour licensing constraints jointly determine the contour shape on the stressed syllable. The surface patterns emerge out of the interaction of positional faithfulness and positional

203

markedness constraints. Other cases of tone movement, like the leftward tone movement in Wenzhou, can be accounted for in similar fashion.

In the next subsection, we will show that when tone preservation and stress occur in one place, there will be no tone movement. The retained tone surfaces in its original position. This is what we observe in the S-W patterns of disyllabic tone sandhi in Zhenhai. The constraint system we have constructed so far will be shown to correctly predict the S-W tone sandhi patterns.

5.3.2   S-W Disyllabic Tone Sandhi

The tone mapping in the S-W disyllabic patterns is summarized in the following table, where T stands for any of the six citation tones in the input. The stress falls on the first syllable in S-W tone sandhi.

(39) Tone mapping in disyllabic tone sandhi

patterns	A: T1 + T	B: T2+ T	C: T4 + T	D: T5 + T
input	HL – T	MH – T	LM – T	H? – T
output	H – L	MH – L	LM – L	H? – L

Except for pattern A, the initial tone is faithfully preserved in the output. The second syllable loses its underlying tone completely. It takes on a low boundary tone, as we proposed for the W-S patterns. The underlying specification for tone register is preserved on the first syllable; it is neutralized to a low register on the second syllable.

In the S-W patterns, the initial-position prominence and metrical prominence coincide. As we mentioned earlier, the initial tone is attracted to the stressed syllable in Zhenhai. When stress is initial, the initial tone surfaces on the initial syllable. The

single-prominence theory is able to explain this pattern by saying that the stressed syllable (i.e. initial in this case) keeps its lexically associated tone. However, as we have already illustrated with the W-S patterns, the initial tone preservation and tone attraction to the stressed syllable are two independent processes.

The surface tone patterns in S-W tone sandhi follow from the ranking of the positional faithfulness and positional markedness constraints that we have already established. Putting aside pattern A for a moment, let us consider patterns B, C and D first.

Patterns B and C can be derived similarly in the following tableaux:

(40)   MH (tone 2) + ML (tone 3) → MH + L-, *ML + L-

tone 2 + tone 3 σ́₁ σ₂ / R₁ F₂	BND(L)/EDGE	COINCIDE (σ́, CONTOUR)	COINCIDE (σ́, H)	MAX(TONE)/INITIAL	IDENT (TONE)/σ́
a. σ́₁ σ₂ / R₁ F₂	*!		*		
b. σ́₁ σ₂ / F₂ L-			*	*!	*
c. σ́₁ σ₂ / L₁ L-		*!		*	*
d. ☞ σ́₁ σ₂ / R₁ L-			*		

Only relevant constraints are shown in the tableau. Candidate (a) is faithful to the input, but it fails to replace the underlying tone of the final syllable with a low boundary tone. (c) simplifies the initial R to L in violation of another high-ranking constraint

COINCIDE($\acute{\sigma}$, CONTOUR). It is also considered a violation of MAX(TONE)/INITIAL. The two remaining candidates (b) and (d) both violate COINCIDE($\acute{\sigma}$, H). (b) is excluded by the positional faithfulness constraint MAX(TONE)/INITIAL. (d) is the winner. The same analysis also works for the cases in which the second syllable is short.

Recall that in short-long combinations in the W-S patterns, the second syllable surfaces with its own underlying tone at the expense of violating MAX(TONE)/INITIAL. We have shown that the violation is compelled by COINCIDE($\acute{\sigma}$, CONTOUR), which requires a contour tone to appear on a stressed syllable. For the short-long combinations in the S-W patterns, the same ranking of the constraints gives rise to a different tonal output as a result of the different stress position (i.e. initial in S-W patterns).

(41) H? (tone 5) + MH (tone 2) → H + L-, *MH + L- ($\acute{\sigma}_1$ = short)

tone 5 + tone 2 $\acute{\sigma}_1$ σ_2 | | H_1 R_2	*CONTOUR /SHORT	COINCIDE ($\acute{\sigma}$, CONTOUR)	MAX(TONE) /INITIAL
a. $\acute{\sigma}_1$ σ_2 | | R_2 L-	*!		*
b. ☞ $\acute{\sigma}_1$ σ_2 | | H_1 L-		*	

We only consider candidates that satisfy BND(L)/EDGE because it is unviolated in Zhenhai. Candidate (a) is penalized by the high-ranking constraint banning contour tone on short syllables. (b) has a violation of the low-ranking COINCIDE($\acute{\sigma}$, CONTOUR) in this case. What is most striking about this pattern is that the phonological short syllable (but stressed) maintains its underlying tone whereas the phonologically long syllable (but

206

unstressed) loses its underlying tone and is realized in a low boundary tone. This pattern poses a serious challenge to the theory of contour tone distribution according to which contour tone licensing constraints make crucial reference to such surface phonetic properties as duration and/or sonority (Gordon 1999, Zhang 2001).

When both syllables are short, the winner is picked by MAX(TONE)/INITIAL.

(42) H? (tone 5) + L? (tone 6) $\rightarrow$ H + L-, *L + L- ($\acute{\sigma}_1$, σ_2 = short)

tone 5 $\acute{\sigma}_1$ σ_2 + tone 6 H_1 L_2	*CONTOUR /SHORT	COINCIDE (σ, CONTOUR)	MAX(TONE) /INITIAL
a. $\quad$ $\acute{\sigma}_1$ σ_2 $\quad$ | | $\quad$ L_2 L-		*	*!
b. ☞ $\quad$ $\acute{\sigma}_1$ σ_2 $\quad$ | | $\quad$ H_1 L-		*	

Since the underlying tone register is preserved on the first syllable, IDENT(REG)/INITIAL has to outrank REG(H)/$\acute{\sigma}$, which in turn outranks IDENT(REG)/$\acute{\sigma}$. Note that since stress is initial in the S-W tone sandhi patterns, IDENT(REG)/INITIAL and IDENT(REG)/$\acute{\sigma}$ both point to the initial syllable. The register neutralization on the second syllable indicates that REG(L)/EDGE has to rank higher than the context-free IDENT(REG). Candidate (b) fatally violates IDENT(REG)/INITIAL. (a) and (c) are tied for REG(H)/$\acute{\sigma}$. The evaluation is passed onto the next constraint, REG(L)/EDGE, which picks (c) as the output. (a) fails to neutralize its register on the second syllable (i.e. at the prosodic edge) to low.

(43) LM (tone 4) + HL (tone 1) → LḾ + L-, *MH + L-

(44) IDENT(REG)/INITIAL >> REG(H)/ớ >> REG(L)/EDGE, IDENT(REG)/ớ

Tone 4 + Tone 1 ớ₁ ớ₂ /\ /\ r c r c | | | | l R₁ h F₂	IDENT(REG) /INITIAL	REG(H)/ớ	REG(L)/EDGE	IDENT(REG)/ớ
a. ớ₁ ớ₂ /\ /\ r c r c | | | | l R₁ h L-		*	*!	░░░
b. ớ₁ ớ₂ /\ /\ r c r c | | | | h R₁ l L-	*!			░*░
c. ☞ ớ₁ ớ₂ /\ /\ r c r c | | | | l R₁ l L-		*		░░░

Lastly we consider pattern A. It is different from the other three patterns in that the underlying tone of the initial syllable is not preserved in the output. Rather we observe a process of contour simplification: HL → H. The simplification does not occur in the W-S pattern in which the initial HL moves to the second syllable. Compare the following mappings:

(45) Mapping of tone 1 in S-W and W-S patterns

input	HL – T
S-W	H – L
W-S	M – HL

(T = any tone)

Recall that in the S-W patterns the tones after tone 1 are neutralized to a low boundary tone L-. Therefore, we propose that this is a case of tone absorption in which the low target of the falling tone is absorbed into the following low tone target. Tone absorption is widely attested in West African languages (Hyman and Schuh 1974) and it is also found in Tianjin dialect. One of the tone sandhi processes is described as changing a falling tone to high level tone before a low level tone (Li and Liu 1985). For the moment, we treat it as a markedness constraint, *FL.[3] It has to be ranked higher than COINCIDE(σ, CONTOUR) so that the initial syllable can surface with a high level tone.

(46)   HL (tone 1) + ML (tone 3) → H + L-, *HL + L-

tone 2 σ̂₁ σ₂ + | | tone 3 F₁ F₂	*FL	COINCIDE (σ, CONTOUR)	COINCIDE(σ, H)

a. σ̂₁ σ₂
    | |
    F₁ L-
	*!		*

b. σ̂₁ σ₂
    | |
    L₁ L-
		*	*!

c. ☞ σ̂₁ σ₂
    | |
    H₁ L-
		*	

When F on the stressed syllable is ruled out by *FL, COINCIDE(σ, H) picks the second best choice, H.[4]

---

[3] Kenstowicz and Akanlig-Pare (2003) illustrates a similar tone absorption process involving a rising tone in Buli: LH-H → L-H.

[4] Alternatively, (b) can also be ruled out by OCP against two adjacent low tones, which is unviolated in Zhenhai.

To summarize, the constraint system we have established correctly explains both W-S and S-W tone sandhi patterns in Zhenhai. The main feature of the theory is the differentiation of prosodic edge and stress. The two kinds of prominent positions are both called for by the phonology in the computation of the surface tone patterns in Zhenhai.

It is noticed that in the W-S patterns, the second stressed syllable inherits the underlying contour of the first syllable and realize it in high register. It is reasonable to consider the register neutralization to high as a case of enhancement (cf. Stevens and Keyser 1989): higher pitch tends to increase the perceptual saliency of stressed syllables. This is reminiscent of iambic lengthening by which the head syllable (i.e. the second syllable) in a foot is lengthened. Register enhancement does not occur in the S-W patterns when the first syllable is stressed.

5.3.3   Trisyllabic Lexical Tone Sandhi

The trisyllabic lexical tone sandhi patterns in Zhenhai provide further support for the separation of prosodic edge and stress. Since there is no description of the complete trisyllabic patterns in the literature, we will limit our discussion to four patterns, described in Rose (1990) and (1994). Although only four patterns are available, they cover all three possible stress positions in trisyllabic words. All the examples used here are lexical compounds.[5] The transcriptions are Rose's.

(47)   Trisyllabic patterns

---

[5]   No specific example is cited for pattern 2 here in Rose (1990).

	accent	input	output	example	gloss	source
1	W-W-S	T4-T^l-T^l	11-22-*34*	nau mĩ vɛ	"glutinous rice"	1990:27-28
2	S-W-W	T2-T-T	*34*-53-21			1990:31-32
3	W-S-W	T1-T-T	33-*44*-31	tɕjɔ dɛ wej	"reception"	1994
4	W-S-W	T3-T-T	11-*44*-31	nej tsz hej	"real man"	1994
			11-*34*-31	wu mĩŋ tsz	"ring finger"	

(T = any tone, T^l = tone in low register, stressed tones are in italics.)

According to Rose (1990), the surface sandhi form in the first pattern is characterized as a low level target [11] on the first syllable, a shift of an initial low rise to a high rise target [34] on the final syllable, and a transitional pitch [22] on the middle syllable. Construed in our theory, the underlying contour tone of the initial unstressed syllable moves to the stressed final syllable where it is realized in a high register. The initial syllable gets a low boundary tone, realized in its original register specification. The middle syllable is toneless, i.e. it is not associated any tonal target. The mapping is described below:

(48) LM (tone 4) + T^l + T^l → L- + Φ + MH, *L- + Φ + LM (Φ = toneless)

This example furnishes evidence for the boundary tone analysis. The default tone analysis predicts that each of the two unstressed syllables gets a default low tone whereas the boundary tone analysis only allows the unstressed syllable at the edge of a prosodic domain to get a low tone target.

In our earlier discussion, *TONE does not seem to play a significant role in the disyllabic tone sandhi patterns. In trisyllabic forms, however, it is responsible for eliminating candidates with any tonal specification on the middle syllable (49). An analysis of this pattern can be found in (14).

(49) LM (tone 4) + T^l + T^l → L- + Φ + MH

211

The second pattern is very similar to the S-W disyllabic pattern. The f0 peak in [53] at the beginning of the second syllable can be argued to be affiliated with the first syllable. In other words, it is the delayed f0 peak of the preceding rise, as observed in the disyllabic patterns in Zhenhai as well as Mandarin Chinese. The low tone target at the end of the final syllable is again the low boundary tone. The mapping is given below:

(50) MH (tone 2) + T + T → MH + Φ + L- (Φ = toneless)

In this pattern, the initial rise is preserved and realized on its original host syllable. A low boundary tone is aligned with the unstressed syllable at the right edge of the trisyllabic word.

The third and fourth patterns only differ in the register specification of the initial syllable: high in the former and low in the latter. Rose (1994) reported that when the first syllable is tone 3, the stressed middle syllable is realized in two "audibly different pitch shapes" – high level or high rise, depending on the voicing of C2. We argue that the stressed syllable has an invariable high level target as surface sandhi form. The high rise [34] is due to the contextual effect of co-articulated tones (Xu 1997). From the perspective of tone implementation, when a high tone target [44] is preceded by a low tone target [11] and if both are aligned at the end of their respective host syllables, a rising f0 transition is inevitable on the second syllable. The high tone target is fully approximated at the end of its host syllable. This period of f0 transition is always there when C2 is voiced, especially a sonorant. Figure 3 in Rose (1994) clearly shows this effect of tone co-articulation. When the first syllable is in tone 1, the f0 transition from [33] to [44] is short and inaudible because the f0 hits its target right at the beginning of the second syllable, as shown in the Figure 2 in Rose (1996).

With that clarified, the surface sandhi forms are unremarkable in the theory we have proposed. The mappings for the third and fourth patterns follow:

(51) a. HL (tone 1) + T + T → M- + H + L-
b. ML (tone 3) + T + T → L- + H + L-

Note that the stress is on the middle syllable, so the syllables at both edges are unstressed. Consequently, each is aligned with a low tone target, as driven by BND(L)/EDGE. In addition, our theory predicts that the boundary tone in the initial position will get the original register specification of the initial syllable because of the high-ranking IDENT(REG)/INITIAL. The register of the boundary tone on the final syllable is in the default low register because of REG(L)/EDGE. Finally, REG(H)/σ́ will compel the stressed syllable to be in high register, which leads to the complete tonal neutralization on the middle syllable. The following tableau illustrates:

(52) ML (tone 3) + T + T → L- + H + L-

σ_1 $\acute\sigma_2$ σ_3 /\ /\ /\ r c r c r c | | | | | | l F_1 l F_2 h R_3	REG(H)/σ́	IDENT(REG)/σ́
a. σ_1 $\acute\sigma_2$ σ_3 /\ /\ /\ r c r c r c | | | | | | l L- l H_2 l L-	*!	
b. ☞ σ_1 $\acute\sigma_2$ σ_3 /\ /\ /\ r c r c r c | | | | | | l L- h H_2 l L-		*

213

In the above tableau, (a) is ruled out by REG(H)/$\acute{\sigma}$ because the middle stressed syllable fails to neutralize its input register to high. The winner is (b) in which both unstressed syllables (initial and final) default to low register.

The tonal output for this pattern also follows straightforwardly from our theory.

(53) ML (tone 3) + T + T → L- + H + L-

	σ_1 $\acute{\sigma}_2$ σ_3 | | | F_1 R_2 R_3	*FL	COINCIDE ($\acute{\sigma}$, CONTOUR)
a.	σ_1 $\acute{\sigma}_2$ σ_3 | | | L- F_1 L-	*!	
b. ☞	σ_1 $\acute{\sigma}_2$ σ_3 | | | L- H_1 L-		*

As we have demonstrated, the falling tone simplification is triggered by *FL when a fall is followed by a level low tone. This is exactly what is observed in this pattern.

So far we have accounted for the four available tone sandhi patterns in trisyllabic lexical compounds. It remains to see whether our theory is able to handle the other trisyllabic combinations which should be published in future research. Our general prediction is clear: the initial tone will be preserved and shifted to the stressed syllable when phonologically long syllables are involved. More specific predictions can be readily made regarding register neutralization and contour tone distributions in Zhenhai.

A remaining issue is that we did not find clear evidence for the initial syllable lengthening in the trisyllabic patterns. Rose (1990) only provides transcriptions for the first two trisyllabic patterns we have discussed. Estimating from the two f0 plots in Rose (1994), we find that the vowels of the first two syllables have similar durations, on the

214

order of 170 ms to 220 ms, much smaller than what we observed in the disyllabic patterns. We will leave this issue open. However, as we have shown before, it is clear that the initial tone moves to the stressed position, and the initial syllable always keeps its underlying register specification regardless of whether it is stressed or not, overriding the constraint in favor of a high register on stressed syllable. In addition, even though the initial syllable does not exhibit longer duration, the stressed syllable does not show durational advantage over unstressed syllables either. Nonetheless, the retained tonal contour always seeks out the stressed syllable. All these facts point to the conclusion that the initial position represents a kind of prominence distinguishable from the metrical prominence.

5.3.4   Chen's (2000) Analysis of Zhenhai Lexical Tone Sandhi

In this subsection, we review Chen's (2000:64ff) analysis of Zhenhai tone sandhi. As we present his analysis, it will become clear that our analysis crucially differs from his in a number of aspects.

5.3.4.1   W-S Patterns

Following Rose (1990), Chen also recognizes the fact that in W-S patterns, the tonal contour (rise and fall) of the initial unstressed syllable is retained, but realized on the stressed, second syllable. The tone movement exhibits what he calls an "integrity effect" in that the whole contour tone moves as an integral unit. The tone movement is regarded as an instantiation of the Tone-to-Stress Attraction (TSA) principle, stated as follows:

(54)   Tone-to-Stress Attraction

```
W       S
|       |
T       T
‡_____‡
|      /|
c       c       (T=tonal root node, c=contour, ‡=delink)
```
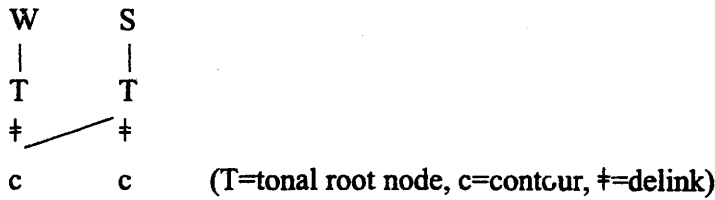
TSA has the effect of detaching the contour of the initial unstressed syllable, relinking it

to the second stressed syllable, and at the same time, supplanting the original contour of

the latter. It is modeled as the interaction of MAX(TONE)/INITIAL and contour tone

licensing constraints in our analysis.

Chen has a different interpretation of the tone register in both strong and weak

positions. He put forth the following principles in the determination of register:

(55) a. in strong position: the register is high if it is associated with a falling tone or

checked syllable, otherwise it retains its underlying specification;

b. in weak position: the register is uniformly low before a rising tone (i.e. MH or

LH), otherwise the register retains its underlying specification.

To facilitate our discussion, we repeat Rose's (1990) transcriptions for W-S

patterns below.

(56) W-S disyllabic patterns

σ_1	σ_2					
	T1:441	T2:323	T3:231	T4:213	T5:5	T6:23
A T1:441	33-441				33-4	
B T3:231	11-441				11-4	
C T4:213	11-334		11-24		11-4	
D T6:23	1-441	1-35	1-242	1-114	1-4	1-25

In our analysis, the register is always high in metrically strong position. As seen

from his classification of sandhi forms (Chen 2000:68), Chen treats 334 and 24 (pattern C)

as high and low rising tones respectively. We have examined the f0 tracks corresponding to these two tones in section 4.4.2.3, and concluded that the difference lies in the onset voicing of the second syllable. When the second syllable has a voiced onset (in the 11-24 case), the f0 is continuous at the syllable boundary. When it has a voiceless onset (in the 11-334 case), the f0 is broken at the syllable boundary. Further, a voiceless onset tends to raise the f0 a little bit. The two tones (334 and 24) have almost identical f0 peaks, as seen from the relevant averaged f0 tracks in section 4.4.2.3, chapter 4. We also showed that the tones after the checked tone (tone 6) are all realized in high register after examining the corresponding f0 tracks. In a metrically weak position, the register retains its underlying specification because of initial-position prominence.

Chen maintains that TSA also occurs with a checked syllable in the second position (patterns A, B, C), but a rising or falling contour is simplified to a high level tone. Our analysis is different from his in that the checked syllable is not able to carry a contour tone because of its short duration. Consequently, the contour licensing constraints restrict a high level tone to the checked syllables.

Chen's analysis is not explicit about cases in which the checked syllable in the first position (pattern D) does not shift its underlying tone to the long syllable in the second position. This is expected in our analysis since a contour tone is preferred on a long syllable given the ranking of constraints we have established.

5.3.4.2   S-W Patterns

We first repeat Rose's transcriptions below.

(57) S-W disyllabic patterns

217

σ₁	σ₂					
	T1:441	T2:323	T3:231	T4:213	T5:5	T6:23
A T1:441	44-31				44-32	
B T2:323	334-51				334-5	
C T4:213	114-51				114-5	
D T5:5	5-51				5-5	

Another difference between our analysis and Chen's is that he regards 31 and 51 as real low and high falling tones while in our analysis they are f0 transitions from the preceding high pitch target to the low pitch target. The difference between 31 and 51 is due to the f0 peak delay in the preceding rising tones (334 and 114 in B and C): when the f0 peak delays into the beginning of the following syllable, the falling f0 transition starts at a higher f0 value. There is no f0 peak delay when the preceding tone is H (pattern A), resulting in a falling f0 transition starting at a lower f0 value. We further attributed the low pitch target on the second syllable to the low boundary tone.

Chen's analysis can be summarized in the following autosegmental representations.

(58) a. T2-T: rightward migration of the tonal contour

```
   S    W        S    W
  /\   /\   →   /\   /\
 μ μ  μ μ      μ μ  μ μ
 | | \ /       |  \ /  |
 M H  L        M  H   L
```
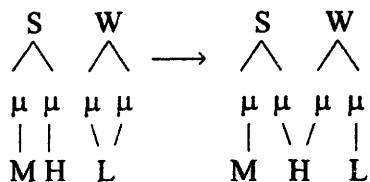
b. T1-T: melodic interpretation + rightward migration of the tone contour

```
   S    W        S    W        S    W
  /\   /\   →   /\   /\   →   /\   /\
 μ μ  μ μ      μ μ  μ μ      μ μ  μ μ
 | | \ /       | | \ /       \ /  | |
 H L  L        H M  L        H   M L
```
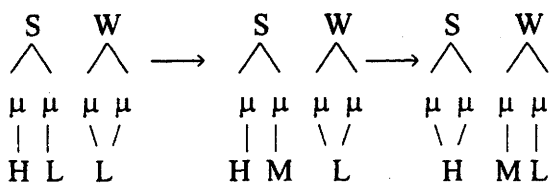
In Chen's view, what is going on is a rightward migration of the tone contour from the initial stressed syllable to the weak syllable which has lost its underlying tone. In (a), the weak syllable acquires a L default tone. Then the H of the stressed syllable slides into the second syllable, creating a high falling tone. In (b), there is one extra step, which he calls "melodic interpretation", whereby a L between a H and a L is raised halfway to a mid-level pitch. The resulting M then migrates rightwards, and shifts altogether into the next syllable.

Chen's analysis captures the phonetic facts of S-W tone sandhi, but an immediate question is why M shifts altogether into the next syllable in (b), but H only partly shifts into the next syllable in (a). In addition, by positing a low boundary tone, we are able to propose a much simpler analysis, which is independently justified in other Chinese dialects (for example, Mandarin Chinese).

A more general critique of Chen's approach is that it does not explain why the initial tone is attracted to the stressed syllable (via his TSA). Rather it is stipulated that the initial tone moves to the stressed syllable, supplanting the underlying tone of the stressed syllable. In our theory, tone movement emerges from interaction of positional faithfulness and positional markedness constraints, both of which make reference to two kinds of prominent positions.

5.4 Theoretical Implications

5.4.1 Tonal Representation

One of the major theoretical issues in Chinese tonology is concerned with the geometrical representation of tones, particularly contour tones, in the phonology. Reviews of different proposals can be found in Bao (1990, 1999), Duanmu (1990), Yip (1995) and Chen (1996, 2000). In a nutshell, they differ as to the status of a contour tone unit and the relation between register and terminal tonal features. As a refinement to Yip's (1989) model of tonal representation, Bao (1990, 1999) proposes a tonal geometry in which the register (r) and contour (c) nodes, which stand in a sisterhood relation, are dominated by the tonal root node (t), depicted below.

(59) Tonal representation (Bao 1990, 1999)

$$
\begin{array}{c}
t \\
\diagdown \diagup \\
r \quad c
\end{array}
$$
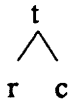
Crucial evidence in support of Bao's model has been sought from contour movement/spreading. As Chen (2000) points out, Zhenhai tone sandhi provides a convincing case for contour movement as a unit and also the separation of register and contour. As shown in the following derivation, the initial contour (c_1) moves to the second syllable, leaving behind its own register specification (r_1). Consequently, the underlying contour (c_2) of the second syllable is replaced.

(60)

$$
\begin{array}{ccccc}
\sigma_1 \quad \sigma_2 & \rightarrow & \sigma_1 \quad \sigma_2 & \rightarrow & \sigma_1 \quad \sigma_2 \\
r_1\, c_1 \quad r_2\, c_2 & & r_1\, c_1 \; r_2\, c_2 & & r_1\, c \quad r\, c_1 \\
& & & & L\text{-} \quad h
\end{array}
$$

220
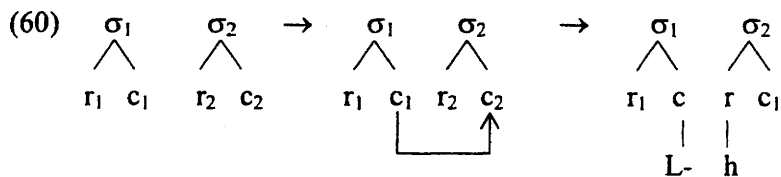
Tone movement in Zhenhai would be inexplicable without positing a contour node independent of a register node in the tonal representation. At the same time, Bao's model is also able to capture the composite nature of contour tones because terminal tonal features like H and L are dominated by the contour node. In other words, a rising tone still consists of L and H at the terminal level.

5.4.2   Positional Prominence, Positional Faithfulness and Positional Markedness

Beckman's (1998) theory of positional faithfulness is keyed to the idea of positional prominence in phonology.   She suggests that there are three distinct patterns of phonological asymmetry which can be used as a diagnostic of positional prominence or privilege.   For example, prominent positions license phonological contrasts which are neutralized elsewhere; they also resist phonological processes which apply elsewhere. According to her, prominent positions include stressed syllables, root-initial syllables, long vowels, syllable onsets and possibly final syllables.   It deserves mentioning that in Beckman's model a phonological position is rendered prominent only by the surface properties associated with that position, for example, maintenance of phonological contrasts.

Tone movement in Zhenhai is remarkable in that the surface properties which render the initial syllable prominent (i.e. longer syllable rime duration and initial tone preservation) do not license the realization of the preserved tone.   Rather the preserved tone is realized in a position which is phonologically stressed.   Contrasting tonal contours, originating, in the initial position, are maintained on the stressed syllable.   Stressed syllables fail to retain their underlying tones.   The altruistic nature of the initial position

in Zhenhai calls for a re-appraisal of positional faithfulness. In Zhenhai W-S disyllabic tone sandhi, the most unfaithful mapping takes place: both syllables acquire tonal output different from the input. The initial tone moves to the second syllable, supplanting its underlying tone.

Our approach heavily depends on the interaction of positional faithfulness and positional markedness constraints, arguing for the admission of both into the phonology. More specifically, positional faithfulness retains an input tone in one prominent position (domain-initial in Zhenhai); positional markedness requires a stressed syllable to take on certain tonal shape on the surface.

5.4.3   Phonetically-driven Contour Tone Distribution

Zhang (2001) puts forth a careful proposal of contour tone distribution in which phonetic duration (and sonority) of a syllable rime plays a vital role. His theory predicts that a longer syllable is more likely to license a contour tone than a shorter syllable. By longer and shorter, he refers to the phonetic duration (which covers more diverse categories than phonological duration). Specifically, four properties of a syllable make it a better licensor: (a) having a long vowel, (b) being stressed, (c) being in the final position of a prosodic domain, and (d) belonging to a short word. He argues that these properties can be unified in the sense that they enjoy certain durational advantage over other positions. He further argues that phonetic duration (and sonority of a syllable rime) is the most relevant phonetic parameter in contour tone distribution.

The contour tone distribution pattern observed in Zhenhai poses a serious challenge to Zhang's phonetically-grounded approach to contour tone distribution. As

we mentioned earlier, when a phonological short syllable (i.e. checked syllable) occurs in the initial stressed position, the following phonological long syllable (i.e. smooth syllable) still loses its underlying tone and takes on a low boundary tone. In addition, in both S-W and W-S patterns involving only smooth syllables, the first syllable is always phonetically longer than the second, regardless of where stress is and where contrasting tonal contours reside. This pattern also presents a problem to Zhang's approach. In trisyllablic words, the initial syllable lengthening is not as obvious as in disyllable word, but still the syllable which carries contrasting tonal contours does not enjoy clear durational advantage over other syllables.

5.5   Concluding Remarks

We have laid out a formal theory of tone movement in Optimality Theory, using the lexical tone sandhi in Zhenhai. In our analysis of tone movement, different kinds of prominent positions are singled out by the phonology and they are referred to by the input-oriented positional faithfulness and output-oriented positional markedness constraints. The positional faithfulness constraints, when properly ranked with other constraints, retain the underlying tones in the relevant prominent positions. The positional markedness constraints, taking the form of contour licensing constraints, restrict the retained tonal contour to a specific prosodic position. Tone movement cannot be easily accommodated by the single-prominence theory in which stress is the only relevant parameter.

223

In the next chapter, we discuss in detail how the retained tones in the prominent prosodic positions are distributed in tone sandhi domains by drawing on evidence from a number of Chinese dialects.

Chapter 6 Contour Re-association and Spreading: Effects of FO Timing and Prosodic Positions

6.1 Introduction

We have discussed at length the separation of tone retention and tone realization in tone sandhi phenomena in Chinese dialects. Our main contention is that a lexical tone can be retained in one position, but realized in another. We have proposed that tone retention can be dealt with by general and positional faithfulness constraints while tone realization can be regulated by general and positional markedness constraints. Tone movement, as we have observed in Zhenhai and Wenzhou, furnishes an example in which a tone is preserved at one edge within a prosodic tone sandhi domain, but moves to another carrying metrical prominence. In the OT analysis that we have proposed for the tone movement in Zhenhai in chapter 5, the initial tone is preserved due to the high-ranking positional faithfulness constraint, MAX(TONE)/INITIAL, but where it is realized is determined by the contour licensing constraints which restrict certain contour shapes to relevant prosodic positions (initial, final or stressed syllable).

Although the tone movement in Zhenhai is remarkable in Chinese dialects, how the preserved tone is realized is quite straightforward. In most cases, the initial tonal contour either stays put or moves to another position, depending on where the stress is. There is no obvious contour re-association as one observes in Shanghai. In other words, the preserved rising tone is realized as a rise on its eventual host, and the preserved

falling tone is realized as a fall unless followed by a low boundary tone. We attributed the simplification of the falling tone to a high before a low boundary tone to a general tonal process, tone absorption, attested in both West African languages (Hyman and Schuh 1974) and Chinese dialects (Tianjin dialect, reported by Li and Liu 1985).

In this chapter, we examine a variety of patterns in contour re-association and spreading in Chinese dialects, most notably Wu dialects, and show how they can be derived in a constraint-based approach adopting both positional faithfulness and positional markedness constraints. We will focus our attention on the effects of f0 timing and prosodic positions on how the preserved tone is realized on its eventual host syllable, in the sense to be made clear in the following sections. In the meanwhile, it will be shown that the notion of edge-in association in tone mapping is derivable by invoking multiple prominent positions in tone mapping.

6.2　Phonetic Timing of F0 Target: Local Migration

In chapter 2, we studied the phonetic timing of f0 target in Mandarin rising tone. Our results, combined with Xu's (2001) findings, showed different degrees of f0 peak delay as a function of the following context. For example, when the syllable in rising tone is followed by a toneless syllable, its f0 peak is realized near the beginning of the following syllable rime. When it is followed by two or three toneless syllables, its f0 peak is found in the middle of the following syllable rime. Although the f0 peak of the rise moves rightward, the beginning of the rise is consistently aligned with the middle of its host syllable. In addition, the f0 peak is not found to encroach on the third syllable, i.e. it is

always delayed into the immediately following syllable and maintains a local relation with its original host syllable.

The above phonetic process of pitch target implementation finds its phonological parallel in contour re-association. We start with an illustration of the contour re-association process that we have been discussing in Shanghai.

(1) σ_1 σ_2 σ_3 σ_4 $\rightarrow$ σ_1 σ_2 σ_3 σ_4 $\rightarrow$ σ_1 σ_2 σ_3 σ_4

$\quad\quad$ $\wedge$ $\quad$ $\wedge$ $\quad$ $\wedge$ $\quad$ $\wedge$ $\quad\quad$ $\wedge$ $\quad\quad\quad\quad$ $|$ $\quad$ $|$ $\quad\quad$ $|$

$\quad$ $T_1 T_2$ $\;$ $T_3 T_4$ $\;$ $T_5 T_6$ $\;$ $T_7 T_8$ $\quad$ $T_1 T_2$ $\quad\quad\quad$ T_1 $\quad$ T_2 $\quad$ L-
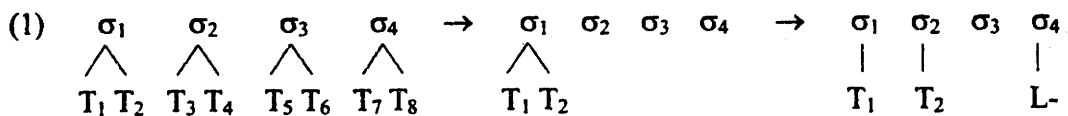
In (1) the first syllable is phonologically associated with a contour tone consisting of T_1 and T_2 (e.g. LH); the second syllable with a contour tone consisting of T_3 and T_4, and so on. After the phonologically associated tones of the non-initial syllables (σ_2, σ_3, σ_4) are deleted, T_1 and T_2 are mapped to the segmental strings from left to right in one-to-one fashion. When there are more than two syllables in the domain (i.e. compounds), the domain-final syllable gets a boundary L- tone, following our analysis of Mandarin neutral tone in chapter 3. Other syllables outside the first two syllable window are simply toneless and their pitch is transitional.

The f0 peak delay in Mandarin is a phonetic process because it does not incur any phonological alternation or neutralization. The rising tone is still phonologically associated with its host syllable despite the fact that its f0 peak migrates onto the following toneless syllable. By contrast, contour re-association can be considered as a case of f0 target delay[1], but phonologized as a process of phonologically aligning the pitch targets with the following syllable. Another similarity between Mandarin f0 peak

---

[1] F0 targets refer to f0 peak in the case of a rise, and f0 valley in the case of a fall.

delay and Shanghai contour re-association is that only the immediately following syllable is involved, illustrated below.

(2) a. σ_1 σ_2 σ_3 σ_4 $\rightarrow$ σ_1 σ_2 σ_3 σ_4

      $\bigwedge$ $\bigwedge$ $\bigwedge$ $\bigwedge$ $|$ $|$ $|$

    $T_1 T_2$ $T_3 T_4$ $T_5 T_6$ $T_7 T_8$ T_1 T_2 L-

b.*σ_1 σ_2 σ_3 σ_4 $\rightarrow$ σ_1 σ_2 σ_3 σ_4

      $\bigwedge$ $\bigwedge$ $\bigwedge$ $\bigwedge$ $|$ $|$ $|$

    $T_1 T_2$ $T_3 T_4$ $T_5 T_6$ $T_7 T_8$ T_1 T_2 L-

c.*σ_1 σ_2 σ_3 σ_4 $\rightarrow$ σ_1 σ_2 σ_3 σ_4

      $\bigwedge$ $\bigwedge$ $\bigwedge$ $\bigwedge$ $|\diagup$ $|$ $|$

    $T_1 T_2$ $T_3 T_4$ $T_5 T_6$ $T_7 T_8$ T_1 T_2 L-
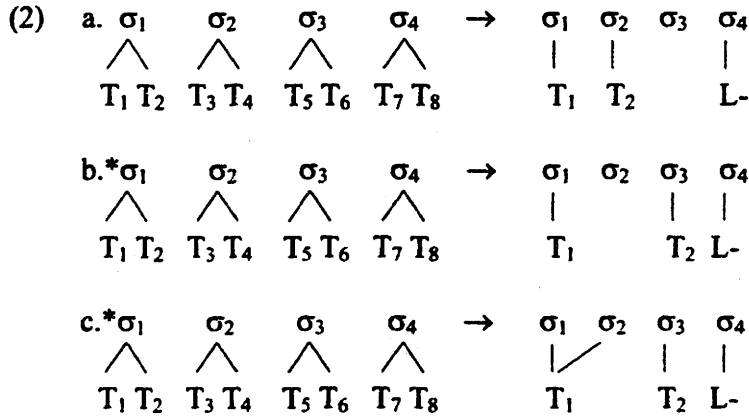
In (a) T_1 and T_2 are re-mapped from left to right in one to one fashion. In (b), T_2 skips the second syllable and seeks out the third syllable. (c) is identical to (b) except that T_1 spreads to the second syllable. A low boundary tone is inserted on the final toneless syllable. In all three candidates, the initial tones are preserved in the output, respecting positional faithfulness constraints, to be substantiated later. We propose that the constraint preventing the non-local migration of the f0 peak in Mandarin is also responsible for ruling out (b) and (c) above in Shanghai, given below.

(3)   CONTIGUITY(SYL-TONE): No syllable intervenes between the syllable that bears the tone in the input and the syllable that realizes the tone in the output.

The constraint excludes skipping in pitch target implementation and contour re-association. It is conceivable that long distance displacement of a pitch target potentially obscures its affiliation with the original host syllable. Its effect can only be overridden when a phonological constraint requires the pitch target to be realized in a specific prosodic position. For that matter, we consider the tone mapping in Tangxi in section 4.

6.3   Effect of Prosodic Positions

We have discussed the relevance of multiple prominent positions and their interaction in tone retention and tone realization. In this section, we provide an overview of the diverse patterns of contour re-association and their correlation to prosodic positions.

In our previous discussion, we have seen examples of positional effects in contour tone distribution. For example, the complex contour tone (tone 3, transcribed as 214) in Mandarin Chinese is only found in phrase-final position or citation form. As we will argue later in this chapter, the tone melody mapping in Danyang presents a case in which the rising tone is restricted to domain-final position. The advantage of the final syllable for contour bearing has already been observed by Clark (1983), who attributes this effect to final lengthening. Zhang (2001) further establishes the correlation between contour tone distribution and final syllables. These facts about contour tone distribution can be captured by making reference to prosodic positions.

When a contour tone is retained in the initial position (either due to initial position prominence or metrical prominence), its realization there displays an array of patterns. We have observed that in Shanghai, both initial rise and fall spread over the first two syllables, resulting in one level tone on each syllable (L-H or H-L).[2] This pattern can be attributed to a positional markedness constraint, which has the effect of ruling out initial syllables in contour tones, defined below.

(4)   a. *CONTOUR/INITIAL: No contour tone is allowed on a domain-initial syllable.

---

[2]   Hyphen indicates syllable boundary.

However, there is evidence that this constraint should be relativized to rise and fall respectively, *RISE/INITIAL and *FALL/INITIAL. For example, in Nantong and Suzhou the initial fall stays intact; only the initial rise and complex contour tone undergo contour re-association. We briefly review the relevant facts in Nantong, originally reported and analyzed in Ao (1993) and Chen (2000). Nantong is another northern Wu dialect, with an inventory of five tones: H, L, MH, LM, and HM. Within a foot, all non-initial syllables lose their underlying tones. While both rising tones (MH and LM) re-associate and then spread to the remaining syllables within a foot, the falling tone stays intact and the following toneless syllables get their pitch either by default or boundary tone insertion. For an underlying H tone, spreading is optional. The following examples are taken from Chen (2000:342ff).

(5) a. [rao le] ta   'pardon him' (pardon + ASP + him)
        (MH.o.o)   (o indicates toneless)
        (M.H.H)   Re-association and spread

     b. xo tə   'gave' (give + ASP)
        (H.o)
        (H. H)   Spread, or alternatively
        (H. L)   Default insertion

     c. tu er   'rabit'
        (HM.o)
        (HM. L)   Default insertion

A more complete analysis of Nantong tone sandhi can be found in Ao (1993), Yip (1999) and Chen (2000).

Although the contour tones in Mandarin Chinese do not undergo contour re-association, the complex contour tone does spread over onto the following toneless

syllable, to be discussed in the next section. Therefore, we observe the following markedness scale based on contour re-association in different dialects:

(6)   Complex contour > Rise > Fall

6.4   Contour Re-association and Boundary Tone Insertion

6.4.1   Tolerance of Contour Tones

Mandarin Chinese and Shanghai represent two contrasting cases with respect to contour re-association. As pointed out by Duanmu (1999:1), a well-known problem in Chinese phonology has been that in some dialects most regular syllables keep their underlying tones, but in others the initial syllable determines the tonal pattern of a multisyllabic domain. The contrast can be illustrated with Mandarin Chinese and Shanghai with the following corresponding examples:

(7)   Tonal stability in Mandarin (transcribed in pinyin, from Duanmu 1999:2)

    a. san55 + bei55 → 55-55   'three cups'
    b. san55 + pan35 → 55-35   'three plates'
    c. si51 + bei55 → 51-55   'four cups'
    d. si51 + pan35 → 51-35   'four plates

(8)   Contour re-association in Shanghai (transcribed in IPA, from Duanmu 1999:2)

    a. se52 + pe52 → 55-21   'three cups'
    b. se52 + bø23 → 55-21   'three plates'
    c. sz34^3 + pe52 → 33-44   'four cups'
    d. sz34 + bø23 → 33-44   'four plates'

---

[3]   According to Duanmu, [z] in [sz] 'four' is a syllabic consonant.

In Mandarin, the output tone pattern is simply a concatenation of the input syllable tones unless the tonal sequence is subject to the third tone sandhi (chapter 2) or there are unstressed functional elements involved (chapter 3). Specifically, both rise (35) and fall (51) are realized as such in multisyllabic spans: there is no contour re-association. In Shanghai, as we have shown in chapter 4, the initial stressed syllable retains its underlying tone and determines the overall output tone pattern. It is maintained by most researchers on Shanghai that underlying tones of the non-initial syllables are deleted, and then the retained tone of the initial syllable re-associates with the first two syllables (Zee and Maddieson 1979, Yip 1980, Selkirk and Shen 1995, Duanmu 1993, 1995, among others). This process can be illustrated as follows.

(9)   se   pe   →   se   pe   →   se   pe
      ∧   ∧   ∧ | |
      H L   H L   H L   H   L
                      (deletion)   (re-association)
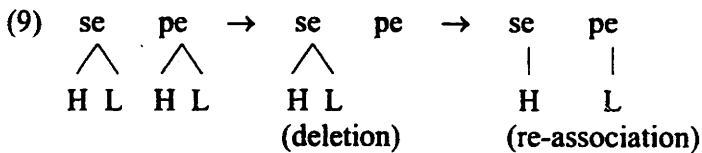
Such a process of contour re-association is not seen in Mandarin Chinese even when the contour tones (rise and fall) are followed by toneless syllables (i.e. neutral tone syllables, discussed in chapter 3). Consider the following examples.

(10) Lack of contour re-association in Mandarin

    a. fei55 + le → 55 + 2   (L)   'fly + ASP'
    a'   *55 + 5   *via spreading*
    b. lai35 + le → 35 + 3 (L)   'come + ASP'
    b'   *35 + 5   *via spreading*
    b"   *33 + 5   *via re-association*
    c. lei51 + le → 53 + 1 (L)   'tire + ASP'
    c'   *53 + 3   *via spreading*
    c"   *55 + 1   *via re-association*

The aspectual marker *le* is said in neutral tone. Tone 4 is realized as 53 in non-final position (c). As we discussed in chapter 3, neutral tone syllables following tone 1 (a), tone 2 (b) and tone 4 (c) acquire a low boundary tone. They do not get their pitch via spreading from or re-association of the preceding tones (a', b', b", c', c").

The following surface-true generalizations emerge from the forgoing discussion. Contour tones are not tolerated in the initial position of a multisyllablic domain in Shanghai: they spread over the first two syllables. Contour tones are allowed everywhere (initial, medial or final) in Mandarin Chinese: they stay with their original host syllables. The question of interest is why there is such a difference. Specifically, why do tone deletion and contour re-association take place in Shanghai, but not in Mandarin?[4] Before we provide our answer to this question, we consider a related question first.

Recall that in Mandarin Chinese tone 3 is realized as a complex contour tone (transcribed as 214) in citation and phrase-final position. In non-final position, it is realized as a low tone (transcribed as 21) when it is not followed by another tone 3. The other three tones do not exhibit this positional asymmetry in terms of tonal distribution. They (i.e. 55, 35 and 51) occur in both phrase-final and non-final positions. In chapter 2, adopting Zhang's (2001) insight on contour licensing, we attributed the restriction of 214 to the phrase-final position to the following ranking of two contour licensing constraints, *COMPLEXCONTOUR/NON-FINAL σ >> *COMPLEXCONTOUR/FINAL σ, which penalize a complex contour tone occurring on a non-final or final stressed syllable. In chapter 3, we noted that the toneless syllable following tone 3 does not acquire a low boundary tone as

---

[4] The same question was asked by Duanmu (1999), but his analysis is different from ours.

it does when it is following the other three tones (4). Instead, it is realized as a high tone target.

(11) Re-association of complex contour tone in Mandarin

mai214 + le → 21 + 4 (H)   'buy + ASP'
              *21 + 1 (L)   *via boundary low insertion*

What happens here is that the complex contour tone slides into the following toneless syllable. The re-association of the complex contour tone is only possible when the following syllable is toneless so that the high f0 ending in 214 can be realized.

We have observed contour re-association in both Mandarin and Shanghai. The difference lies in the fact Mandarin is more tolerant in contour tone distribution than Shanghai. It bans only complex contour tones from occurring in non-phrase-final position but not contour tones (rise and fall). Shanghai has a more stringent condition for contour tones. It disallows contour tones in the initial position of a multisyllabic domain.

There are dialects in which complex contour tones are allowed in non-phrase-final position. We briefly mention two of them, both spoken in Shangdong province. Jinan has an inventory of four citation tones, transcribed as 213, 42, 55, and 21. The distribution of the four tones in disyllabic tonal combinations is given below (from Z. Qian 1995). The four tones are labeled as their corresponding Middle Chinese tonal categories.

(12) Tonal distribution in disyllables in Jinan

σ_1	σ_2			
	Ia: 213	Ib: 42	II: 55	III: 21
Ia: 213	23-213	213-42	213-55	23-21
Ib: 42	42-213	42-42	42-55	55/42-21
II: 55	55-213	42-42	42-55	55-21
III: 21	21-213	21-42	21-55	23/21-21

The complex contour tone 213 is shaded in the table. It occurs in both initial and final positions. Similar distribution of the complex contour tone is seen in Yantai (Z. Qian et al. 1982). Yantai has three citation tones, transcribed as 31, 214 and 55. The distribution of the complex contour tone 214 is given in the following table.

(13) Tonal distribution in disyllables in Yantai

	I: 31	II: 214	III: 55
I: 31	35-31	31-214	31-55
II: 214	35-31	55-214	214-55
III: 55	55-31	55-214	31-55

The contour distribution patterns in the four dialects are summarized below.

(14) Contour tone distribution in Jinan, Yantai, Mandarin and Shanghai

distribution	complex contour		contour	
	initial	final	initial	final
Jinan/Yantai	yes	yes	yes	yes
Mandarin	no	yes	yes	yes
Shanghai			no	yes

The distribution of complex contour tones is unrestricted in Jinan and Yantai unless they are modified by the tone sandhi processes. Mandarin only allows the complex contour tone (214) in phrase-final position. Shanghai does not have complex contour tone. It disallows contour tones in the domain-initial position.

In sum, we have provided evidence that both Mandarin Chinese and Shanghai exhibit the process of contour tone re-association. In the former, the complex contour tone (dipping tone) spreads onto the following toneless syllable and the contour tones

(rise and fall) stay put whereas in the latter, the initial contour tones spread over the first two syllables in a multisyllabic domain.

6.4.2   A Formal Analysis of Contour Re-Association

We have shown in chapter 4 that in Shanghai an initial stressed syllable (smooth or checked alike) retains its underlying tone; non-initial syllables lose their underlying tones. This surface-true generalization can be captured by ranking MAX(TONE)/$\acute{\sigma}$ over *TONE.

(15)   a.   MAX(TONE)/$\acute{\sigma}$: If T is a tone on the stressed syllable in the input, then T has a

correspondent in the output. ('Do not delete a stressed tone.')

b.   *TONE: Any tone bearing unit with a tone in the output is banned.

c.   MAX(TONE)/$\acute{\sigma}$ >> *TONE

$\acute{\sigma}_1$ σ_2 $\wedge$ $\wedge$ $T_1 T_2$ $T_3 T_4$	MAX(TONE)/$\acute{\sigma}$	*TONE
a. $\acute{\sigma}_1$ σ_2 $\wedge$ $\wedge$ $T_1 T_2$ $T_3 T_4$		**!
b. $\acute{\sigma}_1$ σ_2 $\wedge$ $T_3 T_4$	*!	*
c. ☞ $\acute{\sigma}_1$ σ_2 $\wedge$ $T_1 T_2$		*

The ranking in (c) has the effect of deleting the tones on the unstressed syllables and keeping the stressed syllable tone. Candidate (a) is the most faithful to the input, but it

has two violations of *TONE since both syllables are associated with a tone. (b) violates the high-ranking MAX(TONE)/ớ, so it loses to (c).

Although candidate (c) in the above tableau preserves the underlying tone of the initial stressed syllable, it is still not the attested output tone pattern. In the actual output, the preserved contour tone re-associates with the first two syllables. In our constraint system, (c) is ruled out by *CONTOUR/INITIAL. The following tableau establishes the ranking for *CONTOUR/INITIAL and IDENT(TONE)/ớ. Note that the high-register rising tone has been transcribed as MH, which is actually a short hand. If both register and contour nodes are considered, MH is specified as register(high) and contour(LH). From now on, we will ignore the register distinction between high rise and low rise in Shanghai, and refer to both as LH.

(16) a. IDENT(TONE)/ớ: If a stressed syllable has a tonal specification in the input, then its correspondent has identical specification in the output. ('Do not change the tonal specification of a stressed syllable.')

b. *CONTOUR/INITIAL >> IDENT(TONE)/ớ, *TONE

ớ₁ σ₂ ∧ ∧ L₁ H₁ L₂ H₂	*CONTOUR/ INITIAL	IDENT(TONE)/ớ	*TONE
a. ớ₁ σ₂ ∧ L₁ H₁	*!		*
b. ☞ ớ₁ σ₂ | | L₁ H₁		*	**

237

*CONTOUR/INITIAL forces re-association of the preserved the underlying tone of the initial stressed syllable, which is penalized by IDENT(TONE)/ớ.

Since the low boundary tone is only inserted when the compound is longer than two syllables, it follows that BND(L)/EDGE outranks DEP(TONE) because of inserting a tone not in the input, but it is outranked by MAX(TONE)/ớ. Since *CONTOUR/UNSTRESSED σ is unviolated, any candidates with a contour tone on the unstressed syllable will be ruled out in the first place, though they may satisfy both MAX(TONE)/ớ and BND(L)/EDGE, as the candidate (c) (shaded) in the following tableau.

(17) a. BND(L)/EDGE: An unstressed syllable at the edge of a prosodic domain coincides with a low tone target.

b. DEP(TONE): If T is a tone in the output, then T has a correspondent in the input. ('Do not insert a tone.') (after McCarthy and Prince 1995)

c. *CONTOUR/UNSTRESSED σ: No contour tone is allowed on an unstressed syllable.

d. MAX(TONE)/ớ >> BND(L)/EDGE

$\overset{\acute{\sigma}_1}{\underset{L_1\,H_1}{\wedge}}\;\;\overset{\sigma_2}{\underset{L_2\,H_2}{\wedge}}$	MAX(TONE)/ớ	BND(L)/EDGE
a. $\quad\overset{\acute{\sigma}_1}{\underset{L_1}{\mid}}\;\;\overset{\sigma_2}{\underset{L\text{-}}{\mid}}$	*!	
b. ☞ $\overset{\acute{\sigma}_1}{\underset{L_1}{\mid}}\;\;\overset{\sigma_2}{\underset{H_1}{\mid}}$		*
c. $\quad\overset{\acute{\sigma}_1}{\underset{L_1}{\mid}}\;\;\overset{\sigma_2}{\underset{H_1\,L\text{-}}{\wedge}}$		

238

e. BND(L)/EDGE >> DEP(TONE)

$\acute{\sigma}_1$ σ_2 σ_3 ⋀ ⋀ ⋀ L_1 H_1 L_2 H_2 L_3 H_3	BND(L)/EDGE	DEP(TONE)
a. $\acute{\sigma}_1$ σ_2 σ_3 | | ╱ L_1 H_1	*!	
b. ☞ $\acute{\sigma}_1$ σ_2 σ_3 | | | L_1 H_1 L-		*

Another unviolated constraint in Shanghai is CONTIGUITY(SYL-TONE), defined earlier. It penalizes candidates with non-local association of the preserved contour tone, as the next tableau illustrates.

(18)   CONTIGUITY(SYL-TONE), *TONE

$\acute{\sigma}_1$ σ_2 σ_3 σ_4 ⋀ ⋀ ⋀ ⋀ L_1 H_1 L_2 H_2 L_3 H_3 L_4 H_4	CONTIGUITY (SYL-TONE)	*TONE
a. ☞ $\acute{\sigma}_1$ σ_2 σ_3 σ_4 | | | L_1 H_1 L-		***
b. $\acute{\sigma}_1$ σ_2 σ_3 σ_4 | | | L_1 H_1 L-	*!	***
c. $\acute{\sigma}_1$ σ_2 σ_3 σ_4 | | | L_1 H_1 L-	*!	***

(b) and (c) violate CONTIGUITY(SYL-TONE) because H_1 is phonologically associated with σ_1 in the input, but it is realized on σ_3, with an intervening σ_2. Putting L_1 and H_1 both

239

under σ_2 could satisfy CONTIGUITY(SYL-TONE), but it is ruled out by the unviolated constraint *CONTOUR/UNSTRESSED σ since the stress falls on σ_1.

A ranking argument for *CONTOUR/INITIAL and COINCIDE($\acute{\sigma}$, CONTOUR)[5] can be found by observing that the preserved contour tone on the initial stressed syllable always spreads over the first two syllables. As a result, *CONTOUR/INITIAL >> COINCIDE($\acute{\sigma}$, CONTOUR), illustrated as follows.

(19) Contour licensing constraints II: positive (from chapter 5)

　　 a. COINCIDE($\acute{\sigma}$, CONTOUR): If a syllable is stressed, it bears a contour tone.

　　 b. COINCIDE($\acute{\sigma}$, H): If a syllable is stressed, it bears a H tone.

　　 c. COINCIDE($\acute{\sigma}$, L): If a syllable is stressed, it bears a L tone.

(20) *CONTOUR/INITIAL >> COINCIDE($\acute{\sigma}$, CONTOUR), *TONE

$\acute{\sigma}_1$　　σ_2 ∧　　∧ $L_1 H_1 L_2 H_2$	*CONTOUR/ INITIAL	COINCIDE ($\acute{\sigma}$, CONTOUR)	*TONE
a.　$\acute{\sigma}_1$　σ_2　∧　$L_1 H_1$	*!		*
b. ☞　$\acute{\sigma}_1$　σ_2　|　　|　L_1　H_2		*	**

We have seen that the positional faithfulness constraint MAX(TONE)/$\acute{\sigma}$ has the effect of keeping the underlying tone of the stressed syllable in the output, and the

---

[5]　COINCIDE($\acute{\sigma}$, CONTOUR) subsumes both COINCIDE($\acute{\sigma}$, RISE) and COINCIDE($\acute{\sigma}$, FALL) when no distinction between a rise and a fall is necessary.

positional markedness constraint *CONTOUR/INITIAL has the effect of ruling out contour tones on the domain-initial syllable. In order to establish their ranking, we consider two special cases involving monosyllabic domains in phrase-final and phrase-initial positions. The idea is that if contour re-association is a response to the constraint banning domain-initial contour tone, then the ranking MAX(TONE)/σ́ >> *CONTOUR/INITIAL will preserve the contour tone in a monosyllabic domain at the cost of violating the latter. The opposite ranking will give rise to leveling of contour. We will show, in what follows, that preservation occurs when the monosyllabic domain is phrase-final, and leveling occurs when the monosyllabic domain is phrase-initial. We argue that *CONTOUR/INITIAL should be split into two more specific contour licensing constraints, referring to different prosodic categories: *CONTOUR/PHRASE-INITIAL, and *CONTOUR/FOOT-INITIAL. We obtain the following ranking: *CONTOUR/PHRASE-INITIAL >> MAX(TONE)/σ́>> *CONTOUR/FOOT-INITIAL.

We first consider the monosyllabic domain in phrase-final position. The following examples are taken from Chen (2000:312, originally due to Xu et al. 1988).

(21) a.
```
          x           phrase-level stress
    (x   ) (x)         word-level stress
    la   la  zā        'to stretch out' (lit. to pull + long)
    HL-HL-LM           citation tone
    (H - L)(LM)        tone deletion and tone re-association
```

b.
```
          x           phrase-level stress
    (x   ) (x)         word-level stress
    k'ue dɤ  du        'to have a good physique' (lit. physique + big)
    MH-LM-LM
    (M - H) (LM)
```

241

```
c.          x                  phrase-level stress
     (x    ) (x)               word-level stress
     dø  dø  tsø               'to turn round and round' (lit. circles + turn)
     LM-LM-MH
     (L - M) (MH)


d.          x                  phrase-level stress
     (x    ) (x   )            word-level stress
     bong dʋ se faʔ            'unkempt' (lit. disheveled head + messy hair)
     LH-LH-MH-Hʔ
     (L - H)(M - H)
```

Duanmu (1993, 1995) argues that tone deletion and tone re-association apply within a left-headed metrical foot, which is usually the size of a disyllabic lexical compound. A multisyllabic compound can form one larger foot via a process of stress reduction, which is typically restricted to lexical compounds, as discussed in section 4.4.3.3. All the examples here are phrasal constructions, consisting of more than one lexical compound. The phrase-level stress in Shanghai is assigned by the Nonhead Stress principle by which the syntactic nonhead gets mores stress than the corresponding head (Cinque 1993, Duanmu 1995). How it works is irrelevant for the purpose of our discussion. As Chen (2000) suggests, stress reduction does not apply at the phrasal level. Therefore, all the above examples constitute two domains. It is worth pointing out that there is no contour re-association within the phrase-final monosyllabic domain. Compare (c) and (d), and one shall see that contour re-association occurs when the stressed syllable is not in domain-final position, i.e. when it is followed by at least one syllable.

More examples of resistance to contour re-association in phrase-final monosyllabic domain are furnished by Selkirk and Shen (1990:321ff). Tonal transcriptions and groupings are theirs, and metrical grids are added.

(22) a.   x   phrase-level stress
       (x   ) (x   )   word-level stress
       z la? zã he   'to live in Shanghai'
       LH-LM-LH-MH   citation tone
       (L - H) (L - H)   tone deletion and tone re-association

b.   x   phrase-level stress
       (x   ) (x)   word-level stress
       tɔ ?i? pe zo   'to pour a cup of tea'
       MH-MH-HL-LH
       (M - H - L)(LH)

The following tableau illustrates how the contour tone in a phrase-final monosyllabic domain is realized.

(23) a. Max(Tone)/ớ >> *Contour/Foot-Initial

la la zã HL-HL-LM	Max(Tone)/ớ	*Contour/ Foot-Initial
a. (H-L)(L)	*! (L)	
b. ☞ (H-L)(LM)		* (LM)

When the monosyllabic domain is phrase-initial, contour simplification occurs, as the following examples demonstrate. They are taken from Chen (2000:313-314, originally due to Tang et al. 1988).

(24) a.   x   phrase-level stress (Nonhead stress on object)
       (x) (x   )   word-level stress
       t'ø çiɔ li   'penny-wise' (lit. seek + small gains)
       HL-MH-LH   citation tone
       (H)(M - H)   contour simplification

b.   x   phrase-level stress
       (x)(x)   word-level stress
       dzi i   'to quit smoking'
       LH-HL
       (M)(HL)

```
c.    x                    phrase-level stress
   (x) (x)                 word-level stress
   ts'ɔ ve                 'to fry rice' (phrase)
   MH-LH
   (H)(LH)


c'. (x    )                compound-level stress (stress clash resolved)
   (x) (x)                 word-level stress (stress clash)
   ts'ɔ ve                 'fried rice' (compound)
   MH-LH
   (M - H)
```

(a), (b) and (c) are all phrasal constructions. (c') is a compound, forming a couplet with (c). Recall that stress clash is obligatorily resolved at the level of compound, as shown in (c'). But at the phrasal level, as Duanmu (1993, 1995) argued, if resolution of stress class would create an ill-formed foot form (the foot form is trochaic in Shanghai) or leave a syllable unparsed, then the stress clash will be tolerated, as in (a), (b) and (c). Remarkably, while the phrase-final monosyllabic foot preserves the whole contour, its phrase-initial counterpart does undergo contour simplification, by which only a level tone surfaces. The positional asymmetry in contour licensing can be accounted for by the following ranking, which penalizes phrase-initial contour.

(25)   a.   *CONTOUR/PHRASE-INITIAL >> MAX(TONE)/ố

ts'ɔ ve MH-LH	*CONTOUR/ PHRASE-INITIAL	MAX(TONE)/ố
a. (MH)(LH)	*! (MH)	
b. ☞ (H)(LH)		* (MH)

One minor detail deserves some comment. Although the initial contour simplifies to a simple level tone, it is not entirely clear whether the actual output is part of the corresponding contour tone. Chen (2000:315) suggests the following rule of thumb: H

244

for high-registered tones; M or L for low-registered tones depending on whether they fall on smooth or checked syllables. We suspect that contour simplification in Shanghai occurs in the initial position of an intonational phrase, rather than a phonological phrase. See Selkirk and Shen (1990) for a discussion of the prosodic domains in Shanghai. If that is the case, the result of contour simplification might be subject to intonational effects as well. Without further available details, we will leave it as an open question.

Our analysis makes correct predictions on the contour distribution of phrase-medial monosyllabic domains. Since *CONTOUR/PHRASE-INITIAL only triggers contour leveling in phrase-initial monosyllabic domain, it does not affect phrase-medial monosyllabic domains. When MAX(TONE)/σ is ranked higher than *CONTOUR/FOOT-INITIAL, as in our analysis, no contour leveling is predicted to happen in a phrase-medial monosyllabic domain. The prediction is borne out in the following example (Duanmu 1995:255), with a minor clarification.

(26) (lo pe) (ŋo)(kɤ)
    boss   bit dog   'the boss bit the dog'
    LH-LH-LH-LH
    (L - H)(LH)(LH)

According to Selkirk and Shen (1990), prosodic word and phonological phrase in Shanghai are projected from the left edge of a lexical X^0 and XP respectively. The above example can be parsed as follows.

(27) Phonological Phrase (left edge of XP):   (((
      Prosodic Word (left edge of X^0):   (((
      Syntactic structure: [lo pe]NP [[ŋo]V[kɤ]NP]VP

This phrasing shows that *CONTOUR/PHRASE-INITIAL actually refers to an intonational phrase rather than a phonological phrase because LH on the verb *ŋo* does not undergo

simplification although it is initial in the phonological phrase. Once this clarification has been made, everything else follows from our analysis.

Up to now, we have covered almost all major aspects of Shanghai contour re-association. The only remaining issue is tone movement in the multisyllabic compounds with initial tone 5. The basic facts are repeated below.

(28) Tone 5 as the initial tone

initial tone	σσ	σσσ	σσσσ	σσσσσ
tone 5: 23	22-23	22-22-23	22-55-33-31 / 22-22-22-23	22-55-33-33-31

Tone 5 is a low rising checked tone in citation form. When it occurs in the initial position in a compound no longer than three syllables, the low rising tone moves as a unit to the final syllable, with its original host syllable filled with a low pitch target. Tone 5 exhibits a different pattern from the other four tones in Shanghai in that it resolves initial contour restriction by moving it to the final syllable, instead of spreading it over the first two syllables. Tone 4, another checked tone in Shanghai, patterns with the other three long tones, and shows contour re-association.

Instead of unifying tone movement and contour re-association in Shanghai, we propose that tone 5 constitutes a subsystem in which the domain-initial contour is resolved by moving it to the final position in the compound. In the constraint system we have been developing, both metrically strong position and peripheral position in a prosodic domain are prominent. What goes on with tone 5 in Shanghai is that when the metrically strong position (domain-initial in Shanghai) is not able to accommodate the

preserved contour (due to *CONTOUR/INITIAL), the contour shifts to another prominent position (domain-final). The process can be illustrated with the follow tableau.

(29) a. COINCIDE(FINAL, CONTOUR): if a syllable is in domain-final position, it bears a contour tone.

   b. MAX(TONE)/σ́, *CONTOUR/INITIAL, COINCIDE(FINAL, CONTOUR)

σ́₁ σ₂ σ₃ ⋀ ⋀ ⋀ L₁ M₁ M₂ H₂ H₃ L₃	MAX(TONE)/σ́	*CONTOUR/INITIAL	COINCIDE (FINAL, CONTOUR)
a. σ́₁ σ₂ σ₃ ⋀ | L₁ M₁ L-		*!	*
b. σ́₁ σ₂ σ₃ ⋀ ⋀ L₁ M₁ H₃ L₃		*!	
c. σ́₁ σ₂ σ₃ | | | L₁ M₁ L-			*!
d. ☞ σ́₁ σ₂ σ₃ | ⋀ L- L₁ M₁			

We only consider candidates which preserve the initial contour tone. Candidates (a) and (b) are the most faithful candidates to the head syllable in the input, but they violate *CONTOUR/INITIAL. Here 'Initial' is interpreted as foot-initial. Candidate (c) undergoes the normal contour re-association, but it is ruled out by another contour licensing constraint COINCIDE(FINAL, CONTOUR). Candidate (d) does not violate any of these three constraints. In addition, in this subsystem, *CONTOUR/UNSTRESSED σ is violated by the output, so it has to be demoted after the three constraints in the tableau. By the same

247

token, CONTIGUITY(SYL-TONE), which penalizes long distance movement of tonal targets, is also violated by the output.

In sum, the predominant tone mapping patterns (i.e. except initial tone 5) can be derived by the rankings we have established so far.

(30) Ranking summary

CONTIGUITY(SYL-TONE)   *CONTOUR/PHRASE-INITIAL *CONTOUR/UNSTRESSED σ
                                          |
                                 MAX(TONE)/σ́
                              ⟋ |
          BND(L)/EDGE *CONTOUR/FOOT-INITIAL
               | | ⟍   ⟍
          DEP(TONE)   IDENT(TONE)/σ́ *TONE COINCIDE(σ́, CONTOUR)

6.4.3   A Comparison with Mandarin Chinese

In Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1995), cross-linguistic variations are modeled as emerging from different rankings of faithfulness and markedness constraints. We propose that the difference in tolerance of contour tones between Mandarin Chinese and Shanghai stems from different rankings of the relevant constraints. We elaborate on this idea in this subsection. As we have discussed in chapter 2, the complex contour tone (214) in Mandarin is banned on non-final stressed syllables because of the constraint *COMPLEXCONTOUR/NON-FINAL σ́, defined below.

(31) a. *COMPLEXCONTOUR/NON-FINAL σ́: A complex contour tone is not allowed on a

non-final stressed syllable.

When a stressed syllable in tone 3 is followed by a toneless syllable, the complex contour re-associates with these two syllables, of which the following example is illustrative.

(32) mai214 + le0 → 21 + 4 (H)   'buy + ASP'

The point worth making is that the low boundary tone insertion, which happens to the neutral tone syllables following level or contour tones, is blocked when there is only one neutral tone syllable following tone 3. The low boundary tone will be inserted to the last neutral tone syllable when there is more than one. The scenario is reminiscent of Shanghai in which the low boundary tone insertion is blocked in a disyllabic domain because the second syllable takes the re-associated tonal target of the first syllable. The difference is the tonal entity involved: complex contour in Mandarin, but simple contour (i.e. rise and fall) in Shanghai. This shows that MAX(TONE)/σ is active in both dialects, but the re-association is triggered by different active contour licensing constraints. In Shanghai, the ranking MAX(TONE)/σ >> *CONTOUR/FOOT-INITIAL is primarily responsible for the contour re-association, assuming that *COMPLEXCONTOUR/NON-FINAL σ is undominated in Shanghai. By contrast, *COMPLEXCONTOUR/NON-FINAL σ will be shown to trigger re-association of the complex contour tone in Mandarin.

(33) *COMPLEXCONTOUR/NON-FINAL σ >> MAX(TONE)/σ, IDENT(TONE)/σ

	214-N	*COMPLEXCONTOUR/ NON-FINAL σ	MAX(TONE)/σ	IDENT(TONE)/σ
a.	214-L-	*!		
b.	21-L-		*!	*
c. ☞	21-4			*

N stands for syllables in neutral tone and L- for the inserted low boundary tone. Since 214 cannot stay on the non-final syllable, (a) is ruled out by the contour licensing constraint, though it satisfies both MAX(TONE)/σ and IDENT(TONE)/σ. (b) fails on

MAX(TONE)/ớ. (c) fails on IDENT(TONE)/ớ, which is inconsequential. The ranking

argument for *COMPLEXCONTOUR/NON-FINAL ớ and MAX(TONE)/ớ was provided in

chapter 2. When 214 is followed by a stressed syllable in any tone except tone 3, it

simplifies to 21, to satisfy *COMPLEXCONTOUR/NON-FINAL ớ, but in violation of

MAX(TONE)/ớ. Like Shanghai, the ranking MAX(TONE)/ớ >> BND(L)/EDGE ensures that

realization of the input tone of the stressed syllable is preferred over low boundary tone

insertion.

(34) MAX(TONE)/ớ >> BND(L)/EDGE

	214-N	MAX(TONE)/ớ	BND(L)/EDGE
a.	21-L-	*!	
b. ☞	21-4		*

Similarly, *COMPLEXCONTOUR/NON-FINAL ớ should also outrank BND(L)/EDGE.

(35) *COMPLEXCONTOUR/NON-FINAL ớ >> BND(L)/EDGE

	214-N	*COMPLEXCONTOUR/ NON-FINAL ớ	BND(L)/EDGE
a.	214-L-	*!	
b. ☞	21-4		*

It is crucial that in Mandarin *CONTOUR/FOOT-INITIAL should rank lower than

IDENT(TONE)/ớ so that simple contour tones can stay put on their host syllables.

(36) MAX(TONE)/ớ, IDENT(TONE)/ớ >> *CONTOUR/FOOT-INITIAL

|	35-N	MAX(TONE)/ớ	IDENT(TONE)/ớ	*CONTOUR/FOOT-INITIAL
a.	33-5		*!	
b. ☞	35-L-			*

Both candidates in the above tableau satisfy MAX(TONE)/σ, but (a) violates Ident(Tone)/σ for spreading the rise over the two syllables. (b) is the winner despite its violation of the low-ranking *CONTOUR/FOOT-INITIAL. We note in passing that BND(L)/EDGE is not responsible for the ill-formedness of candidate (a), though it is also violated by (a). When we are looking at examples with more than one neutral tone syllable, the crucial competing candidates would become 33-5-L- (a') vs. 35-N-L- (b'), corresponding to (a) and (b). Both satisfy BND(L)/EDGE, but still (b') is the winner given the ranking above.

(37) MAX(TONE)/σ, IDENT(TONE)/σ >> *CONTOUR/FOOT-INITIAL

	35-N-N	MAX(TONE)/σ	IDENT(TONE)/σ	*CONTOUR/FOOT-INITIAL
a.	33-5-L-		*!	
b. ☞	35-N-L-			*

The partial ranking of the constraints we have covered for Mandarin Chinese is summarized in (36).

(38) Ranking summary

*COMPLEXCONTOUR/NON-FINAL σ

IDENT(TONE)/σ   MAX(TONE)/σ

*CONTOUR/FOOT-INITIAL   BND(L)/EDGE
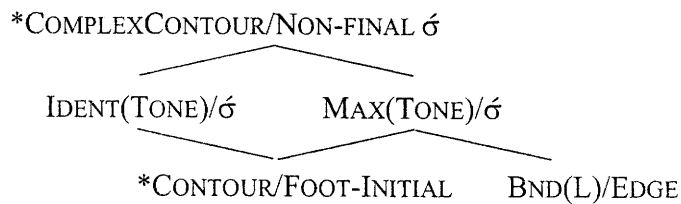
To summarize, the difference between Mandarin Chinese and Shanghai can be attributed to different rankings of the relevant positional faithfulness and contour licensing constraints, shown below. *COMPLEXCONTOUR/NON-FINAL σ is undominated in both Mandarin Chinese and Shanghai.

(39) a. Mandarin:　MAX(TONE)/ớ, IDENT(TONE)/ớ >> *CONTOUR/FOOT-INITIAL

　　　b. Shanghai:　MAX(TONE)/ớ >> *CONTOUR/FOOT-INITIAL >> IDENT(TONE)/ớ

6.4.4　Duanmu's Analyses (1993, 1999)

Duanmu (1993) and (1999) deal with the same research question we asked at the outset of this section. We review his analyses and compare them with ours.

Duanmu (1993) links the difference in contour re-association to the syllabic difference between Mandarin and Shanghai. Assuming that the moraic segment is the tone-bearing unit, and that the basic metrical structure is moraic trochee in Mandarin and Shanghai, he argues that Mandarin full syllables (in contrast with weak syllables, i.e. syllables in neutral tone) are bimoraic while all Shanghai syllables are monomoraic. Consequently each syllable in Mandarin can take two tones (L and H), hence there is no contour re-association. In Shanghai, only a stressed syllable can retain its underlying tone. In disyllables, each syllable can take only one tone, hence the initial syllable tone (LH or HL) re-associates with both syllables. When a Shanghai syllable occurs in citation form, it is lengthened to two moras so that the syllable can take both of its underlying tones.

In a dramatically different analysis, Duanmu (1999) discusses several problems with his previous approach. One of the problems is phonetic. According to Zhu (1995), in a disyllabic Shanghai word, the first syllable is much longer than the second. For smooth syllables, the average duration of the first rime is on the order of 200 ms while that of the second rime is about 100 ms. This may seem to support his earlier metrical analysis that the first syllable is stressed and the second is unstressed. However, the rime

duration of a stressed syllable in Shanghai is almost identical to that of a stressed syllable in Mandarin, as reported in Lin et al. (1984), Lin and Yan (1988), Wang and Wang (1993). Consider the following facts (Duanmu 1999:6).

(40) Rime duration in disyllables in Shanghai and Mandarin (normal read speech)

	First Rime	Second Rime
Shanghai	200 ms	100 ms
Mandarin (heavy – light)	200 ms	100 ms
Mandarin (heavy – heavy)	200 ms	200 ms

The data shows that phonetically a disyllable in Shanghai matches the heavy-light in Mandarin. As Duanmu points out, if the first syllable is considered heavy (i.e. bimoraic), one must explain why it takes just one of its underlying tones and shifts the other to the following syllable.

His present analysis assumes that the initial syllable gets stress, hence it is heavy. The metrical pattern for a disyllabic word is now heavy-light, not light-light as in his early analysis. He further assumes that there are only two syllable tones H and L, instead of HL and LH. He proposes a tonal polarity constraint which requires an initial tone to be followed by an opposite tone. Another constraint is proposed which requires a syllable to carry a simple level tone (SIMPLE TONE). This constraint forces the polarity tone to occur on the following syllable, even though the initial syllable has two moras. An example of his analysis is given below.

(41) POLARITY, SIMPLE TONE

	se pe H H	POLARITY	SIMPLE TONE
a. ☞	see pe H L		
b.	see pe H H	*!	
c.	see pe HL		*!
d.	see pe HL H		*!

In (a) H is mapped to the two moras of the first syllable, satisfying SIMPLE TONE; L is mapped to the second syllable, satisfying POLARITY. (c) and (d) satisfy POLARITY, but since the polarity tone is realized on the second mora of the first syllable, creating a contour tone, they are ruled out by SIMPLE TONE. (b) fails to observe POLARITY.

The idea of polarity tone is adopted to predict the tonal targets on the neutral tone syllables after the four tones in Mandarin. The underlying form for tone 3 is taken to be a simple low. He suggests that due to POLARITY neutral tone syllables surface as L after H (tone 1), and H after L (tone 3). After LH (tone 2) and HL (tone 4), which already satisfy POLARITY, the neutral tone syllables should take the default value L. He does not illustrate how his analysis assigns the default value to the neutral tone syllables after tone 2 and tone 4. One may ask when the default L is plugged in after tone 4, giving rise to HL-L, does it violate POLARITY for having a sequence of L-L? What constraint will rule out the unattested HL-H?

6.4.5 Non-local Migration of Tonal Targets

The contour re-association in Shanghai can be characterized as a local process of tone target due to the constraint CONTIGUITY(SYL-TONE). However, in Tangxi the ban on initial contour tone is resolved by a non-local process: the first tone stays on the first syllable and the second tone is aligned with the final syllable in the compound.

The tone sandhi pattern of this northern Wu dialect is reported by Kennedy (1953). According to him, in disyllabic expressions the first syllable carries stress in lexical constructions and the second carries stress in phrasal constructions. In compounds, the initial syllable preserves its underlying rise or fall. Then the surviving contour tones spread rightward to span over the entire domain, shown below.

(42) a. tseo sea   'disinfectant'
          LH HL   citation tone
          L   H   sandhi tone (deletion and re-association)

      b. kwen sea   'boiling water'
          HL   HL
          H   L

      c. bhao moa then   'racecourse'
          LH M   LH
          L   o   H   (o indicates toneless syllable.)

      d. pao hyie kong sz   'insurance company'
          HL HL   HL   HL
          H   o   o   L

Stress falls uniformly on the leftmost syllable in lexical compounds. The rising or falling tone of the first syllable spreads over the entire compound by re-associating the first tone segment of the contour to the initial syllable and the second tone segment to the final syllable. CONTIGUITY(SYL-TONE) is always violated in longer compounds in Tangxi.

Chen (2000) implements this tone spread by an edge-in association (Yip 1988, 1989): link the tone segments to the initial and final syllables. In our theory, the tone spread in Tangxi can be derived by invoking the positional markedness constraints, which will restrict H or L to the two prominent positions, initial stressed syllable and final syllable.

(43) a. COINCIDE(FINAL, L): If a syllable is final, it bears a low tone.

b. COINCIDE(FINAL, H): If a syllable is final, it bears a high tone.

As in Shanghai, MAX(TONE)/$\acute{\sigma}$ will preserve the underlying tone of the initial stressed syllable, and *CONTOUR/INITIAL will exclude contour tones on the initial syllable. Crucially, CONTIGUITY(SYL-TONE) has to be outranked by the above two constraints, which defines how the retained tones are actually distributed.

(44) COINCIDE($\acute{\sigma}$, L/H), COINCIDE(FINAL, L/H) >> CONTIGUITY(SYL-TONE)

$\acute{\sigma}_1$ σ_2 σ_3 σ_4 $\wedge$ $\wedge$ $\wedge$ $\wedge$ $L_1 H_1 L_2 H_2 L_3 H_3 L_4 H_4$	COINCIDE ($\acute{\sigma}$, L/H)	COINCIDE (FINAL, L/H)	CONTIGUITY (SYL-TONE)
a. $\acute{\sigma}_1$ σ_2 σ_3 σ_4 $\mid$ $\mid$ L_1 H_1		*!	
b. $\acute{\sigma}_1$ σ_2 σ_3 σ_4 $\mid$ $\mid$ L_1 H_1	*!		*(H_1)
c. ☞ $\acute{\sigma}_1$ σ_2 σ_3 σ_4 $\mid$ $\mid$ L_1 H_1			*(H_1)

Our analysis shows that the effect of edge-in association can be derived from interacting constraints on contour distribution.

As Kennedy (1953) reports, the stress is on the second syllable in disyllabic phrasal constructions. Accordingly, the tone sandhi patterns differ from the compounds which have initial stress.

(45) a. yao zoe   'to row a boat'
      LH LH   citation tone
      o   LH   surface form

    b. bhao moa   'to race horses'
      LH HL
      o   HL

The data shows that it is the stressed syllable, rather than the initial syllable, that preserves its underlying tone. In compounds, the stress is on the first syllable, whose underlying tone is preserved. In verb-object constructions, the stress is on the second syllable and its underlying tone surfaces. The difference between initial stress and final stress lies in how the preserved contour is finally realized. In our analysis, MAX(TONE)/σ́ has the effect of preserving the underlying tone on the stressed syllable, no matter whether it is initial or final. Contour re-association in the initial position is triggered by *CONTOUR/INITIAL, which is irrelevant to the non-initial syllable. COINCIDE(σ́, CONTOUR) will ensure that the stressed syllable carries a contour. When it is ranked below *CONTOUR/INITIAL, only the non-initial stressed syllable can be realized in a contour. The toneless syllables in the following tableaux will be supplied with a low boundary tone, a detail which does not concern us here.

(46) *CONTOUR/INITIAL >> COINCIDE(σ́, CONTOUR)

a. Initial stress in compounds

σ́₁ σ₂ / L₁ H₁ L₂ H₂	*CONTOUR/INITIAL	COINCIDE (σ́, CONTOUR)
a. σ́₁ σ₂ / L₁ H₁	*!	
c. ☞ σ́₁ σ₂ / L₁ H₁		*

b. Final stress in phrases

σ₁ σ́₂ / L₁ H₁ L₂ H₂	*CONTOUR/INITIAL	COINCIDE (σ́, CONTOUR)
a. σ₁ σ́₂ / L₂ H₂		*
c. ☞ σ₁ σ́₂ / L₂ H₂		

6.4.6 Summary

We have shown that both Mandarin and Shanghai exhibit the phonological process of contour re-association, the former targeting the complex contour tone and the latter targeting the simple contour tone. There is a parallel between Mandarin f0 peak delay in rising tone and Shanghai contour tone re-association. It is best characterized as phonetic timing of pitch targets in Mandarin, which finds striking resemblance in the contour tone re-association in Shanghai. In both Mandarin complex contour re-association and Shanghai contour re-association, the contour splits into two parts, one part being realized

on the original host syllable and the other being shifted onto the immediately following syllable. The contour tone in Mandarin does not undergo re-association. In both cases, the low boundary tone is inserted on the last toneless syllable in the relevant domain, and the intervening toneless syllables (if there are) remain toneless and get their transition pitch from both sides.

(47) a. Shanghai: contour re-association and low boundary tone insertion

$$\begin{array}{cccccccc}
\sigma_1 & \sigma_2 & \sigma_3 & \rightarrow & \sigma_1 & \sigma_2 & \sigma_3 \\
\wedge & \wedge & \wedge & & | & | & | \\
T_1\,T_2 & T_3\,T_4 & T_5\,T_6 & & T_1 & T_2 & L\text{-}
\end{array}$$
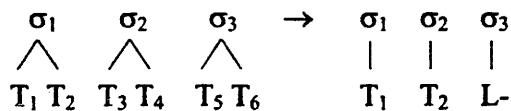
b. Mandarin: contour stability and low boundary tone insertion

$$\begin{array}{cccccccc}
\sigma_1 & \sigma_2 & \sigma_3 & \rightarrow & \sigma_1 & \sigma_2 & \sigma_4 \\
\wedge & \wedge & \wedge & & \wedge & & | \\
T_1\,T_2 & T_3\,T_4 & T_5\,T_6 & & T_1\,T_2 & & L\text{-}
\end{array}$$
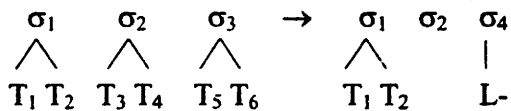
Since realization of the underlying tone on the stressed syllable has priority over low boundary tone insertion in Mandarin and Shanghai, MAX(TONE)/$\acute{\sigma}$ outranks BND(L)/EDGE, which in turn outranks DEP(TONE) to allow boundary tone insertion. The next pattern we examine displays both contour re-association and spreading to the following toneless syllable. Conceivably, either DEP(TONE) or other tonal markedness constraints have to outrank BND(L)/EDGE to block the boundary tone insertion.

6.5   Contour Re-association and Spreading

6.5.1   Neutral Tone in Hai'an

Hai'an county is located northwest of the city of Nantong. Despite its geographical proximity to the vast Wu-speaking region, Hai'an dialect belongs to Jianghuai Mandarin. It has an inventory of six citation tones. All data are taken from Y. Wang (1998).

(48)   Citation tones in Hai'an dialect (Checked tones are underlined.)

register	I	II	III	IV
a. *yin*	21	213	33	3
b. *yang*	35			35

There are three types of disyllabic tone sandhi. Type A involves partial tone neutralization in the first syllable. Type B seems to be a morphologically-triggered tone sandhi, in which the second syllable in some nouns changes its tone to 213 regardless of its tonal context. Type C is triggered by neutral tone. It is the Type C tone sandhi in Hai'an that concerns us in this section.

(49)   Type A tone sandhi

  a.   213 → 21 / __ 35, 35

  b.   213 → 35 / __ 213

The distribution of neutral tone in Hai'an is similar to Mandarin Chinese in that the toneless syllable appears in the non-initial position in a word and its occurrence is a lexical property of the word. Hai'an is different from Mandarin Chinese in how the surface tone pattern is derived in the context of neutral tone syllables. In Mandarin Chinese the stressed syllable before a neutral tone syllable keeps its underlying tone and

the neutral tone syllable is assigned a low boundary tone. In Hai'an underlying contour tones re-associate, and underlying level tones (H, M, L) simply spread to the following toneless syllable. The following table summarizes the surface tone patterns in disyllabic words in which the first syllable carries one of the six citation tones and the second is in neutral tone.

(50) Type C tone sandhi

		σ_1-N	
Ia	21 L	21-1	L-L
Ib	35 MH	33-5	M-H
II	213 LM	21-3	L-M
III	33 M	33-3	M-M
IVa	3 M?	3-3	M-M
IVb	35 MH?	3-5	M-H

(N stands for syllable in neutral tone.)

What we observe immediately is the striking similarity of contour re-association between Hai'an and Shanghai. The neutral tone syllables lose their underlying tones (if they have one). When the first syllable carries a rising tone in citation form, the rising tone spreads over the two syllables; when it carries a level tone, the level tone spreads to the following neutral tone syllable.

(50) Contour re-association and level tone spreading (transcriptions are in pinyin)

a. xue sheng → xue sheng "monk"
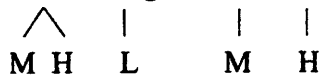
b. ben shi → ben shi "skill"

c. fan ren → fan ren "criminal"

261

It is worth mentioning that we treat 213 as a low rise (LM) in the preceding discussion. There is limited evidence reported in Y. Wang (1998) implying that 213 is a complex contour tone. In a trisyllabic word, the initial 213 spreads to the following two neutral tone syllables.

(52) wo men de → wo men de 'ours'
   ⟋\   ⟋\ | |
   2 1 3   2 1 1 3

However, there are only two trisyllabic examples reported in the paper. It is hard to decide whether this process is due to complex contour spreading or tone attraction to the domain-final position. It is also not clear from the paper whether 213 is significantly longer than the other tones in citation form.

We propose that contour re-association in Hai'an can be derived in similar way as Shanghai.

(53) MAX(TONE)/ớ >> *CONTOUR/INITIAL >> IDENT(TONE)/ớ

	xue sheng MH L	MAX(TONE)/ớ	*CONTOUR/ INITIAL	IDENT(TONE)/ớ
a.	xue sheng MH L		*!	
b.	xue sheng M L	*!		*
c. ☞	xue sheng M H			*

(a) keeps both underlying tones, so it violates *CONTOUR/INITIAL. (b) satisfies the contour licensing constraint, but deleting part of the first tone violates MAX(TONE)/ớ. The rise spreads over the two syllables in (c), in violation of IDENT(TONE)/ớ.

When the stressed syllable carries a level tone, it spreads to the following neutral tone syllable. We propose that tone spreading is triggered by an AGREE constraint, defined below.

(54) a. AGREE(TONE): The tone-bearing units in the same prosodic domain agree in tonal specifications.

   b. AGREE(TONE) >> BND(L)/EDGE

	fan ren M MH	AGREE(TONE)	BND(L)/EDGE
a.	fan ren M L-	*!	
☞b.	fan ren | / M		*

We assume that the relevant domain is a prosodic word, consisting of a stressed syllable and one or two neutral tone syllables. In order to block low boundary tone insertion, AGREE(TONE) has to outrank BND(L)/EDGE.

AGREE(TONE) is dominated by MAX(TONE)/σ so that there is no spreading in a disyllable prosodic word either from the first tone (a) or from the second tone (b), when the first syllable carries a contour tone.

(55) MAX(TONE)/ǒ >> AGREE(TONE)

	xue sheng MH L	MAX(TONE)/ǒ	AGREE(TONE)
a.	xue sheng |/ M	*!	
b.	xue sheng \\\| L	*!	
c. ☞	xue sheng M H		*

The rankings for the constraints we have discussed so far are summarized below:

(56) Ranking summary for Hai'an

$$
\begin{array}{cc}
& \text{MAX(TONE)/ǒ} \\
\diagup & | \\
\text{AGREE(TONE)} & \text{*CONTOUR/INITIAL} \\
| & | \\
\text{BND(L)/EDGE} & \text{IDENT(TONE)/ǒ}
\end{array}
$$
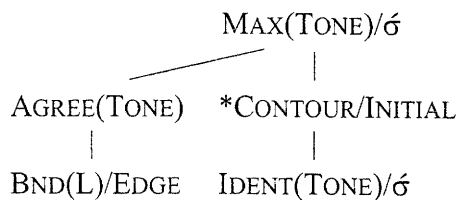
6.5.2   Conflicting Directionality of Tone Spreading in Danyang

Danyang is a northern Wu dialect, which exhibits far more extensive tone spreading than Hai'an. Its complex tone sandhi was reported in extraordinary detail and analyzed in great depth by Lü (1980). Since then, Danyang has been the subject of considerable theoretical interest (Chen 1986, 1991, 1996, 2000, Chan 1989, 1991, 1995, Yip 1989, 1995, Bao 1990, 1999, Duanmu 1990, 1991, 1994, H. Zhang 1992). Our discussion focuses on how the six tone melodies are mapped to a polysyllabic domain.

Danyang distinguishes two slightly different tone systems, one in the literary layer and one in the colloquial layer. Almost all previous analyses of Danyang deal with the

colloquial reading. There are six citation tones in Danyang, described below with their correspondence to the Middle Chinese tonal categories.

(57) Citation tones in Danyang (Colloquial reading)

register	I	II	III	IV
a. *yin*	33 (M)	55 (H)	24 (LH)	3 (M)
b. *yang*	24 (LH)	24 (LH)	11 (LH)	4 (H)

In addition to the six citation tones, a falling tone (HL), transcribed as 42, shows up in tone sandhi context. For example, Lü (1980) mentions a dissimilation process which changes 24 to 42 before another 24: 24 → 42 /___ 24.

Danyang has six tone melodies which may be mapped onto polysyllabic compounds or set phrases. Following Chen (1996, 2000), these tone melodies are labeled from A to F.

(58) Basic tone melodies in Danyang

melodies	σ-σ	σ-σ-σ	σ-σ-σ-σ
A H	H-H	H-H-H	H-H-H-H
B M	M-M	M-M-M	M-M-M-M
C L	L-L	L-L-L	L-L-L-L
D H.L	HL-L	HL-L-L	HL-L-L-L
E L.H	LH-H	LH-H-H	LH-H-H
F HL.LH	HL-LH	HL-HL-LH	HL-HL-HL-LH

According to Lü (1980), all polysyllabic patterns are derived from disyllabic tone sandhi patterns. Disyllabic tone sandhi yields six patterns, identified as A through F in the above table. The six basic tone melodies are based on disyllabic tone sandhi. When two syllables are combined, the choice of one of these six tone sandhi patterns is dependent on the historical source of the input tone and onset voicing of the initial syllable. For

265

more details, please refer to Lü (1980). Duanmu (1994) and Chan (1995) also have in-depth discussions of this issue.

In the OT analysis of tone mapping that we have developed, Danyang tone melody mapping receives a straightforward explanation. We assume that the selection of a specific tone melody is due to an allomorphy process (cf. Tsay and Myers 1996 for an account of Taiwanese tone sandhi), which does not concern us here. Lü (1980), as a native speaker of Danyang, points out that in all cases stress falls on the first syllable, with duration as the most perceptually salient cue and the unstressed syllables are perceived as short and light (Lü 1980:89).[6]

Since no tone insertion or deletion is observed, DEP(TONE) and MAX(TONE) are undominated in Danyang. Any candidates, surfacing with a tone not in the input or not surfacing with a tone in the input, are not considered in the following tableaux. The mappings of patterns A through C are unremarkable and uncontroversial. The input tones spread to all syllables in the domain. In order to make sure that every syllable gets a tone (i.e. no toneless syllable in the output), SPEC(TONE) is also undominated. The tone spreading can be derived as follows.

---

[6] Lü (1980) also describes alternate tone sandhi patterns for trisyllabic and quadrisyllabic compounds, which stress the final syllable.

(59) a. SPEC(TONE): Every tone-bearing unit dominates a tone.

b. MAX(TONE), DEP(TONE), SPEC(TONE) >> BND(L)/EDGE

σ σ σ H	DEP(TONE)	SPEC(TONE)	BND(L)/EDGE

a. σ σ σ
  | |
  H L-
	*!	*	

b. σ σ σ
  ⟋ |
  H L-
	*!		*

c. ☞ σ σ-σ
  ⟋⟍
  H
			*

The tableau shows that boundary tone insertion and toneless syllables are not allowed given the ranking. The winner (c) only violates the low-ranking BND(L)/EDGE. Patterns B and C can be derived similarly.

Patterns D and E are characterized as spreading from the terminal feature node of a contour tone rather than from the higher tonal root node in Yip's (1989a) model of contour tone. This idea is spelled out in Yip (1989b), and illustrated in the following autosegmental representation.

(60) σ   σ   σ   σ   →   σ   σ   σ   σ
     | | | | |
     o   o   o   o   o   tonal root node
    /\   /\ | | |
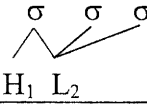   H L   H L L   L   L   terminal feature node
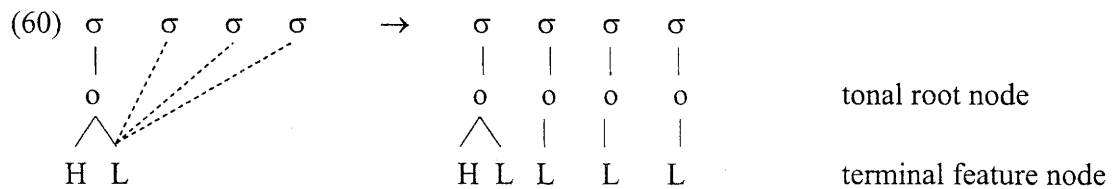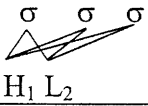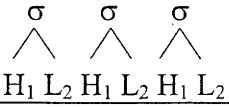
After spreading from the terminal feature node, a process of tier conflation applies, splitting multiple associations into individual associations. With or without tier conflation, this derivation yields the correct output. But one may wonder what prevents spreading from the tonal root node.

Following Chan (1995), we posit that the tone melodies for patterns D and E consist of two level tones: H.L for D, L.H for E. We assume the following high-ranking constraints, derived from the well-formedness conditions proposed by Goldsmith (1976, see also Pulleyblank 1997). Output candidates violating these constraints never survive in Danyang.

(61) a. NO LINE CROSSING: Association lines do not cross.

   b. LINEARITY(TONE): Input precedence relations in a tone melody are preserved in the output. ('No tone metathesis') (after McCarthy and Prince 1995)

   c. INTEGRITY(TONE): No tone in the input has multiple correspondents in the output. (after McCarthy and Prince 1995)

LINEARITY(TONE) maintains the linear order of adjacent tones and rules out metathesis. INTEGRITY(TONE) rules out mapping a tone into multiple copies. The effects of these constraints are illustrated below:

(62) INTEGRITY(TONE), NOLINECROSSING >> AGREE(TONE)

$\sigma\ \sigma\ \sigma$ $H_1.L_2$	INTEGRITY (TONE)	NOLINECROSSING	AGREE(TONE)
a. $\sigma\ \ \sigma\ \ \sigma$ $H_1\ L_2$		**!	
b. $\sigma\ \ \ \sigma\ \ \ \sigma$ $H_1\ L_2\ H_1\ L_2\ H_1\ L_2$	**,**		
c. ☞ $\sigma\ \ \sigma\ \ \sigma$ $H_1\ L_2$			*

The input tones and their output correspondents are indexed. (a) and (b) satisfy AGREE(TONE), but they are ruled out by the unviolated constraints: (a) has two crossing

association lines; (b) makes two more copies of H and L respectively. The winning candidate (c) has one violation of AGREE(TONE) because the first syllable carries a fall (HL), differing from the cther two syllables each carrying a L.

Observing that the initial stressed syllable always surfaces with a contour tone in patterns D and E (the input melody has two tones), we invoke one of the contour licensing constraint COINCIDE(ó, CONTOUR), which restricts a contour tone (rise or fall) to the stressed syllable. This constraint will rule out a crucial competing candidate.

(63) COINCIDE(ó, CONTOUR), AGREE(TONE)

σ σ σ $H_1.L_2$	COINCIDE (ó, CONTOUR)	AGREE(TONE)
a. σ σ σ | |/ H_1 L_2	*!	*
b. ☞ σ σ σ /\\/ H_1 L_2		*

Candidates (a) and (b) both violate AGREE(TONE) once because the first syllable has different tone than other syllables. A violation of COINCIDE(ó, CONTOUR) is assigned to (a) because the first syllable does not carry a contour tone. Recall that stress falls on the first syllable in Danyang polysyllabic compounds. The contour licensing constraint is satisfied in (b). Conceivably, COINCIDE(ó, CONTOUR) can also be satisfied by a candidate with an initial rising tone, as illustrated below. Since DEP(TONE) is unviolated, the only way to get a rise is to metathesize H_1 and L_2 in the input, yielding L_2H_1 on the first syllable. Such a candidate crucially violates LINEARITY, and does not survive.

(64) LINEARITY(TONE), COINCIDE($\acute{\sigma}$, CONTOUR), AGREE(TONE)

$\sigma\ \sigma\ \sigma$ $H_1.L_2$	LINEARITY (TONE)	COINCIDE ($\acute{\sigma}$, CONTOUR)	AGREE(TONE)
a. $\sigma\quad\sigma\quad\sigma$ $L_2\ H_1$	*!		*
b. ☞ $\sigma\quad\sigma\quad\sigma$ $H_1\ L_2$			*

Introducing COINCIDE($\acute{\sigma}$, CONTOUR) does not affect the computation of patterns A,

B and C. Since the input melody has only one level tone (H, M, L), there is no way to

satisfy COINCIDE($\acute{\sigma}$, CONTOUR) except by inserting a tone with opposite value.[7] But any

tone insertion will be ruled out by the undominated Dep(TONE). The following tableau

illustrates this point.

(65) DEP(TONE) >> COINCIDE($\acute{\sigma}$, CONTOUR)

$\sigma\ \sigma\ \sigma$ H_1	DEP(TONE)	COINCIDE ($\acute{\sigma}$, CONTOUR)	AGREE(TONE)
a. $\sigma\quad\sigma\quad\sigma$ $H_1\ L$	*! (L)		*
b. ☞ $\sigma\ \sigma\text{-}\sigma$ H_1		*	

Pattern F has generated a great deal of interest in the literature. A number of

different analyses have been proposed, most of which require a special treatment for

pattern F. Chen (1986) is the earliest formal analysis of Danyang tone melody. He

---

[7] For M, insertion of L or H would create a contour.

makes the following proposal for pattern F: "The most obvious solution that presents itself entails an underlying representation like /HL-LH/ together with a lexically marked *right-to-left mapping* (emphasis added)". The proposal is illustrated in Yip's model of contour tone unit as follows. Given the input as HL-LH, LH as a gestalt is mapped to the rightmost syllable and HL mapped to the penult. Then HL spreads to the leftover syllables to the left.[8]

(66) Chen's (1986) treatment of pattern F
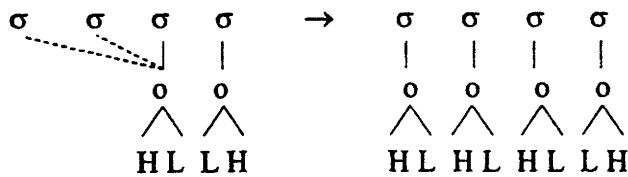
This analysis is unsatisfactory in one respect, as Chen notes himself, namely, "its idiosyncratic directionality of tone spread, from right to left, out of synch with all the other patterns" (Chen 1996:28).

Yip (1989) offers an 'edge-in' account for patter F, according to which HL and LH are first associated with the two end points of the polysyllabic domain, and then the initial HL spreads rightwards to the remaining syllables. This solution gets rid of the idiosyncratic right-to-left mapping unique to pattern F, but it still calls for a special mode of association for pattern F (edge-in, then left-to-right).[9] In addition, it is still not clear in

---

[8] Chen (1986) does not provide a complete analysis of the Danyang case. For other patterns, he simply states that tone mapping goes from left to right, with the rightmost tone segment extended over the leftover syllables, if any.

[9] In a footnote in Chan (1995:182), she quotes Yip (written personal communication) as saying to the effect that tone associations in other five patterns can also be edge-in. See Bao (1999:65ff) for a relevant discussion.

this account why tone spread targets the tonal root node (HL and LH) in pattern F, but only extends the terminal tonal element in patterns D and E.
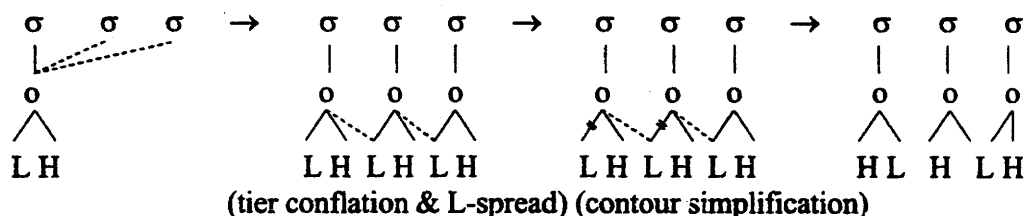
A third analysis we consider is Chan (1989, 1991), whose basic insights are followed with minor modifications in Bao (1990, 1999), and further pursued in Chan (1995). Chan's analysis circumvents the problem of directionality of association. Based on a dissimilation process described by Lü (1980), 24 → 42 /___ 24, she formulates a HL formation process, stated below.

(67) HL formation (Chan 1995:171)
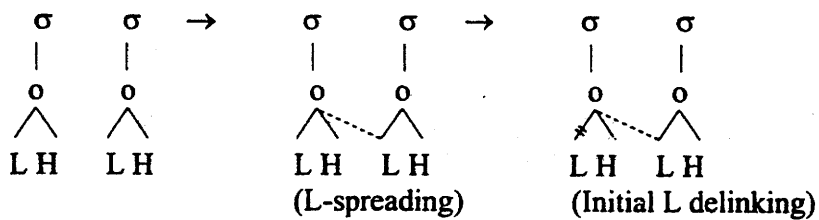
Chan reduces the initial L delinking to a more general process of contour simplification, which prohibits a contour node associated with three tones. The resulting surface sequence after HL formation is HL-LH, which is identical to the disyllabic sequence in pattern E. Consequently, Chan assumes that the tone melody for pattern E is LH, which spreads as a unit across a polysyllabic domain. The idea of tier conflation is crucial in the derivation. This is because HL formation only applies to derive the surface sequence for pattern F *after* each syllable gets a copy of the multiply linked contour tone, illustrated as follows.

(68) Chan's derivation of pattern F (Chan 1995:172)

σ σ σ → σ σ σ → σ σ σ → σ σ σ

o   o   o   o   o   o   o   o   o   o

LH   LH LH LH   LH LH LH   HL H LH

(tier conflation & L-spread) (contour simplification)

Bao (1990, 1999) treats contour dissimilation as metathesis. After rightward spread of LH, contour metathesis applies to change non-final LH to HL, leaving the final LH intact.

In the OT analysis that we are pursuing, pattern F is no different from the other five patterns in that they all are derived from interacting constraints on faithfulness and markedness. We assume that the tone melody for pattern F is HL.LH (alternatively F.R). Pattern F is interesting because one expects F-R-R-R via left-to-right mapping, but the attested sequence is F-F-F-R, conceivably via right-to-left mapping. The analyses reviewed above all try to circumvent the problem of directionality. Recent studies on contour tone distribution (Zoll, 1996, 2003, Gordon 1999, Zhang 2001) reveal the positional asymmetries of contour licensing. For example, following Clark (1983), Zhang (2001) argues that contour tones are more likely to be licensed in domain-final position than non-domain-final position, largely because the former enjoys durational advantage due to final lengthening. The distribution of Mandarin tone 3, discussed earlier, is an example. In addition, there is also crosslinguistic evidence that rising tone is more marked than falling tone in terms of contour distribution (Gordon 1999, Zhang 2001). It also takes longer time to be implemented (Ohaha and Ewan 1973, Sundberg 1979, Xu and Sun 2002).

According to the dual-prominence theory, both stress and edge (initial and final) positions in a prosodic domain are prominent. The input tone seeks out prominent positions in tone realization. We have presented cases in which the retained tone is realized on the stressed syllable. We argue that this is also the case in Danyang: the initial syllable carries the stress and the input melody seeks out the stressed syllable first. In pattern E, the first part of the input melody F is hosted by the initial stressed syllable; the second part R is hosted by another prominent position, domain-final position. In our constraint system, the contour licensing constraints are responsible for restricting contour shapes to different prosodic positions. We propose *RISE/NON-FINAL, defined below. Everything else follows from the constraints and rankings we have established. F and R are used for the sake of visual clarity.

(69) *RISE/NON-FINAL: A rising tone is not allowed on a non-domain-final syllable.

σ σ σ σ F.R	COINCIDE (σ́, CONTOUR)	*RISE/NON-FINAL	AGREE(TONE)

a. σ σ σ σ
   │ \│/
   F   R
		**!	*

b. ☞ σ σ σ σ
   \│/ │
   F   R
			*

Candidate (a) is what one expects to get from left to right mapping, but it violates *RISE/NON-FINAL twice because the second and third syllables are linked with a rising tone. AGREE(TONE) is violated once in both candidates since there are the tone bearing units are associated with two tones instead of one. (b) satisfies *RISE/NON-FINAL by aligning R only with the final syllable, and then spread F to the remaining syllables.

When *RISE/NON-FINAL is satisfied by, it is also possible to spread from the terminal feature node of F rather than from the contour node, as we observed in patterns D and E. Consider the next tableau.

(70) Spreading from terminal feature node

σ σ σ σ F.R	COINCIDE (σ́, CONTOUR)	*RISE/NON-FINAL	AGREE(TONE)

a.   σ   σ   σ   σ
     /\⟋⟋ |
     H L   R
			**!

b. ☞   σ σ σ σ
    \\\/ |
     F   R
			*

In candidate (a), L of the initial F (i.e. HL) spreads to the remaining syllables after R is aligned with the final syllable. However, spreading from the terminal feature node incurs one more violation of AGREE(TONE). Therefore (a) is less harmonic than (b) on the ground that the syllables in (a) have more different tonal specifications.

COINCIDE(σ́, CONTOUR) is not violated by these three crucial competitors. It is ranked over *RISE/NON-FINAL, to be justified next.

*RISE/NON-FINAL has to be ranked lower than COINCIDE(σ́, CONTOUR) so that the initial rise can appear on the initial stressed syllable. Consider pattern E. Candidate (a) satisfies *RISE/NON-FINAL, but violates the high-ranking COINCIDE(σ́, CONTOUR). (b) is the winner, despite its violation of RISE/NON-FINAL.

(71) COINCIDE($\acute{\sigma}$, CONTOUR) >> *RISE/NON-FINAL

	σ σ σ $L_1.H_2$	COINCIDE ($\acute{\sigma}$, CONTOUR)	*RISE/NON-FINAL
a.	σ σ σ $\mid$ $\mid/$ L_1 H_2	*!	
b. ☞	σ σ σ $\bigwedge\!\!\diagup$ L_1 H_2		*

In sum, our analysis provides a simple and straightforward explanation for the intriguing tone mapping in Danyang. The problem of conflicting directionality in tone mapping vanishes in a constraint-based approach (cf. Zoll 1997). The ranking of constraints is summarized in (72):

(72) Ranking summary

    a. DEP(TONE) >> COINCIDE($\acute{\sigma}$, CONTOUR) >> *RISE/NON-FINAL

    b. MAX(TONE), DEP(TONE), SPEC(TONE) >> BND(L)/EDGE

    c. INTEGRITY(TONE), NOLINECROSSING >> AGREE(TONE)

6.6   Concluding Remarks

In this chapter, we have examined diverse patterns of contour re-association and tone spreading in a number of Chinese dialects. It has been shown that they can be derived by the interacting faithfulness and markedness constraints, referring to different prosodic positions. The affinity between the phonetic timing of f0 target and phonological contour re-association suggests that it is possible to that they are subject to the same tonal markedness constraint regulating the placement of pitch targets. Such a constraint

requires that the rightward migration of pitch targets be local. When it is overridden by other positional markedness constraints, non-local migration of tone targets can emerge, as we saw in Tangxi. We also show that the edge-in association can be derived in a constraint-based approach to tone mapping. Its primary motivation is to solve the problem of conflicting directionality in phonological process, in particular tone mapping. It vanishes in an OT framework.

Chapter 7 Conclusion

This dissertation concerns phonetic and phonological aspects of tone mapping in various Chinese languages. The central issue addressed is the role of contrast and positional prominence and positional neutralization in the realization of tone.

The inventory of tonal contrasts is found to constrain the outputs of contextual neutralization in the realization of tone. Specifically, the phonological allotonic selection in Mandarin Chinese is shown to result from the interaction of constraints on preserving tonal contrasts with tonal markedness constraints. Phonetic timing of pitch targets has figured prominently in the recent literature on the realization of tone. This dissertation shows that the detailed phonetic timing of pitch targets can be modeled as emerging from the interaction of phonological constraints on preserving tonal contrasts with phonetic constraints on contour tone durations.

Neutral tone syllables in Mandarin Chinese are unstressed syllables that do not carry any lexical tones on the surface. Drawing on evidence from the observed f0 patterns on multiple neutral tone syllables and their effect on the location of pitch targets on the preceding full-toned syllables, this dissertation proposes that the low pitch target on the last neutral tone syllable is actually a prosodic boundary tone aligned with the right edge of a prosodic domain.

This dissertation also proposes that in order to explain a wide range of tone sandhi phenomena, two prominent positions are distinguished: peripheral (initial and final) positions and metrically strong positions, and they have to be referred to by the

phonology. Specifically, cases of tone movement in which a tone occupying a peripheral position in the input is realized on the stressed syllable in the output (i.e. different from the input position), strongly suggest that both peripheral positions and metrically strong positions are both implicated in the process. An adequate account of tone movement calls for the interaction of positional faithfulness with positional markedness constraints, both of which make reference to different prominent positions.

The dual-prominence theory predicts the interaction of different kinds of prominent positions in tone mapping. In some cases, where an underlying tone is retained is different from where it is eventually realized. With respect to tone retention, the underlying tone of a syllable can be retained when it occurs in different prominent positions. Zhenhai lexical tone sandhi is an example of initial tone preservation while the pattern F in Wenzhou tone sandhi is a case of final tone preservation. In both cases, stress occurs in a different position. Xiamen and Shanghai furnishes examples of preservation of tones on a stressed syllable. In our theory, initial and final tone preservation can be achieved by ranking MAX(TONE)/INITIAL or MAX(TONE)/FINAL over IDENT(TONE)/$\acute{\sigma}$ and MAX(TONE)/$\acute{\sigma}$. Conversely, when IDENT(TONE)/$\acute{\sigma}$ and MAX(TONE)/$\acute{\sigma}$ are high-ranking, the input tones of stressed syllables will be retained.

After an underlying tone is retained, it is generally realized on a stressed syllable in the relevant tone sandhi domain. Tone movement occurs when a tone is retained in one position, but stress falls elsewhere. Diverse patterns have been observed with respect to how a retained tone is eventually realized. There are cases like Zhenhai and Wenzhou in which the retained contour tone moves as a gestalt onto a stressed syllable. What is

more often encountered is contour tone re-association in a domain-initial position. Shanghai instantiates a local migration of tonal targets by which an initial contour tone spreads over the first two syllables whereas Tangxi represents a non-local migration by which the second tonal target of a contour tone seeks out the domain-final syllable. Tone spreading occurs when AGREE(TONE) is active.

Positional asymmetry in contour tone distribution has been reported in the literature. In Chinese tone languages, contour tones are often prohibited from occurring in a domain-initial position, especially in various Wu dialects; they are more tolerated in a domain-final position. A most complex example of tone spreading in Danyang, which allegedly involves conflicting directional mapping, falls out naturally in recognition of positional asymmetry regarding contour tone distribution.

References

Alderete, John. 1995. Faithfulness to prosodic heads. Ms., University of Massachusetts, Amherst. Available at http://roa.rutgers.edu/files/94-0000/94-0000-ALDERETE-0-0.PDF.

Ao, Benjamin. 1992. Metrical constituents as domains of tone sandhi in Nantong Chinese. In *Proceedings of CLS* 28.

Ao, Benjamin. 1993. Phonetics and phonology of Nantong Chinese. Doctoral dissertation, Ohio State University.

Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26:3-25.

Atterer, M., and D. Robert Ladd. 2002. On the phonetics and phonology of "segmental anchoring" of f0: evidence from German. Ms.

Ballard. W. L. 1988. Distribution and historical significance of varying tone sandhi types in South China. *Computational Analysis of Asian and African Languages* 30:43-55.

Bao, Zhiming. 1990. On the nature of tone. Doctoral dissertation, MIT.

Bao, Zhiming. 1999. *The Structure of Tone*. Oxford: Oxford University Press.

Barnes, Jonathan. 2001. Domain-initial strengthening and the phonetics and phonology of positional neutralization. In *Proceedings of NELS 32*.

Beckman, Jill. 1998. Positional faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.

Beckman, Mary E., and Jan Edwards. 1990. Lengthening and shortening and the nature of prosodic constituency. In *Papers in the Laboratory Phonology I: Between the Grammar and the Physics of Speech*, ed. by John Kingston and Mary E. Beckman, 152-178. Cambridge: Cambridge University Press.

Byrd, Dani. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57:3-16.

Cao, Jianfen. 1986. Putonghua qingsheng yinjie texing fenxi [An analysis of neutral tone syllables in Standard Chinese]. *Yingyong Shengxue [Applied Acoustics]*, *Volume 5, Number 4*.

Cao, Jianfen, and Ian Maddieson. 1992. An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics* 20:77-92.

Chan, Marjorie K. M. 1989. On the status of base tone. *UCLA Working Papers in Phonetics* 63:71-94.

Chan, Marjorie K. M. 1991. Contour-tone spreading and tone sandhi in Danyang Chinese. *Phonology* 8:237-259.

Chan, Marjorie K. M. 1995. An autosegmental analysis of Danyang tone sandhi: some historical and theoretical issues. In Wuyu Yanjiu [Studies of the Wu Dialects], ed. by Xu Yunyang (Eric Zee), 145-183. New Asia Academic Bulletin, Vol. XI, Hongkong Chinese University.

Chan, Marjorie K. M., and Ren Hongmo. 1985. Wuxi tone sandhi: from last to first syllable dominance. *UCLA Working Papers in Phonetics* 63:48-70.

Chao, Yuen-Ren. 1930. A system of tone letters. *La Maitre Phonetique* 45:24-47.

Chao, Yuen-Ren. 1948. *Mandarin Primer: An Intensive Course in Spoken Chinese.* Cambridge, MA: Harvard University Press.

Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese.* Berkeley: University of California Press.

Chen, Matthew Y. 1976. From Middle Chinese to Modern Peking. Journal *of Chinese Linguistics* 4:113-277.

Chen, Matthew Y. 1986. An overview of tone sandhi phenomena across Chinese dialects. Paper presented at the *Conference on Languages and Dialects of Chinese,* University of California, Berkeley.

Chen, Matthew Y. 1987. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4:109-149.

Chen, Matthew Y. 1992. Tone rule typology. In *Proceedings of the Special Session on the Typology of Tone Languages, BLS 18,* 54-66.

Chen, Matthew Y. 1996. Tone geometry – a Chinese perspective. In New Horizons in Chinese Linguistics, ed. by C.-T. James Huang and Audrey Li, 21-48. Dordrecht: Kluwer.

Chen, Matthew Y. 2000. *Tone Sandhi: Patterns Across Chinese Dialects.* Cambridge: Cambridge University Press.

Cheng, Chinchuan. 1973. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

Cho, Taehong, and Patricia Keating. 2001. Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29:155-190.

Clark, Mary M. 1983. On the distribution of contour tones. In *Proceedings of WCCFL 2*, ed. by Michael Barlow, Daniel P. Flickinger, and Michael T. Wescoat, 281-339. Stanford, Calif.: Stanford University, Stanford Linguistics Association.

Cohn, Abigail. 1990. Phonetic and phonological rules of nasalization. Doctoral dissertation, University of California, Los Angeles.

de Lacy, Paul. 2002. The interaction of tone and stress in Optimality Theory. *Phonology* 19:1-32.

Dilley, Laura, Stefanie Shattuck-Hufnagel, and Mari Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24: 423-444.

Dong, Shaowen. 1958. *Yuyin Changshi [Common Sense in Phonetics]*. Culture and Education Press.

Duanmu, San. 1990. A formal study of syllable, tone, stress and domain in Chinese languages. Doctoral dissertation, MIT.

Duanmu, San. 1992. An autosegmental analysis of tone in four Tibetan languages. *Linguistics of the Tibeto-Burman Area* 15.1:65-91.

Duanmu, San. 1993. Rime length, stress, and association domains. *Journal of East Asian Linguistics* 2:1-44.

Duanmu, San. 1994. Against contour tone units. *Linguistic Inquiry* 25:555-608.

Duanmu, San. 1995. Metrical and tonal phonology of compounds in two Chinese dialects. *Language* 71:225-259.

Duanmu, San. 1999. Metrical structure and tone: evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8(1):1-38.

Duanmu, San. 2000. *Phonology of Standard Chinese*. Oxford: Oxford University Press.

Edwards, Jan, Mary E. Beckman, and Janet Fletcher. 1991. The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America* 89:369-382.

Flemming, Edward. 1995. Auditory representation in phonology. Doctoral dissertation, University of California, Los Angeles.

Flemming, Edward. 1997. Phonetic optimization: compromise in speech production. *University of Maryland Working Papers in Linguistics* 5:72-91.

Flemming, Edward. to appear. Contrast and perceptual distinctiveness. In *The Phonetic Bases of Phonological Markedness*, ed. by Bruce Hayes, Robert Kirchner and Donca Steriade. Cambridge: Cambridge University Press.

Fougeron, Cecile, and Patricia Keating. 1996. Articulatory strengthening in prosodic domain-initial position. *UCLA Working Papers in Phonetics* 92:61-87.

Fougeron, Cecile. 1999. Articulatory properties of initial segments in several prosodic constituents in French. *UCLA Working Papers in Phonetics* 97:74-99.

Gårding, E., P. Kratochvil, J.-O. Svantesson, and Zhang Jialu. 1986. Tone 4 and tone 3 discrimination in modern standard Chinese. *Language and Speech* 29-3.

Goldsmith, John. 1976. Autosegmental phonology. Doctoral dissertation, MIT.

Gordon, Matthew. 1999. Syllable weight: phonetics, phonology, and typology. Doctoral dissertation, University of California, Los Angeles.

Gordon, Matthew. 2001. A typology of contour tone restrictions. *Studies in Language* 25:423-462.

Greenberg, Steven, and Eric Zee. 1979. On the perception of contour tones. *UCLA Working Papers in Phonetics* 45:150-164.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago.

House, David. 1990. *Tonal Perception in Speech*. Lund: Lund University Press.

Hsiao, Franny. 2002. Tonal domains are stress domains in Taiwanese – evidence from focus. In *MIT Working Papers in Linguistics* 42: *Phonological Answers (and their Corresponding Questions)*, ed. by Anikó Csirmaz, Zhiqiang Li, Andrew Nevins, Olga Vaysman and Michael Wagner, 109-140..

Huang, Tsan. 2001. The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. *OSU Working Papers in Linguistics* 55, 23-42.

Hyman, Larry M., and Russell G. Schuh. 1974. Universals of tone rules: evidence from West Africa. *Linguistic Inquiry* 5:81-115.

Ito, Junko. 1986. Syllable theory in prosodic phonology. Doctoral dissertation, University of Massachusetts, Amherst.

Ito, Junko. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7:217-259.

Jiang, Hui. 1994. The phonetic description of neutral tone in Tianjin dialect. M.A. dissertation, Tianjin Normal University.

Keating, Patricia. 1985. Universal phonetics and the organization of grammars. In *Phonetic Linguistics: Essays in honor of Peter Ladefoged*, ed. by Victoria A. Fromkin, 115-132. Orlando: Academic Press.

Keating, Patricia. 1996. The phonology-phonetics interface. *UCLA Working Papers in Phonetics* 72:45-60.

Keating, Patricia, Taehong Cho, Cecile Fougeron, Chai-Shue Hsu. 1999. Domain-initial strengthening in four languages. *UCLA Working Papers in Phonetics* 97:139-151.

Kennedy, George A. 1953. Two tone patterns in Tangsic. *Language* 29:367-373.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell Publishers.

Kenstowicz, Michael, and George Akanlig-Pare. 2003. Tone in Buli. *Working Papers on Endangered and Less Familiar Languages* 4, MIT.

Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14:83-111.

Klatt, Dennis. 1975. Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics* 3:129-140.

Klatt, Dennis. 1976. Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59: 1208-1221.

Laniran, Yetunde O. 1992. Intonation in tone languages: the phonetic implementation of tones in Yoruba. Doctoral dissertation, Cornell University.

Ladd, D. Robert, D. Faulkner, H. Faulkner, and Astrid Schepman. 1999. Constant "segmental anchoring" of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106:1543-1554.

Ladd, D. Robert, I. Mennen, and A. Schepman. 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America* 107:2685-2696.

Ladd, D. Robert, and Astrid Schepman. 2003. "Sagging transition" between high pitch accents in English: experimental evidence. *Journal of Phonetics* 31:81-112.

Laniran. Yetunde O., and George N. Clements. 2003. Downstep and high raising: interacting factors in Yoruba tone production. *Journal of Phonetics* 31:203-250.

Laniran, Yetunde O., and G. N. Clements. 2003. Downstep and high raising: interacting factors in Yoruba tone production. *Journal of Phonetics* 31:203-250.

Leben, William. 1973. Suprasegmental phonology. Doctoral dissertation, MIT.

Li, Xingjian, and Liu Sixun. 1985. Tianjin fangyan de liandu biandiao [Tone sandhi in the Tianjin dialect]. *Zhongguo Yuwen* 1:76-80.

Li, Zhiqiang. 2002. A perceptual account of asymmetries in tonal alignment. In *Proceedings of NELS 33*.

Liberman, Mark, and Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In *Language Sound Structure*, ed. by M. Aronoff and R. Oehrle, 157-233. Cambridge MA.: MIT Press.

Lin, Jowang. 1994. Lexical government and tone group formation in Xiamen Chinese. *Phonology* 11:237-276.

Lü, Shuxiang. 1980. Danyang fangyan de shengdiao xitong [The tonal system of Danyang dialect]. *Fangyan* 2:85-122.

Lin, Maocan. 1988. Putonghua shengdiao de shengxue texing he zhijue zhengzhao [Acoustic characteristics and perceptual cues of tones in Standard Chinese]. *Zhongguo Yuwen* 3:182-193.

Lin, Maocan, and Yan Jingzhu. 1980. Beijinghua qingsheng de shengxue xingzhi [Acoustic characteristics of neutral tone in Beijing Mandarin]. *Fangyan* 3:166-78.

Lin, Maocan, and Yan Jingzhu. 1988. The characteristic features of the final reduction in the neutral tone syllable of Beijing Mandarin. *Phonetic Laboratory Annual Report of Phonetic Research*, 37-51. Beijing: Phonetic Laboratory, Institute of Linguistics, Chinese Academy of Social Sciences.

Lin, Maocan, Yan Jingzhu, and Sun Guohua. 1984. Beijinghua liangzizu zhengchang zhongyin de chubu shiyan [Preliminary experiments on the normal stress in Beijing disyllables]. *Fangyan* 1:57-73.

Lin, Tao. 1962. Xiandai haiyu qingyin he jufa jiegou de guanxi [The relation between neutral tone and syntactic structure in Modern Chinese]. *Zhongguo Yuwen* 7.

Lin, Tao. 1985. Tantao Beijinghua qingsheng de chubu shiyan [Preliminary experiments in the exploration of neutral tone in Beijing Mandarin]. In *Beijing Yuyin Shiyanlu [Peking University Working Papers in Phonetics]*, ed. by Lin Tao and Wang Lijia, 1-26. Beijing: Beijing University Press.

Lombardi, Linda. 1991. Laryngeal features and laryngeal neutralization. Doctoral dissertation, University of Massachusetts, Amherst.

Lombardi, Linda. 1995. Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory* 13:39-74.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in optimality theory. *Natural Language and Linguistic Theory* 17:267-302.

Lubowicz, Anna. 2003. Contrast preservation in phonological mappings. Doctoral dissertation, University of Massachusetts, Amherst.

Manuel, Sharon Y. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88:1286-1298.

McCarthy, John, and Alan Prince. 1993. Generalized alignment. *Yearbook of Morphology* 1993:79-153.

McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh Dickey, and Susan Urbanczyk, 249-384. Amherst, MA.: University of Massachusetts, GLSA.

Meredith, Scott. 1990. Issue in the phonology of prominence. Doctoral dissertation, MIT.

Milliken, Stuart. 1989. Why there is no third tone sandhi rule in Standard Mandarin. Paper presented at the *Tianjin International Conference on Phonetics and Phonology*. Tianjin Normal University.

Mei, Tsu-lin. 1977. Tones and tone sandhi in 16th century Mandarin. *Journal of Chinese Linguistics* 5:237-60.

Myers, Scott. 1997. OCP effects in optimality theory. *Natural Language and Linguistics Theory* 15:847-892.

Myers, S. to appear. F0 timing in Kinyarwanda. *Phonetica*.

Ohala, John, and William Ewan. 1973. Speed of pitch change. *Journal of the Acoustical Society of America* 53:345(A).

Oller, D. K. 1973. The duration of speech segments: the effect of position in utterance and word length. *Journal of the Acoustical Society of America* 54:1235-1273.

Padgett, Jaye. 1997. Perceptual distance of contrast: vowel height and nasality. In *Phonology at Santa Cruz, Volume 5*, ed. by R. Walker, M. Katayama and D. Karvonen, 63-78. Linguistic Research Center, University of California, Santa Cruz.

Padgett, Jaye. 2001. Contrast dispersion and Russian palatalization. In *The Role of Speech Perception in Phonology*, ed. by E. V. Hume and K. Johnson, 187-218. San Diego, CA.: Academic Press.

Padgett, Jaye. 2003. Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21:39-87.

Pierrehumbert, Janet, and David Talkin. 1992. Lenition of /h/ and glottal stop. In *Papers in Laboratory Phonology II: Gestures Segment Prosody*, ed. by G. Doherty and D. Robert Ladd, 90-117. Cambridge: Cambridge University Press.

Prieto, P., J. van Santen, and J. Hirschberg. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23:429-451.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms., Rutgers University, New Brunswick, N.J., and University of Colorado, Boulder.

Pulleyblank, Douglas. 1986. *Tone in Lexical Phonology*. Dordrecht: Kluwer.

Pulleyblank, Douglas. 1997. Optimality theory and features. In *Optimality Theory: An Overview*, ed. by Diana Archangeli and D. Terence Langendoen, 59-101. Oxford: Balckwell.

Qian, Nairong. 1992. *Dangdai Wuyu Yanjiu [Studies in Contemporary Wu Dialects]*. Shanghai: Shanghai Jiaoyu Chubanshe [Shanghai Education Press].

Qian, Zengyi. 1995. Jinan fangyan cidian yinlun [Introduction to the Jinan dialect dictionary]. *Fangyan* 4:242-256.

Qian, Zengyi. et al. 1982. *Yantai fangyan baogao* [*A Report on the Yantai dialect*]. Jinan: Qilu Shushe.

Rose, Phil. 1990. Acoustics and phonology of complex tone sandhi. *Phonetica* 47:1-35.

Rose, Phil. 1994. Renewing the connection: acoustics and phonology of two lexical trisyllabic tone patterns in Zhenhai. In *Current Issues in Sino-Tibetan Linguistics*, ed. by H. Kitamura et al., 412-18.

Rose, Phil. 2000. Wenzhou dialect disyllabic lexical tone sandhi with first syllable entering tones. In *Proceedings of the 8th Australian International Conference on Speech Science and Technology*, ed. by M. Barlow, 230-235. Canberra: Australian Speech Science and Technology Association.

Rose, Phil. 2002. Independent depressor and register effects in Wu dialect tonology: evidence from Wenzhou tone sandhi. *Journal of Chinese Linguistics* 30:39-81.

Rose, Phil., and Takako Toda. 1994. A typology of tone sandhi rules in northern Wu. In *Current Issues in Sino-Tibetan Linguistics*, ed. by H. Kitamura et al., 267-273.

Sanders, Nathan. 2002. Dispersion in OT: color contrast in Middle Polish nasal vowels. In *Proceedings of WCCFL 21*, ed. by L. Mikkelsen and C. Potts, 415-428. Somerville, MA.: Cascadilla Press.

Selkirk, Elisabeth. 1994. Optimality theory and featural phenomena. Lectures notes, University of Massachusetts, Amherst.

Selkirk, Elisabeth. 1995. The prosodic structure of function words. In *Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk, 439-70. Amherst, MA.: University of Massachusetts, GLSA.

Selkirk, Elisabeth, and Tong Shen. 1990. Prosodic domains in Shanghai Chinese. In *The Phonology-Syntax Connection*, ed. by Sharon Inkelas and Draga Zec, 313-337. Chicago: University of Chicago Press.

Shen, Tong. 1981a. Laopai Shanghai fangyan de liandu biandiao [Tone sandhi in old Shanghai dialect]. *Fangyan* 2:131-144.

Shen, Tong. 1981b. Shanghaihua laopai xinpai de chabie [Differences between old and new Shanghai dialects]. *Fangyan* 4:275-283

Shen, Tong. 1982. Laopai Shanghai fangyan de liandu biandiao [2] [Tone sandhi in old Shanghai dialect, part 2]. *Fangyan* 2:100-114.

Shih, Chilin. 1986. The prosodic domain of tone sandhi in Chinese. Doctoral dissertation, University of California, San Diego.

Shih, Chilin. 1997. Mandarin third tone sandhi and prosodic structure. In *Studies in Chinese Phonology*, ed. by Wang Jialing and Norval Smith, 81-123. Berlin: Mouton de Gruyter.

Shih, Chilin, and Richard Sproat. 1992. Variations of the Mandarin rising tone. In *Proceedings of the IRCS Workshop on Prosody in Natural Speech* 92(37), 193-200. University of Pennsylvania.

Silverman, Kim, and Janet Pierrehumbert. 1990. The timing of prenuclear high accents in English. In *Papers in the Laboratory Phonology I: Between the Grammar and the Physics of Speech*, ed. by John Kingston and Mary E. Beckman, 72-106. Cambridge: Cambridge University Press.

Steele, S. A. 1986. Nuclear accent f0 peak location: effects of rate, vowel, and number of following syllables. *Journal of the Acoustical Society of America*, Supplement 1, 80:S51.

Steriade, Donca. 1995. Positional neutralization and the expression of contrast. Ms., University of California, Los Angeles.

Steriade, Donca. 1997. Phonetics in phonology: the case of laryngeal neutralization. Ms., University of California, Los Angeles.

Steriade, Donca. 2001. The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms., University of California, Los Angeles.

Struijke, Caro. 2002. *Existential Faithfulness: A Study of Reduplicative TETU, Feature Movement, and Dissimilation*. London: Routlege.

Sundberg, Jodhan. 1979. Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7:71-79.

Tsay, Jane, and James Myers. 1996. Taiwanese tone sandhi as allomorph selection. In *Proceedings of BLS* 22, 394-405. Berkeley, California.

Wang, Futang. 1959. Shaoxinghua Jiyin [Transcriptions of Shaoxing dialect]. *Linguistic Forum, Volume 3*.

Wang, Jialing. 1997. The representation of the neutral tone in Chinese Putonghua. In *Studies in Chinese Phonology*, ed. by Wang Jialing and Norval Smith, 157-183. Berlin: Mouton de Gruyter.

Wang, Maolin. 1995. An analysis of tones in Cangzhou dialect. M.A. dissertation, Tianjin Normal University.

Wang, William S.-Y. 1967. Phonological features of tone. *International Journal of American Linguistics* 33.2:93-105.

Wang, William S.-Y., and Li Kung-Pu. 1967. Tone 3 in Pekinese. *Journal of Speech and Hearing Research* 10.3:629-36.

Wang, Yunjia. 1998. Hai'an qingsheng yu feiqingsheng guanxi chutan [A preliminary study of the relation between neutral tone and non neutral tone in Hai'an]. *Fangyan* 3:211-217.

Wightman, Collin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91:1707-1717.

Woo, Nancy. 1969. Prosody and phonology. Doctoral dissertation, MIT.

Wright, Martha. 1983. A metrical approach to tone sandhi in Chinese dialects. Doctoral dissertation, University of Massachusetts, Amherst.

Xu, Baohua, Tang Zhenzhu, and Qian Nairong. 1981. Xinpai Shanghai fangyan de liandu biandiao [Tone sandhi in new Shanghai]. *Fangyan* 2:145-155.

Xu, Baohua, Tang Zhenzhu, You Rujie, Qian Nairong, Shi Rujie, and Shen Yaming. 1988. *Shanghai Shiqu Fangyanzhi* [*Urban Shanghai Dialects*]. Shanghai: Shanghai Jiaoyu Chubanshe [Shanghai Education Press].

Xu, Yi. 1993. Contextual tonal variation in Mandarin Chinese. Doctoral dissertation, University of Connecticut.

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

Xu, Yi. 1998. Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica* 55:179-203.

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

Xu, Yi. 2001. Fundamental frequency peak delay in Mandarin. *Phonetica* 58:26-52.

Xu, Yi, and E. Wang. 2001. Pitch targets and their realization: evidence from Mandarin Chinese. *Speech Communication* 33:319-337.

Xu, Yi, and X. Sun. 2002. Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111:1399-1413.

Yip, Moira. 1980. Tonal phonology of Chinese. Doctoral dissertation, MIT.

Yip, Moira. 1988. Template morphology and the direction of association. *Natural Language and Linguistic Theory* 6:551-577.

Yip, Moira. 1989. Contour tones. *Phonology* 6:149-174.

Yip, Moira. 1995. Tone in East Asian Languages. In *Handbook of Theoretical Phonology*, ed. by John Goldsmith, 476-494. Oxford: Blackwell Publishers.

Yip, Moira. 1999. Feet, tonal reduction and speech rate at the word and phrase level in Chinese. In *Phrasal Phonology*, ed. by Rene Kager and Wim Zonneveld, 171-194. Nijmegen: Nijmegen University Press.

Yip, Moira. 2001. The complex interaction of tone and prominence. In *Proceedings of NELS 31*, ed. by Minjoo Kim and Uri Strauss, 531-545. Amherst, MA: University of Massachusetts, GLSA.

Yip, Moira, 2002. *Tone.* Cambridge: Cambridge University Press.

Yue-Hashimoto, Anne O. 1987. Tone sandhi across Chinese dialects. In *Wang Li Memorial Volumes: English Volume*, ed. by the Chinese Language Society of Hong Kong, 445-474. (Original paper presented at the 13th International Conference on Sino-Tibetan Languages and Linguistics, Charlottesville, Virginia, 1980.)

Zee, Eric, and Ian Maddieson. 1979. Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *UCLA Working Papers in Phonetics* 45:93-129.

Zhang, Jie. 2001. The effects of duration and sonority on contour tone distribution: Typological survey and formal analysis. Doctoral dissertation. University of California, Los Angeles.

Zhang, Hongming. 1992. Topics in Chinese phrasal phonology. Doctoral dissertation, University of California, San Diego.

Zhengzhang, Shangfang. 1964a. Wenzhou yinxi [The sound system of Wenzhou]. *Zhongguo Yuwen*, 28-49.

Zhengzhang, Shangfang. 1964b. Wenzhou fangyan de liandu biandiao [Tone sandhi in the Wenzhou dialect]. *Zhongguo Yuwen*, 106-152.

Zhengzhang, Shangfang. 1980. Wenzhou fangyan erweici de yuyin bianhua [Sound change of the er-suffix in the Wenzhou dialect]. *Fangyan* 4:245-62.

Zhu, Xiaonong. 1995. Shanghai tonetics. Doctoral dissertation, Australian National University, Canberra.

Zoll, Cheryl. 1996. Parsing below the segment in a constraint based framework. Doctoral dissertation, University of California, Berkeley.

Zoll, Cheryl. 1997. A note on multiple prominence and tone mapping. In *PF: Papers at the Interface*, ed. by Benjamin Bruening, Yoonjung Kang and Martha McGinnis, 97-111. *MIT Working Papers in Linguistics*, Vol. 30.

Zoll, Cheryl. 1998. Positional asymmetries and licensing. Ms., MIT. Available at http://roa.rutgers.edu/files/282-0998/roa-282-zoll-4.pdf.

Zoll, Cheryl. 2003. Optimal tone mapping. *Linguistic Inquiry* 34:225-268.