5 The Tonal System of Shaoxing

5.1 Introduction

Tone exists in all languages, but it is not always phonemic. According to the definition that a tone language is a language in which pitch is used to contrast individual lexical items or words (Gandour 1978: 41), or in which an indication of pitch enters into the lexical realization of at least some morphemes (Hyman 2001: 1367), about 60 to 70 % of the world's languages are tone languages (Yip 2002). In tone languages, pitch is divorced from stress and prominence. It has been recognized since McCawley (1970, 1978) that tone realization in phonological phrases has properties analogous to stress-accent realization (Downing 2003). De Lacy (2002) explores the interaction of tone and stress and claims that tone can influence main stress placement. The relationship between tone and stress will be discussed later in this chapter. However, sometimes the term of tone language is restricted to languages in which virtually every syllable receives a tone, such as Chinese, while languages in which only some syllables receive tones are partial tone languages. In the canonical case, the tone on each syllable is independent of the tones on other syllables and hence the tone of each syllable must be specified separately. SX, as one of the more than 900 Chinese dialects, is a tone language which is believed still to retain the full tone system of Middle Chinese (spoken between the 6th and 10th centuries A.D.).

In a tone language, a difference in pitch may correspond with a difference in the lexical meaning of a word which is otherwise segmentally identical, as shown in the following examples from SX:

(1)	[tuŋ^{52}]	'east'
	[tuŋ^{35}]	'understand'
	[tuŋ^{33}]	'freeze'

The examples in (1) show that the three words form a minimal triplet where the only difference is tone. In the past few decades, it has been widely accepted in phonology that tones are autosegmentally represented and that tones are independent of the segments on which they realized (Leben 1973; Goldsmith 1976; Bao 1999). It has also been assumed that contour tones are composed of sequences of level tones (Leben 1973;

Goldsmith 1976), although the internal structure of tones might differ from language to language. There are also cross-linguistic communalities with respect to the representation of tonal behaviour, such as the Universal Association Conventions (UACs) (Pulleyblank 1986) and the Well-formedness Conditions (WFCs) [1] (Goldsmith 1976), which are well supported from a number of tone language studies.

In this chapter, I will briefly introduce some proposals on the topic of tonal structure and will then present my approach to SX tone structure in feature geometry. I will also present my analysis of the tonal inventory of SX and propose feature specifications for the tones in SX. I make an attempt to formalize the tone sandhi processes that operate in SX, assuming that tone sandhi is phonologically realized by tone feature delinking and spreading, while observing the constraint against crossing association lines. I also assume that tone sandhi in SX may be lexically or syntactically conditioned by metrical structure, which provides the stress foot as the domain for tone sandhi.

5.2 Traditional Tone Representations

In traditional Chinese phonology, tones are divided into a *yin* register and a *yang* register, referring to the high register and the low register, respectively. Historically, the *yin* tones occur on syllables with voiceless initial obstruents and the *yang* tones occur on syllables with voiced initial obstruents. Both the *yin* and *yang* registers are further classified into four tonal categories: *ping*, *shang*, *qu* and *ru*, literally ‘even’, ‘rising’, ‘going’ and ‘entering’, respectively.[2] This can be summarized as in (2):

[1] The UACs (Pulleyblank 1986) hold that tones are associated with syllables from left to right, in a one-to-one fashion. The WFCs (Goldsmith 1976) postulate that association lines may not cross and that every mora must be associated with (at least) one tone.

[2] It is believed that the terms for the four Chinese tones were first introduced in the Six Dynasties period (220-589A.D.). In Ancient Chinese, the *ping* tone was probably a mid flat tone; the *shang* tone a high flat tone; the *qu* tone a long and low tone; the *ru* tone a short and clipped tone. If some modern dialects have the same tones as the ancient ones, it is completely by coincidence, as these tones have probably changed many times before reaching their present forms. However, the terms have been retained, though they no longer mean what they sound like, except for the *ru* tone which almost retains the same Ancient Chinese tonal properties in the dialects which have it.

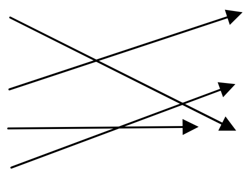
(2)

yin (high) register	yang (low) register
a. *ping* (even)	e. *ping* (even)
b. *shang* (rising)	f. *shang* (rising)
c. *qu* (going)	g. *qu* (going)
d. *ru* (entering)	h. *ru* (entering)

The phonetic tone pitches of Chinese were first transcribed in a numeric notational system by the Chinese scholar Chao in 1930. Chao uses a scale of pitch within an individual speaker's tone range. For graphical representation, a vertical reference line extending from points 1 to 5 is set up, to which a simplified tone graph is attached (see Chao 1930: 24-27). Both the actual intervals and the absolute pitch are relative to the individual voice and the key and mood at the moment of speaking (Chao 1968: 25-26). The five levels are therefore relative pitches. Since Chinese (Mandarin) uses the complete pitch range, all five levels are needed for the description of Chinese tones. Chao's system divides the pitch scale into five distinct levels, from the highest [5] to the lowest [1], as shown with some tones in SX in (3):

(3)

5-level tone pitch scale	tones
5	[52]
4	
3	[35]
2	[22]
1	[13]



There have been many discussions on how many tone levels are needed to describe all languages (e.g. Chao 1930; Wang 1967; Halle & Stevens 1971; Anderson 1978; Hyman 1986; among others). Phonetically, pitch is the primary perceptual correlate of tone and in real speech there can be many pitch levels (Duanmu 2000b). When it comes to phonemic levels which are distinctive, the number of levels is quite small. In most African languages, there are two phonemic levels, H and L. In Asian languages, three or four contrastive levels are quite common (Duanmu 2000b), while five contrastive levels have also been reported.[3]

[3] It is extremely rare that a language has five distinctive level tones. However, it is assumed (Anderson 1978: 145) that Black Miao (a language of a minority nationality in South China) contains five distinctive level tones, i.e. [11], [22], [33], [44] and [55].

Let us see how the different pitch levels can be represented by using distinctive features. If a language has five level tone pitches and, like SX, has a high-low register division, in the high register, tones at level 5 and 4 are usually regarded as high pitches, and the levels from 3 to 1 are regarded as low pitches. In the low register, the highest pitch is level 3, which is regarded to have the [high] tone feature in that register and level 2 and 1 are low pitches. If a tone has a contour pitch [42], it is regarded as a high-register (*yin*) tone, while a [13] tone is likely to be classified as a low-register (*yang*) tone. Obviously, [55] is a high level tone and [22] is a low level tone.

Tones in Middle Chinese were strictly divided into the *yin* and *yang* registers (which correlate historically with voiceless initial obstruents and voiced initial obstruents), classified as in the table in (2). In later times, great changes took place in the phonology of the Chinese languages and the four Middle Chinese tones underwent various splits and mergers. Tone split is sensitive to various phonological conditions, most notably the voicing contrast in the syllable onset, as illustrated by modern SX. In some Chinese dialects, the eight tones merged into a smaller number and also lost the division between high and low registers. For example, in Beijing Mandarin, all voiced obstruents became voiceless, all checked syllables (ending in stops) lost their stop endings entirely, while *yang shang* merged into *yin qu*, and the *ru* tones redistributed among other tonal categories (see Chen 2000: 9). SX has still retained the historically voiced and voiceless distinction in its initial obstruents and has eight tones, as shown in (2), strictly divided into the four high-register tones and the other four low-register tones, which resulted from the historical tonogenesis of Middle Chinese. We will return to this in detail later in this chapter.

5.3 Specification for Tones

There are two obvious and important reasons for considering the question how tones should be represented phonologically and phonetically: feature specifications may be required in the formulation of the phonological rules of a language, including rules that affect tones; second, feature specifications may be required for the phonetic implementation of the tones in a language (Pulleyblank 1986). There have been different proposals with respect to tonal specifications. Chen (2000: 96) gives a brief introduction of different tone feature systems proposed by different authors. These include feature systems incorporating [high], [central] and

[mid] (Wang 1967), [high], [low] and [central] (Sampson 1969), [high], [low] and [modify] (Woo 1969), [stiff vocal cords] and [slack vocal cords] (Halle & Stevens 1971), [high], [low] and [extreme] (Maddieson 1972), and [upper] and [raised] (Yip 1980). Chen (2000) employs H, M and L in his tonal representation. We will not discuss these proposals in detail but concentrate on those aspects of tonal representation that we need for SX. However, on purpose, I present my analysis of tonal representation by means of different feature systems, viz. in a three-level system (H, M, L), register approach ([upper] and [raised]) and laryngeal features ([stiff] and [slack]). Through the comparison below, I argue that the three-level system cannot adequately represent the SX tonal system. Mainly, I employ the laryngeal feature system to explain the consonant-tone correlation and register-tone feature system to formalize the tone sandhi rules.

Phonetically, the production of tone is a function of the vocal cords, which, phonologically, involves two features, [stiff] and [slack]. In articulatory terms, stiffness of the vocal cords induces high tones (and voiceless segments) and slackness induces low tones (and voiced obstruents) (see Halle & Stevens 1971; Maddieson 1984b). The tone features and the consonant-tone correlation will be discussed later. Yip (1980, 1989, 2002) proposes that the features [stiff] and [slack] are equivalent to [upper] and [raised] in her Register theory, in which [upper] is used as a register feature and [raised] is used for particular tone features, as shown in (4):

(4) Yip's revised (2002) proposal

Register	Tone
+Upper (H)	+ raised (h)
	− raised (l)
−Upper (L)	+ raised (h)
	− raised (l)

Where H = [+stiff]
L = [−stiff]
h = [−slack]
l = [+slack]

The features in (4) can specify up to four level tones, which are classified into two different registers. Usually, the number of features needed to represent tones in a language depends on the number of distinctive level tones it has underlyingly, because contours can often be analyzed as sequences of level tones. Snider (1999) proposes four level-tone features, [h.H], [h.L], [l.H] and [l.L],[4] to specify the four level tones Hi, Mid_2, Mid_1 and Lo, respectively, in his analysis of some African languages (see the discussion in §5.3 in this chapter). However, in most cases, tones can be specified with just three features: [High], [Mid] and [Low], as Chen (2000) proposes in his analysis of tone sandhi in Chinese. In a five-level system, level 3 is usually specified as [M], level 5 and 4 as [H], and level 2 and 1 as [L], as shown in (5) below, because it is rare for a language to have contrastive tones at levels [2] and [1], or two distinctive falling tones such as [52] and [42]. The five level tone pitches can be distinguished by the three features [H], [M] and [L], as shown in (5):

(5)

	5	4	3	2	1
H	+	+	–	–	–
M	–	+	+	+	–
L	–	–	–	+	+

The feature matrix in (5) shows that only [5], [3] and [1] are distinctive from one another for one tonal feature. [4] and [2] both share the two features, [M] and [H] or [L], respectively, which might indicate some variation in phonetic realization in tone pitches. For example, an underlying [13] tone may be phonetically realized as [12] by some people and [23] by others. Based on the data of the Chinese dialects and the analyses of tone features argued for by Halle & Stevens (1971), Maddieson (1974, 1984b), Yip (1980, 1989, 2002), Bao (1991), Chen (2000) and many others, I present, in the light of the feature specifications in (5), the formalization of the feature inventory for commonly-occurring level tones and contours in the five-level scale of most Chinese languages in surface representation, as shown in (6):

[4] In Snider's (1999) tone specifications, [h/l] refers to register feature and [H/L] refers to tone feature. A dot '.' in [h.H] indicates the separation between register and tone.

(6)

2nd tone / 1st tone	5	4	3	2	1
5	H	H	HM	HL	HL
4	H	H	M	HL	HL
3	MH	M	M	M	ML
2	LH	LH	M	L	L
1	LH	LH	LM	L	L

The feature formalization in (6) shows that the tone pitches [5] and [4] are specified with the same feature [H], those of [2] and [1] are specified with [L], and [3] with [M], capturing the insight that in the five-level scale, there are usually three distinctive level tones. However, the model in (6) is only a general pattern. The tone pitch [4] may have specification of either the [H] tone or the [M] tone, and [2] may have specification of either the [M] tone or the [L] tone, according to the tonal system of a language in question in that [4] has the features [H] and [M] and [2] has the features [M] and [L], as shown in (5). The model in (6) also predicts that there can be no phonological distinction between level tones [55], [44] and contour tones [45], [54], etc. It is not common that a language has more than three phonemic level tones without resorting to other phonetic or phonological means, because the difference with only one level in pitch (e.g. between [4] and [5]) is slight. For example, in most African languages, two phonemic levels, H and L, are often sufficient. Mandarin has only one level tone [55] and three contours: [35], [21(4)][5] and [51], specified as [H], [MH], [ML][6] and [HL], respectively. Cantonese has three level tones, differing in pitch: [33]/[3] and [5] in the high register and [22]/[2] in the low register, and nine tones in all, which can be specified as follows:

[5] The third tone [214] in Mandarin is controversial and the last tone target [4] is regarded as having no phonological realization so that the tone is usually treated as low falling (see Yip 2002), or as low level (see Woo 1969), rather than dipping.

[6] Woo (1969) claims that the third tone is basically a level low tone and that the rise which appears at the end of a phrase-final third tone is inserted by a phonological rule. She then proceeds to represent the third tone as a low level tone. In her view, there are two level tones (high and low) in Mandarin.

(7) Specification for tones in Cantonese (Chen 2000: 16/33):

	level		rising	falling
	CVN	CVq[7]		
high (yin)	33[M]	5[Hq]; 3[Mq]	35[MH]	53[HM]
low (yang)	22[L]	2[Lq]	23[LM]	21[ML][8]

The three tone features [H], [M] and [L], as shown in (6), are more suitable for languages which have no more than three phonemic level tones. However, some Asian languages may have four or five contrastive levels, as was mentioned above (see Shi *et al* 1987).

In phonological feature theory, the *yin* and *yang* registers and the five-level tone pitches can also be represented in the feature system, with the representation [H] for the *yin* register and [L] for the *yang* register, and [h] for high-pitch tones and [l] for low-pitch tones (see Yip 1980, 2002; Bao 1999). In this proposal, which I will follow here, register and tone have independent feature specifications. SX has four level tones, two in the high register and the other two in the low register. The three-level-tone system (H, M, L) cannot properly specify the four level tones in SX (which will be discussed in §5.6). In the high register, the tone pitches [5] and [4] are usually specified as [h] and [3], [2] and [1] are all specified as [l], while in the low register [3] is specified as [h] and [2] and [1] with [l]. Throughout my dissertation, I use the features H and L for register and h and l for tone pitches in my analysis of the SX tonal system, following Bao (1999) and Yip (2002).

Phonologically, the tone pitches of [5] and [55] are identically specified: they will both bear the same feature [h], instead of *[hh]. Cross-linguistic evidence shows that *[hh] as a sequence of two tone-bearing units with two tonal features violates the OCP (Goldsmith 1976; McCarthy 1986; Pulleyblank 1986; Snider 1999; Chen 2000; among others). In some other languages, the same applies, as in Snider's (1999: 9) analysis of Mende[9]. For example, [bȅlȅ] 'trouble':

[7] Chen (2000) specifies the three *entering* tones which only occur in the syllables ending in a stop with an extra feature marker [q] in Cantonese.
[8] Chen (2000) specifies [2] with [M] in [21] but [2] with [L] in [23] in one language, which seems questionable.
[9] One of the Niger-Congo languages, spoken in Sierra Leone, Guinea, etc.

(8) Monomorphemic form

Snider (1999) assumes that monomorphemic forms like [bɛ̀lɛ̀] do not have a sequence of two identical tones, so that [bɛ̀lɛ̀] has only one Low tone which spreads over two TBUs, rather than two Low tones, as shown in (8), because of the OCP. In short, tone feature sequences as *[ll], *[hh], *HH and *LL in one monomorphemic form are ruled out by this universal constraint.

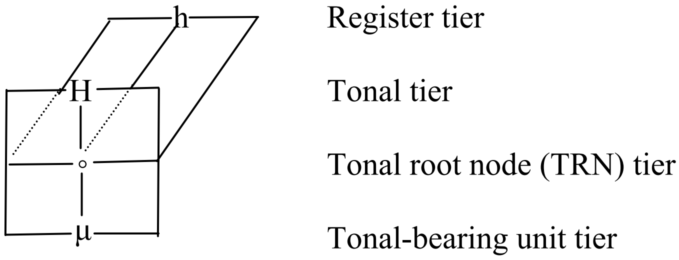
5.4 The Geometry of Tone

The phonological properties of tones in tone languages are best represented in terms of autosegmental features, as various studies have shown. Various distinctive features and geometrical arrangements thereof have been proposed. These competing proposals have been reviewed critically and in considerable detail in Hyman (1986, 1993), Snider (1988, 1999), Bao (1999), Chen (2000), and Yip (1989, 2002). In this subsection, I will briefly introduce some influential proposals of geometrical tone structure. I will claim that feature geometry is universal but internal tone structure can be language-specific, because different tone languages have different types of tones (e.g. falling contour, rising contour or/and level tones), different numbers of tones, and different TBUs, so that tone may behave differently. This is not unlike the account of syllable structure, for which a universal X-bar structure has been postulated but with language-specific differences in internal sub-syllabic constituents, as was discussed in chapter 4.

5.4.1 Snider's proposal

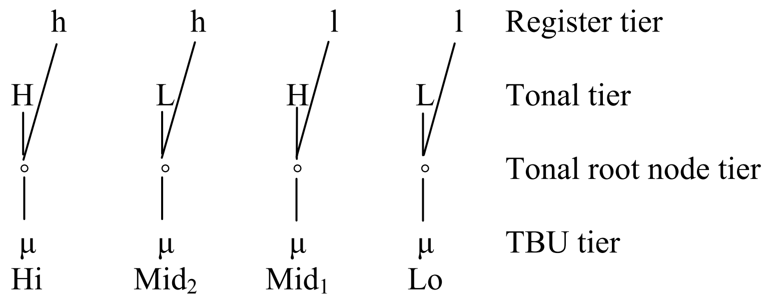
Snider (1999) proposes a register tier theory for his analysis of some African languages. In Snider's view, the register feature associated to any given TBU specifies whether the register of that TBU is higher or lower than the preceding register. The tonal feature associated to any given TBU specifies whether the tone is high or low, relative to the current register (Snider 1999: 25). This is presented in a geometrical structure as in (9):

(9) Geometry of tone (Snider 1999: 23)



Snider's tone geometry in (9) shows that features on the Register tier and the Tonal tier are linked to structural nodes on the TRN tier. The two tiers are like two pages linked together at the binding edge in an open book. On this assumption, the two types of features (H/L for tone features and h/l for register features) are independent of each other. Either can spread or delink on its own. With these two types of features and the geometry structure in (9), Snider (1999:24) postulates four level tone phonemes: Hi, Mid_2, Mid_1, and Lo, as shown in two-dimensional representation below:

(10)



The configurations of the four level tone phonemes presented in (10) may well capture the phonological behaviour of tones in some African languages.[10] However, in Snider's (1999) proposal, the Tonal tier cannot branch so that a contour under one TRN tier is not allowed. This does not fit the tonal structure of SX because in SX a rising contour or a falling contour under one tonal root node is very common, e.g. [35], [13], [52] and [31].

[10] Snider (1999) presents his analysis of Chumburung, Bimoba, Engenni, Mende, Acatlan Mixtec, etc. with the configurations of the four level tones in (10).

5.4.2 Yip's proposal

In Yip's (1980, 1989) theory, mainly based on her analyses of Asian languages, especially the Chinese dialects, a tone is not an indivisible entity in phonological representation. Rather, just as in Snider's proposal, it consists of two parts: Register and Tone. Register indicates the imaginary pitch band in which a tone is realized, and tone specifies the way the tone behaves over the duration of the tone-bearing unit (Bao 1999: 22). These two features, Register and Tone, combine to define four pitch levels phonologically, as was shown in (4). Yip's proposal in (4) shows that the two features play different roles. The Register feature ([±Upper]) first splits the entire pitch range into two halves, each of which is subdivided by the feature [High]. One motivation for the above proposal is that for the vast majority of languages four is the maximum number of contrastive level tones, without necessitating the notion of 'Mid'. Yip's register approach, as shown in (4), prevails over the three-level (H, M, L) systems.

In Yip's theory, the register feature [+Upper] (marked by H) of a high rising tone is the Tonal root node, which dominates tone features (which may branch for contours) (marked by l and h or h and l), as illustrated in the geometrical structure below (Yip 1989, 2002):

(11) σ

H Tonal Node Register

l h Tone

According to Yip's proposal in (11), register dominates tone and tone can be complex (i.e. a contour), but register can never be complex one syllable, one register (this will be discussed later). Based on the tone features in (4), the geometric structure in (11) allows a tone language to have maximally eight tones, including level tones and contours (rises and falls but not concave and convex),[11] which can be formalized as follows:[12]

[11] It is still a matter of controversy whether concave and convex tones exist underlyingly in the languages which are claimed to have them (see Yip 2002).

[12] The following configurations of internal tone structure do not include a toneless syllable which is often assumed to have a neutral tone, or a default tone, i.e. a low level tone.

(12) [+Upper] Register

a.	b.	c.	d.
H	H	H	H
h	l	h l	l h

[−Upper] Register

e.	f.	g.	h.
L	L	L	L
h	l	h l	l h

Yip's [±Upper] approach to Register allows eight distinctive tones, which quite precisely captures the tonal system of SX, whose tones have a clear register division, four in the [+Upper] register and the other four in the [−Upper] register, strictly divided. On Yip's assumption that register features dominate tone features, as shown in (12), the register feature cannot spread independently of tone feature(s); rather, when the register feature spreads, the tone features have to spread along, because the latter is dominated by the former. We will see, however, that this prediction is not always borne out by Chinese dialects, including SX.

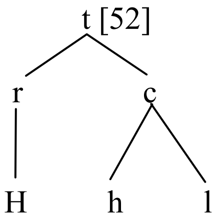
5.4.3 Bao's proposal

Following Yip (1980, 1989) and Halle and Stevens (1971), let us assume that tone allows a register division ([+Upper] and [−Upper]) and is specified by either [+stiff] (H) or [−stiff] (L). Bao (1999) proposes a geometry of tone in which a tonal root node (t) dominates both a register (r) and a contour (c) node. Under the t node, r can dominate either H or L and c can dominate either h or l. If c branches, it may dominate a sequence of lh or hl. The geometry of tone is then as represented in (13):

(13)
t
r c
[α slack] [−α slack] where α = + or −

With the geometry of tone in (13), the internal structure of a contour *yin ping* tone [52] in SX can be illustrated as in (14):

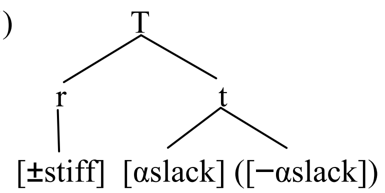
(14)



The tone structure in (14) shows that [52] is a high-register tone and a contour, specified as [H.hl]. Bao's (1999) geometry of tone in (13) shows that r and c are in sister relationship, both dominated by a t node, which allows each feature, [H], [h] and [l], to spread independently. The geometry of tone in (13) captures the fact that the register feature can spread on its own (an example of register feature spreading alone will be presented in the following subsection) while leaving the contour feature(s) behind in SX (contour feature spreading in SX tone sandhi will be discussed later in this chapter). In this sense, Bao's proposal in (13) is more helpful than Yip's proposal in (11) with regard to the SX tonal structure.

However, Bao's geometry of tone in (13) has some conceptual problems. For example, to represent the tone [H.l] in SX it is awkward to call [l] a 'contour' feature since [H.l] is in fact a low level tone. Following Bao's (1999) geometry of tone, I would propose an adaptation of some nodes, as shown in (15):

(15)



where α = + or −; [+stiff] = H; [−stiff] = L; [+slack] = h; [−slack] = l

In (15), 'T', the Tonal Node, which refers to the whole tone, dominates 'r' (the register node), which has the feature H or L, and 't' (tone node), which has the feature h, or l, or a sequence of features lh or hl, so that [l] is a tone feature, separate of the register feature. The features [hl] in (14) are also referred to as tone features. In this chapter, I adopt the geometry in (15), as it is especially suitable for the analysis of tone sandhi in SX.

5.4.4 Register feature spreading

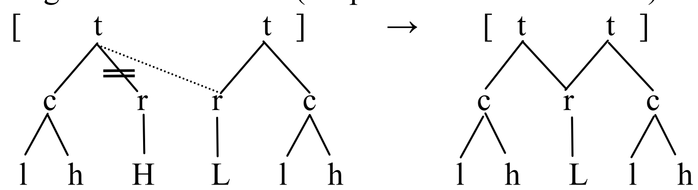
Cross-linguistic evidence shows that tonal register can spread independently of tone features. Bao (1999) presents some examples of regressive anticipatory register feature spreading in Chaozhou,[13] such as a sequence of a syllables with underlying /35.13/ tones, which surfaces as [13.13], as shown in (16) (Bao 1999: 76):[14]

(16)

	Base form		Sandhi form	
	[t^{h}wæ^{35}paŋ^{13}]	→	[t^{h}wæ^{13}paŋ^{13}]	‘quit class’
	[xa^{35}kwei^{13}]	→	[xa^{13}kwei^{13}]	‘start cooking’
	[nja^{35}tɕ^{h}iŋ^{13}]	→	[nja^{13}tɕ^{h}iŋ^{13}]	‘play music instrument’
	[yɛ^{35}səŋ^{13}]	→	[yɛ^{13}səŋ^{13}]	‘the courtyard (is) deep’

The examples in (16) show that only register feature changes (from H to L) in tone sandhi in Chaozhou. Bao formalizes this in a register assimilation rule, as shown in (17):

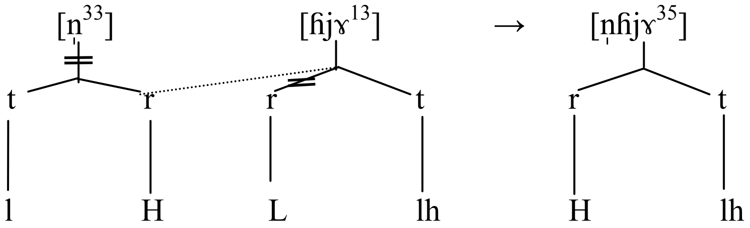
(17) Register Assimilation (adapted from Bao 1999:78):



In (17), the register feature [L] of the second tone spreads regressively to the tonal node and the register feature [H] of the first tone is delinked, changing from [H] to [L], while the contour features remain unchanged, and thus the two tones have the same register feature [L]. Rule (17) says that register spreads independently of the tone features and that the register feature of one syllable can be replaced by the register feature from another syllable, leaving the tone feature(s) unchanged. This also happens in SX, in which a clitic can undergo register feature spreading, when it is merged with a host. In SX, there is a lexical item [n̩jɤ^{35}] ‘haven’t/have no’, which is underlyingly composed of two syllables, [n̩^{33}] ‘not’ and [ɦjɤ^{13}] ‘have’. [n̩^{33}] in SX is a syllabic nasal and forms the lexical monosyllable meaning *not*. The lexical syllable [n̩^{33}] ‘not’ cannot occur alone but al-

[13] A Southern Min dialect known as Teochow in the local vernacular (Bao 1999: 75).
[14] In Bao (1999), the prenuclear glides are transcribed as [i] and [u]. I will use the symbols [j] and [w], respectively, in (16).

ways occurs in combination with [ɦjɤ^{13}] 'have'. The underlying [n̩^{33}ɦjɤ^{13}] surfaces as [n̩jɤ^{35}] in the utterance, which involves a series of phonological changes. First, as a clitic, which cannot be stressed (Katamba 1993; Manfredi 1993; Trask 1996), the unstressed [n̩^{33}] is no longer a TBU (recall the discussion in chapter 2 that an unstressed syllable cannot be a full-tone TBU in SX). Secondly, the tone [33], specified as [H.l], when losing its TBU, spreads its register feature [H] to the tone of the host syllable, replacing the original register feature [L] with [H], changing [13] to [35] in phonetic realization. This series of changes can be illustrated as in (18) (see Zhang 2005: 69-79 for more details):

(18)


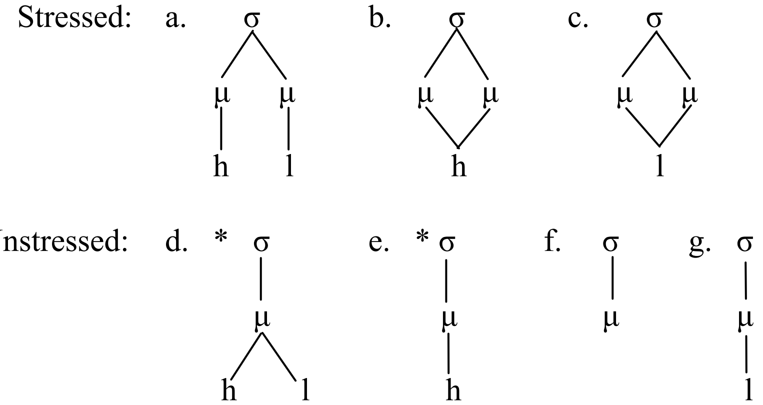
Although there are different approaches to tonal structure in geometry features, it has been agreed that there are minimally two binary features for tone, hierarchically arranged so that all possible combinations exist. One feature splits the pitch range into two registers, and the other feature subdivided each register into two tones (see Yip 2003: 26-35).

5.5 The TBU in SX

Bao (1999) assumes that tone is realized on segments that serve as syllabic nuclei. The canonical tone-bearing segments are vowels, which means vowels are TBUs. Cross-linguistic evidence shows that TBUs can be different from language to language. For example, in Swedish, the entire syllable is the TBU (Gussenhoven & Bruce 1999); in Burmese, the rhyme is the TBU (Zhang 2002); in Kikuyu, the nuclear vowel is the TBU (Clements & Ford 1979); in Luganda, the mora is the TBU (Clements 1986). I assume that in SX the TBU is also the mora, rather than the nuclear vowel.

As was discussed in chapter 4, SX is subject to the Stress-to-Weight principle (Prince & Smolensky 1993), which says that stressed syllables must be heavy and bimoraic and unstressed syllables must be light and monomoraic. In addition, all stressed syllables must bear full tones and unstressed syllables are either toneless or bear a neutral tone. This strongly suggests that moras are TBUs in SX. A syllable with a full tone must be stressed and bimoraic, each mora bearing at most one tone. Thus, one mora bears at most one level tone, but one level tone can be linked to two moras, as shown in (19):

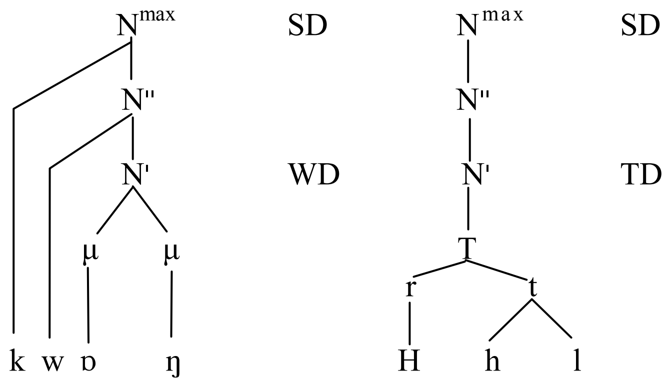
(19)



The configurations in (19) show that a stressed bimoraic syllable can bear a high/low level tone or a contour, as shown in (a, b, c). This suggests that a contour is a sequence of level tones. An unstressed monomoraic syllable cannot bear a contour or a high level tone, although it may bear a low tone as a neutral tone, as shown in (d, e, f). This is also suggested by the tone sandhi phenomena in SX that we will examine below. Thus, an unstressed syllable in SX is either toneless or bears a neutral tone, which is usually a default low tone [l] (Chen 2000).

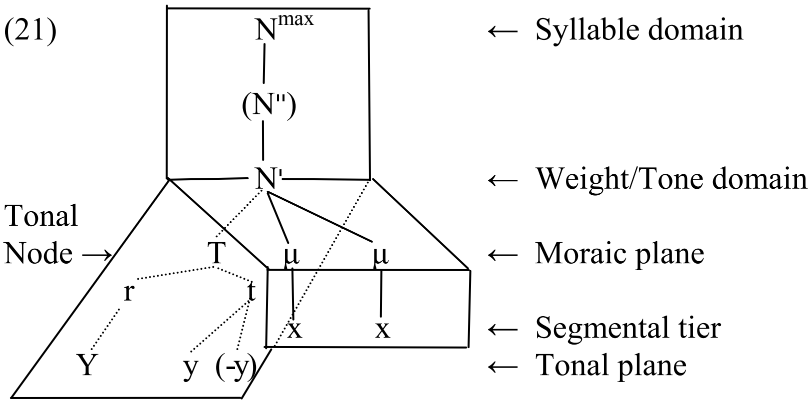
As was discussed in chapter 4, the nucleus and coda are moraic in SX, so that the weight domain is N' in the X-bar structure, as was shown in ((36), ch.4). Thus, N' is also the tone domain in tonal structure. The moraic structure and the tonal structure are two aspects of one and the same syllable structure. For example, the moraic structure and tonal structure of the syllable [kwɒŋ^{52}] 'light' in SX can be illustrated as in (20):

(20) a. Moraic structure b. Tonal structure



In (20), 'SD' refers to the syllable domain. 'WD' refers to the weight domain and 'TD' to the tone domain. The N' domain is both the WD and the TD, which can also be referred to as Rhyme. A three-dimensional structure of a syllable with both mora and tone domains can be presented as follows:

(21)



In (21), 'Y' refers to the register feature (H/L), 'y'[15] refers to the tone feature (h/l), and 'x' refers to a segment slot. The three-dimensional structure in (21) shows that the moraic plane and the tonal plane are like two pages of an open book, joining at the same node, N'. For the moraic plane, N' is the weight domain and for the tonal plane, N' is the tone domain. There is

---

[15] '-y' means the opposite feature of 'y'. If 'y' is [h], '-y' is [l]; if 'y' is [l], '-y' is [h], which means *yy (either *[hh] or *[ll]) under the contour node is unacceptable, since it violates the OCP.

a clear relation between the moraic plane and tonal plane through the N' node, viz. if there is only one mora in the moraic plane, N' is not heavy enough to “support” the tonal plane. The geometric structure in (21) shows the relations between tonal structure and moraic structure when presented in different planes. The geometry of tones and moras in (21) quite precisely captures the tone structure in the syllable domain and its interrelations with moraic structure.

5.6 The Tone Inventory of SX

Cross-linguistic phonetic experiments show that the same tone in a given language may be realized by different tone pitches. A high tone does not have a fixed F_0—it will vary from speaker to speaker, and even for a single speaker, depending on such factors as whether the speaker is male or female, young or old, calm or agitated, whether the word occurs at the start or the end of the utterance, and so on (Chao 1928; Yip 2002). For example, a rising *yin* tone is phonologically specified with [H.lh], in which [H] represents high register, and [lh] the rising contour. Phonetically, however, it may be realized as [25] or [35], differing from language to language, or even from person to person. It is always more difficult to identify distinctive tones than distinctive sounds in a language (Pike 1949).[16] However, the preferred type of lexical tone seems to be (roughly) level underlyingly, while contour tones seem to be added to tonal inventories only in languages with a large number of tonal contrasts (Yip 2002). That is, if a language has only two distinctive tones, for example, it will usually contrast two level tones, rather than a rising and a falling tone. If SX had only four distinctive tones, it would be more likely to have two level tones and two contour tones (one rising, one falling). In this subsection I will present my analysis of how the eight tones in SX are identified and specified for register and tone features.

There have been a number of earlier approaches to the transcription of the eight tones in SX. For example, Yang & Yang’s (2000) transcription can be presented in (22):

[16] Pike (1949) explains that the tonemic analysis is more difficult not because the system is more complicated, but rather because it is a different type of a system (see Pike 1949: ch. 2).

(22) Yang & Yang (2000)

	ping	*shang*	*qu*	*ru*
High Register	a. 42	b. 35	c. 33	d. 4
Low Register	e. 21	f. 13	g. 22	h. 2

The eight tones presented by Yang & Yang in (22) include *ping*, *shang*, *qu* and *ru*, the four tones in each register, which agrees with the general view on the tone types of SX. What differs is the tone pitches. Consider Campbell's (2003) transcription, as shown in (23):

(23) Campbell (2003)

	ping	*shang*	*qu*	*ru*
High Register	a'. 52	b'. 334	c'. 33	d'. 55
Low Register	e'. 31	f'. 113	g'. 22	h'. 23

In (23), Campbell presents a similar set of eight tones in SX but with some different tone pitches. The differences between the exact tone pitches in (22) and (23) are possible because different speakers may produce different pitches for the same tone (cf. above) or different field-workers may have perceived different pitches for the same production. Tones, as suprasegmental units, tend to have variants in phonetic realization. However, there are remarkable differences in tones in SX between speakers of the young generation and those of the old. According to the specifications for tones discussed above, the tones in (22) and (23) can be specified as follows: both 'a' and 'a'' can be specified as [H.hl] because [4] is more likely to be a variant of [5]; 'b' and 'b'' can be specified with [H.lh] because [33] and [3] have the same feature [l], as was discussed in (8); 'c' and 'c'' are the same; 'd' and 'd'' should also be specified with the same feature [H.h], because [4], [5] and [55] can all be specified as [h], according to the feature formalization in (6). The above four tones are high-register tones and the next four are low-register tones, in which 'e' and 'e'' both can be specified as [L.hl] because [2] is more likely to be a variant of [3], as discussed previously; 'f' and 'f'' also have the same specification, because [1] and [11] are both specified with [l]; 'g' and 'g'' are exactly the same; 'h' and 'h'' are very different and it is hard to say which one is correct. Since we already have [hl], [lh], [l] and [h] in the high register and [hl], [lh] and [l] in the low register, we suspect that there should also be a [h] tone in low register, leading to a symmetric system between high and low registers. Thus, we end up with the feature specifications for SX tones as shown in (24).

(24)

	ping	*shang*	*qu*	*ru*
High Register	H.hl	H.lh	H.l	H.h
Low Register	L.hl	L. lh	L.l	L.h

Table (24) presents the nicely symmetrical feature specifications for the eight tones in SX, divided into four tones in the high register and four in the low register. The feature specifications of the eight SX tones are consistent with Yip's proposal for tone features, which was shown in (12). The feature specifications for SX tones in (24) allow us to describe the phonetic tone pitches for the different data.

As was discussed previously, in surface representation, the same tone in a language can be realized in different ways from speaker to speaker. However, as for high level tones in SX, specified as [H.h] in high register and [L.h] in low register, the representation as [23] is obviously not adequate. Maddieson (1978) suggests that the definition of a level tone is 'one for which a level pitch is an acceptable variant'. The high level tones in SX are *ru* tones which only occur in checked syllables ending with the glottal stop and are acoustically shorter in duration than other tones (Bao 1999; Chen 2000; Duanmu 200b). Thus, it is more reasonable to have [5] and [3] in phonetic realization than [55] and [23] (as assumed by Campbell 2003).

Any contour with a two-digit difference between starting and ending points, such as 13 or 53, is probably phonologically a contour, but the ones with only a one digit difference, like 21 or 45, should be approached with a degree of caution (Yip 2002: 23). Based on the comparison between (22) and (23) presented above and on my own observation of the utterances of the tones by the native speakers of SX, I assume that the eight tones of SX are specified as in (24), i.e. [52] for *yin ping* and [31] for *yang ping*, which are falling tones, [35] for *yin shang* and [13] for *yang shang*, which are rising tones, [33] for *yin qu* and [22] for *yang qu*, and [5] for *yin ru* and [3] for *yang ru*, which are all level tones. The tone inventory of SX can be then presented as in (25):

(25)

	Falling	Rising	Level	
	ping	*shang*	*qu*	*ru*
High Register (*yin*)	52	35	33	5
Low Register (*yang*)	31	13	22	3

The tone inventory of SX in (25) shows that there are four different level tones in [5], [33], [3] and [22], and four contours in [52], [35], [31] and [13], which strongly suggests that contours in SX are composed of level tones, viz. [52] by level [5] and [2], [35] by level [3] and [5]. As for [31] and [13], [1] has the same phonological property as [2], both specified with [l]. As contour tones, the phonetic realization is [13] and [31], lowering level [2] to [1] for a distinct contour. Table (25) presents the tone inventory of SX in surface representation, which still allows some difference in pitches from person to person. However, the feature specifications of the eight tones in SX should exclude any other possibility if in the same approach. The tone inventory of SX in (25) also shows that the two level tones of *yin qu* [33] and *yang ru* [3] have the same tone pitch. In phonetic realization, there is not much difference between the high-register [33] and the low-register [3] in terms of pitch level, but phonologically in different registers. However, [3] and [5] are *ru* tones which only occur on checked syllables ending in glottal stop [ʔ]. That is to say, it is predictable from the tone features [H.h] or [L.h] that the syllable must end in a glottal stop, or vice versa. Here the question arises why checked syllables cannot have [l]. Some linguists believe that coda consonants may also affect the tone on the preceding vowels cross-linguistically (e.g. Hombert 1978). Baxter (1992) assumes that in Old Chinese, a final fricative (e.g. [s] or [h]) gave rise to a falling tone; a syllable with no obstruent coda gave rise to a level tone; a syllable ending in [p], [t], or [k] gave rise to a *ru* tone, which is characterized by its shortness.

According to the feature formalization in (6), [3] or [33] is a mid tone, which can be either [H.l] or [L.h], meaning that mid tones can occur either in high register or low register (see Yip 1980). The phonological difference between [H.l] for [33] and [L.h] for [3], as shown in (24) and (25), is in the syllable structure rather than in tone or pitch. Thus, the concept of register is more phonological than phonetic cross-linguistically.

5.7 Consonant-tone Correlation

One of the characteristics of the Wu dialects is that the actual pitch of the upper or *yin* series of tones is higher – *ceteris paribus* – than that of the lower or *yang* series (see Chao 1967; Yan 1994; among many others). This means that the tonal system has a register division between high and low; tones are classified strictly into either register. As discussed above, SX has eight tones; four in the high register and four in the low register, as shown in (25). Their distribution can be exemplified as in (26):

(26) Tone distribution in SX:

[$t^h i^{52}$]	'replace'	[ti^{52}]	'low'	[di^{31}]	'lift up'
[$t^h i^{35}$]	'body'	[ti^{35}]	'bottom'	[di^{13}]	'younger brother'
[$t^h i^{33}$]	'shave'	[ti^{33}]	'weep loudly'	[di^{22}]	'earth'
[t^hɪʔ^5]	'iron'	[tɪʔ^5]	'stumble'	[dɪʔ^3]	'fold'

The fact that high-register tones occur with voiceless initial obstruents and low-register tones occur with voiced initial obstruents gives rise to the question whether the tones of one register are allotones of the other, since the tones of the two registers are in complementary distribution. There are three possibilities. First, low-register tones could be allotones derived from the high-register tones when they occur on syllables with a voiced initial obstruent. *Prima facie* motivation for this could be that high-register tones are more common, since they occur on syllables with both aspirated and unaspirated voiceless initial obstruents, as shown in (26). Second, the voiced obstruents could be allophones derived from voiceless obstruents when they occur with low-register tones, because voicing in obstruents can be predicted from the low-register tones, while voicelessness is predictable but aspiration vs. non-aspiration cannot be predicted on the basis of (high) register, as shown in (26). Third, both voiced obstruents and low-register tones could be underlying forms, though both are in complementary distribution with voiceless obstruents and high-register tones, respectively, because the consonant-tone correlation is determined by phonetic mechanisms, rather than by phonological constraints. In this subsection I will present my analysis of the three possibilities and, in the end, choose for the third possibility.

5.7.1 Allotones?

It has long been known that consonant and tone can interact. Many linguists (e.g. Halle & Stevens 1971; Bradshaw 1979; Duanmu 1990; Bao 1990, 1999) claim that tone is associated to the laryngeal node and that tonal pitch is therefore related to voicing in consonants.

The data in SX, as shown in (26), seem to show that high and low-register tones are in complementary distribution and that low-register tones are triggered by the presence of voiced initial obstruents. This is well expressed by Halle & Stevens' (1971) feature system in which [+stiff] is present in voiceless obstruents and in high tones and [+slack] is used for voiced obstruents and low tones. However, if low tones are allotones of high tones, (i) high tones and low tones should never contrast with each other in any case, and (ii) a change in voicing status should change the tonal register, but not vice versa. The facts in SX, however, provide evidence to the contrary. One piece of evidence is that low tones and high tones contrast in syllables with initial sonorants. Consider the examples in (27):

(27)

a. High register		b. Low register	
[lɛ̃^{52}]	'block'	[lɛ̃^{31}]	'blue'
[lɤ^{52}]	'hollow'	[lɤ^{31}]	'building'
[mɪʔ^{5}]	'wind (*vt.*)'	[mɪʔ^{3}]	'put out'
[mɑɒ^{33}]	'cat'	[mɑɒ^{22}]	'cap'

The examples in (27) show that both high-register tones such as [52], [5] and [33], as in (27a), and low-register tones such as [31], [3] and [22], as in (27b), occur on syllables with the same initial sonorants. This strongly suggests that low tones and high tones contrast with each other and thus should both be present in the underlying representation.

Another piece of evidence comes from the tone merger in cliticization in SX. There is a negation particle, [vəʔ^{3}] 'not', which is very frequent in collocations with many different verbs of SX, as shown in the following examples:

(28) Negator in XP structure in SX

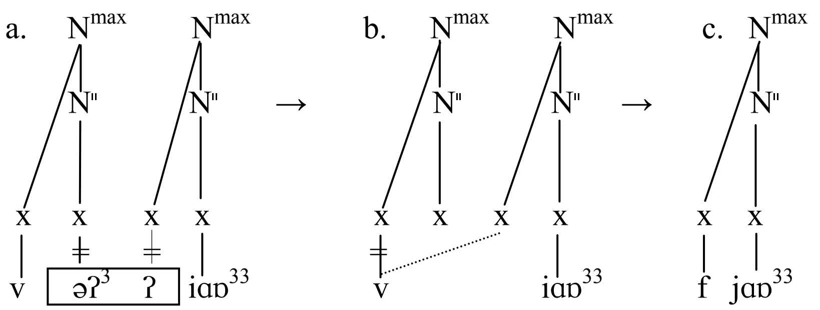
a. [vəʔ^{3}ʔjɑɒ^{33}]	not want	'don't want (to)'
b. [vəʔ^{3}ɦjoŋ^{22}]	not use	'don't use/don't have to'
c. [vəʔ^{3}vɛ̃^{31}]	not naughty	'don't be naughty/well-behaved'

However, the phrase [vəʔ^{3}ʔjɑɒ^{33}] 'don't want (to)' in (28a) always appears in a merged syllable [fjɑɒ^{33}], in which the negator [vəʔ^{3}] 'not', as a clitic, merges phonetically and phonologically into the host syllable [ʔjɑɒ^{33}] 'want', resulting in a new syllable [fjɑɒ^{33}] 'don't want (to)'. In Zhang (2005), I assume that there are some phonological changes from [vəʔ^{3}ʔjɑɒ^{33}] to [fjɑɒ^{33}], as shown in (29):

(29) The phonological process in cliticization

[vəʔ^{3}ʔjɑɒ^{33}] → [v(əʔ^{3})+(ʔ)jɑɒ^{33}] → [vjɑɒ^{33}] → [fjɑɒ^{33}]

In synchronic analysis as in (29), the negator syllable [vəʔ^{3}] has the voiced initial fricative [v] and the low-register tone [3], and the verb syllable [ʔjɑɒ^{33}] has the high-register tone [33] and the voiceless glottal stop [ʔ] as the phonetic ‘filler’ onset. In the process of cliticization, first, the Final of the first syllable and the Initial of the second syllable are deleted; secondly, the remaining Initial of the first syllable becomes onset of the second syllable, merging into [vjɑɒ^{33}]; thirdly, the voiced initial fricative [v] changes into voiceless [f] because of the high-register tone, resulting in a merged new syllable [fjɑɒ^{33}]. This can be formalized in the following rules:

(30)


These phonological rules, as shown in (30), strongly suggest that the low-register tone exists underlyingly, so that it cannot be the allotone of the high-register tone when with voiced initial obstruent; otherwise, the merged syllable should be [vjɑɒ^{22}], instead of [fjɑɒ^{33}]. This phenomenon is similar to another cliticization process, involving [n̩jɤ^{35}], as discussed in chapter 2 (for more details, see Zhang 2005: 69-79).

Diachronically speaking, many scholars (e.g. Xu & Tang 1988; Liu 2002) assume that the original form of the negator syllable [vəʔ^{3}] in the Wu dialects was [fəʔ^{5}] with the voiceless initial [f] and high-register tone [5]. Xu & Tang (1988: 451) assume that [vəʔ^{3}] ‘not’ was pronounced as [fəʔ^{5}] by the old generation of Shanghai speakers, as it is in modern Suzhou. Some other Wu dialects, such as Yuyao (which used to be a county affiliated to Shaoxing City), still have [fəʔ^{5}] for *not*. This phenomenon suggests that the merged syllable [fjɑɒ^{33}] occurred before the voicing of the initial fricative in [fəʔ^{5}] in the old Wu dialects. This assumption is strongly supported by the other merged syllables of cliticization in SX, as shown in (31):

(31)	Phrasal forms		Merged syllables	
	[vəʔ^{3}ɦjoŋ^{22}]	↔	[foŋ^{33}]	'needn't'
	[vəʔ^{3}vɛ̃^{31}]	↔	[fɛ̃^{52}]	'don't be naughty'

In (31), the two syllables of each phrase have voiced initial obstruents and low-register tones; whereas the two merged syllables both have voiceless initial fricative [f] and high-register tones. This phenomenon may suggest that the negator syllable [vəʔ^{3}] used to be [fəʔ^{5}], as assumed by Xu and Tang (1988). However, we have not found any strong evidence for the change from [fəʔ^{5}] to [vəʔ^{3}], viz. whether the voicing of the initial fricative caused the tone to be low-registered or the lowing of the tone caused the voiceless initial fricative to be voiced. However, in another Wu dialect, Longyou,[17] diminutive is realized by changing the tones. For example, [213] changes into [45] and [45] changes into [21] to express diminutive. When the register changes, either from low to high or from high to low, the initial obstruent will also change from voiced to voiceless or from voiceless to voiced correspondingly, as shown in (32) (Cao 2002: 152-160):

(32)	Base tone		Diminutive tone	
a.	[mei^{231}mei^{231}]	→	[mei^{33}mei^{45}]	'younger sister'
b.	[ɕiɑ^{52}kuei^{45}dəɯ^{21}]	→	[ɕiɑ^{33}guei^{21}dəɯ^{213}]	'little boy'
c.	[tsɿ^{45}ȵi^{45}]	→	[dzɿ^{21}ȵi^{45}]	'small earth-worm'

In (32a), [231] changes to [45] and the initial nasal does not change because sonorants can correlate with high-register tones or low-register tones (which will be discussed later in this section). In (32b) and (32c), when [45] changes to [21], the initial [k] changes to [g] and [ts] to [dz], respectively. This phenomenon shows that the voicing status of the initial obstruents is determined by the diminutive tone change.

On the whole, either the synchronic or the diachronic analysis of the merged syllables of cliticization in SX and the tone change for diminutive in Longyou suggest that the low-register tones cannot be the allotones of the high-register tones.

[17] A southern Wu dialect, which has eight tones: [434], [45], [52] and [5] in high register, and [21], [213], [231] and [23] in low register (Cao 2002: 100).

5.7.2 Allophones?

The analysis I presented above gives rise to the question whether the voiced initial obstruents are allophones, since they are in complementary distribution with voiceless initial obstruents and could be predicted on the basis of tone register. However, I argue that the voiced initial obstruents also occur underlyingly since there is evidence in SX that shows that the voiced initial obstruents are not allophones of the voiceless ones. As was mentioned in chapter 2, vowels in high-register syllables are always preceded by a phonetic onset glottal stop [ʔ] when there is no other initial consonant, while vowels in low-register syllables are preceded by a voiced glottal fricative [ɦ] as the onset when there is no other initial consonant, as shown in (33):

(33)	[ʔɛʔ^{5}]	'duck'	[ɦɛʔ^{3}]	'narrow'
	[ʔi^{52}]	'clothes'	[ɦi^{31}]	'move'
	[ʔjaŋ^{33}]	'sprout'	[ɦjaŋ^{22}]	'sheep'

The examples in (33) show that [ʔ] and [ɦ] occur with high and low-register tones, respectively, before a syllable-initial vowel or glide. [ʔ] and [ɦ] are regarded as a pair of phonetic onsets in complementary distribution. Usually, when a pair of phones are in complementary distribution and are allophones in surface representation, (only) one of them must be the underlying phoneme. As for [ʔ] and [ɦ], [ɦ] cannot be the allophone of [ʔ] when occurring on low-register tones, because [ʔ] is not an underlying phoneme, but only a 'filler' onset in phonetic realization. In articulatory terms, [h] and [ɦ] are a pair of fricatives. However, [ɦ] cannot be the allophone of [h] because the former is more frequent and has a wider distribution than the latter, which can be exemplified as in (34):

(34)	voiceless		voiced	
	[ho^{52}]	'shrimp'	[ɦo^{31}]	'river'
	[hɛʔ^{5}]	'blind'	[ɦɛʔ^{3}]	'narrow'
	[hu^{33}]	'call'	[ɦu^{22}]	'unclear'
	[həŋ^{52}]	'groan'	[ɦəŋ^{31}]	'stable'
	*[hi$^{52/35/33}$]		[ɦi^{31}]	'move'
	*[hjɤ$^{52/35/33}$]		[ɦjɤ^{31}]	'oil'
	*[hy$^{35/52/33}$]		[ɦy^{13}]	'rain'

The examples in (34) show that there is a constraint which regulates the distribution of [h], viz. *[h][+high, −back] (see (62), ch. 4), which stipulates that [h] cannot occur before any high front vowel or the front glide, whilst there is no constraint on the distribution of [ɦ]. The different phonological behaviour with respect to the distribution between [h] and [ɦ] strongly suggests that [ɦ] cannot be an allophone of [h].

The fact that full-tone syllables have a phonetic onset ([ʔ] or [ɦ]) when there is no other initial consonant present suggests that an onset consonant is required to assign [+stiff] or [+slack] to the tone on the following vowel. This satisfies the consonant-tone correlation in that tones on the vowels receive [+stiff] for [H] or [+slack] for [L], and not vice versa.

Another piece of evidence comes from tone sandhi in SX, in which the register feature is never involved. I hypothesize that the reason is that the sandhi is sensitive to the voicing status of the initial consonants, since if the register feature were changed, the initial obstruents would have to be changed accordingly, which would lead to a change in segmental structure, potentially resulting in different lexical meaning. We will return to this topic in detail later in this chapter. The analysis I have presented above suggests that voiced initial obstruents cannot be allophones conditioned by low-register tones in SX, but, rather, must be present underlyingly.

5.7.3 Voiced/L in tonogenesis?

Tonogenesis refers to the development of tone, for instance under the influence of neighbouring consonants as a result of language change. There are two theories of tonogenesis, a listener-based theory (Hombert *et al* 1979) and an earlier articulatory-based theory (Halle & Stevens 1971).[18] Hombert *et al.* provide extensive evidence that voiceless consonants raise the F_0 of a vowel and voiced consonants lower the F_0 of the vowel. This (physiological) effect is then exaggerated by the members of the language community so as to mark the difference more clearly for the listener. The effect of various laryngeal configurations of obstruents on the F_0 of a following vowel has been well documented cross-linguistically (also see Mohr 1971; Hombert 1978; Maddieson 1984b; Ohde 1984, cited

[18] In fact, Halle & Stevens' proposal is the basic explanation (physiology) for both theories. The point is that the effect of [voice] on the F_0 is rather small. For tonogenesis, the speakers need to exaggerate the effect, which is normally difficult to perceive. Thus, the listener-based theory is also based on the physiological effect.

from Shryock 1995). Halle & Stevens (1971) offer a different theory of tonogenesis. Their main proposal is that tone and voicing are different realizations of the same articulatory gesture, viz. the stiffness of the vocal cords. Specifically, vocal cord tension is realized in obstruent consonants as devoicing, while in vowels it is realized as tone. They provide interpretations for these relations in terms of traditional phonetic categories for obstruents, glides and vowels through the classification with their proposed laryngeal features (Halle & Stevens 1971 (reprinted in 2002: 51)), as shown in (35):

(35) Classification of obstruents, glides, and vowels in terms of proposed features[19]:

	1	2	3	4	5	6	7	8	9
Obstruents	b_l	b	p	p_k	b^h	p^h	ɓ	ʔb	p
Glides	w,y				ɦ	h,W,Y		ʡ	ʔ, ʔw, ʔy
Vowel	V	V̀	V́	Voiceless vowels	Breathy vowels			Creaky vowels	Glottalized vowels
Spread glottis	–	–	–	+	+	+	–	–	–
Constricted glottis	–	–	–	–	–	–	+	+	+
Stiff vocal cords	–	–	+	–	–	+	–	–	+
Slack vocal cords	–	+	–	–	+	–	–	+	–

Their proposal in (35) shows how the features [stiff] and [slack] vocal cords capture the relationship between low tone and voiced consonants on the one hand, and high tone and voiceless consonants on the other. Halle & Stevens (1971) propose that in the plain vowels, [+stiff vocal cords] is the articulatory correlate of high pitch, whereas [+slack vocal cords] is the articulatory correlate of low pitch. Neutral pitch for the vowels is produced by the configuration [–slack, –stiff]. However, it is a well-documented type of tonogenesis that a relatively lower pitch register develops on vowels following a previously voiced series, and a relatively higher pitch is found after previously voiceless (or voiceless aspirated) series. This process can lead to a multiplication by two of the number of tones.

[19] In (35), there are two interesting obstruents, [b_l] and [p_k]. In Halle & Stevens' (1971) interpretation, [b_l], which probably represents what has sometimes been called a lax voiceless stop, appears in Danish and may occur in initial position for many speakers of English; [p_k] is a moderately aspirated (opposed to fully aspirated) stop of Korean.

Phonetically speaking, vocal fold abduction is a common mechanism in the production of voiceless consonants. The abduction gesture is usually produced in combination with a supralaryngeal constriction, which facilitates the cessation of voicing by decreasing the transglottal airflow (Shryock 1995). The listener-based theory of tonogenesis provides a phonetic mechanism which configures obstruent-tone interaction and the articulatory-based theory explains the phonological motivation for the voiceless-H and voiced-L correlation. However, both assume that the articulation of voicing inherently affects F_0. Following Halle & Stevens's (1971) proposal, I assume that the register is represented by the laryngeal feature, just like syllable-initial consonants: high register is compatible with [+stiff] from [+stiff] initial obstruents and low register is compatible with [+slack] from [+slack] initial obstruents.

Halle & Stevens (1971) also propose that the feature configuration [−stiff, −slack] corresponds to phonetically voiceless stops and phonetically voiced sonorants in that sonorants are spontaneously voiced. Halle (2005, forthcoming) provides a further explanation for the relation between obstruents and sonorants and the vocal folds:

> "Both voicing and pitch are, of course, produced by actions of the vocal folds, but the two classes of sound differ fundamentally with respect to the pressure drop across the folds: the pressure drop is relatively large in sonorants, but significantly smaller in obstruents, and this difference has important consequences for the behaviour of the folds. When slack, the folds vibrate in both obstruents and sonorants. On the other hand, when the folds are stiffened, vocal fold vibration depends on the pressure drop across them. In sonorants, with their large pressure drop, the folds vibrate as before; in fact, the increase in stiffness causes the rate of vibration to increase. By contrast, in obstruents, where the pressure drop across the folds is small, the increased stiffness prevents the folds from being set into motion, and as result the sound is voiceless."

Articulatorily speaking, there is difference between obstruents and sonorants with respect to the voicing property. In obstruents, voicing is active, while in sonorants, voicing is passive. Cross-linguistic evidence strongly supports this asymmetry in voicing behaviour between obstruents and sonorants. For example, in many languages, voicing assimilation is triggered by voiced obstruents, but not by sonorants, e.g. Dutch (van der Torre 2003), Russian (Padgett 2003), etc. As for tonogenesis, Hyman (1978: 266), when writing about West African languages, also notes that among the voiced consonants, it is particularly the voiced obstruents and breathy voiced stops that tend to lower pitch. Thurgood (1996) also finds that in Southeast Asia the lower tone usually occurs after voiced stops, but not after voiced sonorants. The property of passive voicing of sonorants in SX is seen in the correlation with high-register tones and low-register tones, or alternatively for the [−stiff, −slack] feature configuration. Based on Halle & Stevens' (1971) model, the consonants and tones in SX can be specified, using the features [stiff] and [slack], as shown in (36):

(36)

	Consonants			Tones		
	voiceless obstruents	voiced obstruents	sonorants[20]	H	M	L
[stiff]	+	−	−	+	−	−
[slack]	−	+	−	−	−	+

The feature matrix in (36) shows the feature specifications of consonants and tones, from which the following correlation between vocal-cord features and tones holds:

(37) (i) [+stiff] → H
(ii) [+slack] → L
(iii) [−stiff] → L, M
(iv) [−slack] → H, M

These configurations capture the facts of consonant-tone interaction in SX. Tone [3] is a mid tone, which occurs either with voiceless initial obstruents or voiced initial obstruents, viz. either in high or in low register, e.g. [$t^h i^{33}$] 'shave', [ti^{33}] 'weep loudly', and [dɪʔ^3] 'fold'. Moreover, syllables with initial sonorants can have either high-register tones or low-register tones, as was shown in (27). According to the feature matrix in (36), we can also specify H as [+stiff] and L as [+slack] for simplicity. Based on these feature specifications, I propose the following well-formedness conditions for the SX syllable structure in terms of consonant-tone correlation:

[20] As was discussed above, sonorants have default voicing so that the stiffness and slackness of vocal cords do not have an active effect on sonorants, whose [−stiff, −slack] properties allow either high-register tones or low-register tones to occur, at least in SX.

(38) Well-formedness conditions for consonant-tone correlation (WFC(CT)):

(i) T: Every stressed syllable must have a tone.

(ii) *TT: One syllable cannot have two Tonal nodes (no two registers). The following structures are unacceptable: *[HL]$_\sigma$; *[LH]$_\sigma$; *[HH]$_\sigma$; *[LL]$_\sigma$.[21]

(iii) Every full-tone syllable must have an onset to satisfy the consonant-tone correlation (Cf. (26) in ch.2).

(iv) Voiced obstruents have low register and voiceless obstruents have high register, which can be formulated as [+slack]/L or [+stiff]/H, respectively (see also (26), ch.2).

(v) Sonorants have low register or high register, formulated as [son]/L/H.

The WFC(CT) in (38) captures the facts of tone inventory in SX, as shown in (24), which I repeat as follows:

(39)

	ping	*shang*	*qu*	*ru*
High Register	H.hl	H.lh	H.l	H.h
Low Register	L.hl	L. lh	L.l	L.h

The tone inventory in (39) shows that in SX, every tone has only one register and in the feature unary system, SX has maximally eight tones, which is stipulated by the WFC(CT). According to the WFC(CT) in ((38), the configurations in (40) are well-formed:

(40) a. C V^{H} b. C V^{L} c. C V$^{H/L}$

[+stiff] [+slack] [−stiff, −slack]

The configurations above show that a vowel has a H register, sharing [+stiff] with the preceding [+stiff] consonant; a vowel has a L register, sharing [+slack] with the preceding [+slack] consonant; a vowel can have either a H register or a L register, sharing [−stiff] and [−slack] with the preceding [−stiff, −slack] sonorant. This phenomenon suggests that every syllable with a full tone must have an onset consonant to satisfy the

[21] H and L refer to high or low register; *HH and *LL can also be ruled out on account of the OCP.

consonant-tone correlation. Also according to the WFC(CT) in (41), we can tell the following representation are ill-formed in SX:

(41) a.*C V^{L} b.*C V^{H} c. *o $V^{H/L}$ d. * C $V^{H/L}$

[+stiff] [+slack] [+stiff/+slack] [stiff/slack]

(41a) is ill-formed because a [+stiff] consonant is not compatible with a low-register tone; (41b) is ill-formed because a [+slack] consonant is not compatible with a high-register tone; (41c) is ill-formed because a zero onset does not have a laryngeal feature to license [+stiff] or [+slack] of the register on the following vowel; (41d) is also ill-formed because an onset consonant cannot stay unspecified for either [stiff] or [slack] when the following vowel has a high-register or low-register tone. The configurations in (40) and (41) correctly capture the realization of all the SX syllables in terms of consonant-tone correlation.

In short, I assume that in SX every tone has to be specified for [stiff] or [slack], viz. every syllable has a register feature which must keep the agreement of the laryngeal features between the onset consonant and the tone. The perhaps somewhat surprising conclusion is that, in surface representation, every stressed syllable in SX must have an onset. When underlyingly there is no onset consonant, [ʔ] or [ɦ] will be inserted as phonetic onsets of a high-tone syllable or a low-tone syllable, respectively, to satisfy the consonant-tone correlation. This phenomenon can be captured by an OT analysis. Bearing in mind the constraints of WFC(CT) in ((38), we can establish a clear picture of the consonant-tone correlation in SX, based on the constraint ranking ONSET, *TT ≫ DEP-IO. If it is true that [ɦ] and L are [slack] and [ʔ] and H are [stiff], the input /V^{H}/ (underlyingly a high-register tone has no onset consonant) has the surface form in (42):

(42)

/V^{H}/	ONSET	*TT	DEP-IO
a. [V^{H}]	*!		
b. [ɦV^{H}]		*!	*
c. ☞ [ʔV^{H}]			*

In (42), candidate (a) has H register, but has no onset, which violates ONSET and is ruled out; candidate (b) has [ɦ] for [slack] and H for [stiff], so that it violates *TT and is also ruled out; candidate (c) has [ʔ] and H in agreement on [stiff]. Thus, it is the optimal output. The surface form of the input /V^L^/ can also be worked out by the same constraint ranking, as shown in (43):

(43)

/V^L/	ONSET	*TT	DEP-IO
a. [V^L]	*!		
b. [ʔV^L]		*	*
c. ☞ [ɦV^L]			*

In (43), candidate (a) is ruled out because it violates ONSET; candidate (b) has [slack] and [stiff] for [ʔ] and H, respectively, so that it violates *TT and is also ruled out; candidate (c) has a voiced initial obstruent and L register, satisfying WFC(CT), and it is the winner. The tableaux in (42) and (43) show that the surface representation of an underlying vowel with a high-register tone or a low-register tone is [ʔV^H] or [ɦV^L], respectively, as attested by the data in SX, as shown in (44):

(44)				
	[ʔɛʔ^5]	'duck'	[ɦɛʔ^3]	'narrow'
	[ʔi^{52}]	'clothes'	[ɦi^{31}]	'move'
	[ʔa^{52}]	'crowded'	[ɦa^{31}]	'shoes'
	[ʔjɤ^{52}]	'low (voice)'	[ɦjɤ^{31}]	'oil'
	[ʔu^{33}]	'black'	[ɦu^{22}]	'unclear'

The examples in (44) show that the glottal obstruents [ʔ] and [ɦ] occur in the onset position of syllables with high-register tones and low-register tones, respectively, when there is no other onset consonant. In this case, the two phonetic glottal obstruents have no Place component, so that there is no constraint on their distribution in terms of segment sequences, viz. their relations with the following vowels or prenuclear glides. As was discussed in chapter 4, I assumed that the phonetic onset [ʔ] and [ɦ] simply stand for the glottis features [+stiff] and [+slack], respectively, as shown in Halle & Stevens' (1971) proposal in (35), as required by the consonant-tone correlation in SX.

In an utterance, when a syllable with a high tone and the phonetic glottal stop in onset position becomes unstressed in a certain environment, the initial glottal stop automatically drops off, which makes it possible for liaison to take place between a syllable ending in a nasal and a syllable beginning with a vowel, as was discussed in the previous chapter. Accordingly, I would term the phonetic initials [ʔ] and [ɦ] in SX as 'filler' onsets to help realize tones or demarcate a high or low register tone, without any phonological properties in segmental syllable structure. However, the phonetic realization of [ʔ] and [ɦ] helps the tonal system of SX fit in exactly with the general pattern of consonant-tone interaction.

The analysis I have presented above gives rise to the question whether only consonants can affect tone and not vice versa, i.e. whether tone does not affect consonants as assumed by Hyman (1973, 1978). Hombert (1978: 95) also thinks that "it is extremely difficult to find a single case in the literature in which it is clear either from the author's presentation or from our own reanalysis that voiceless consonants, for example, became voiced before a low tone, or voiced consonants became voiceless before a high tone". However, there is an example of tone merger in SX cliticization, in which the change of a tone from low register to high register intrinsically devoices initial voiced obstruents, as was shown in (29). To sum up, I hypothesize that both voiceless and voiced obstruents as well as the high and low register specifications must be present in underlying forms.

5.8 Tone Sandhi

Tone sandhi involves changes in the phonetic realization of a tone under the influence of a neighbouring tone, and sometimes more specifically to the replacement of one toneme by another in such circumstances. In this subsection, I will present my analysis of various forms of tone sandhi in SX, attempting to formalize the phonological rules which regulate the tone sandhi processes in SX. I find that tone sandhi in SX is the result of phonetic dissimilation in tone pitches and phonological assimilation in tone features (other languages may be more restrictive). Moreover, I assume that tone-pitch changes in the surface representation are best captured by phonological rules of tone feature delinking and spreading, observing at all times the prohibition against crossing association lines (following Goldsmith 1976: 27). I will furthermore assume that the phonological motivation for all the sandhi forms, which involve contour dissimilation, contour simplification and contour formation, lies in prosodic requirements regarding the rhythmic structure. The configurations of disyllabic tone sandhi rules in SX can be formalized in the following table:

(45) Configurations of tone sandhi in SX disyllabic sequences T_1+T_2:[22]

T1 \ T2	[hl] falling	[lh] rising	[l] l-level	[h] h-level
[hl] falling	[l(h).hl]	[l.lh]	[l.l]	[l.h]
[lh] rising	[l(h).hl]	[lh.hl]	[l(h).(h)l]	[l(h).h]
[l] l-level	[l.(h)l]	[l.lh]	[l.l]	[l.h]
[h] h-level	[h.hl]	[h.lh]	[h.(h)l]	[h.h]

In the following subsections, I will explain the configurations of tone sandhi in (45) with data in SX and I will also present an OT analysis of all the disyllabic tone sandhi rules in SX, to explain how and why tone sandhi occurs in this language.

5.8.1 Contour dissimilation

Articulatorily speaking, assimilation is often preferred to dissimilation in utterances, because assimilation serves to make sequences of articulatory gestures easier to produce, while dissimilation makes sequences that sound alike more unlike. Why is dissimilation preferred in tone sandhi to assimilation cross-linguistically? I assume that dissimilation in tone sandhi is metrically motivated so as to produce prosodic rhythm.

In many Asian tone languages, there is evidence that two adjacent identical contours are not allowed, especially in Chinese (Yip 1989, 2002; Bao 1999; Chen 2000). The same is true in SX, which has four contours: [52] and [35] in the high register and [31] and [13] in the low register. When two identical rising contours occur in one disyllabic lexical compound or phrasal expression, the tone of the right-hand syllable always changes to a falling contour so that dissimilation takes place, as shown in (46):[23]

[22] Since register features are not involved in tone sandhi in SX, only tone features are cross-tabulated in (45), in which the tones of the first syllable are listed in the first columns and those of the second syllable in the rows. In both columns and rows, ‘l-level’ means low-level; h-level means high-level; the blank ( ) means two alternatives.

[23] All the examples of tone sandhi in SX in this chapter are based on Yang & Yang (2000), with some alternations and additions which were made after consultation of SX native speakers.

(46) a. [35.35] → [35.52] ([H.lh][H.lh] → [H.lh][H.hl])

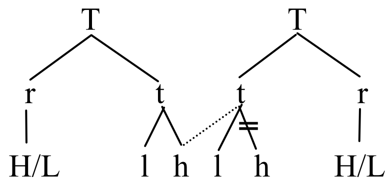
[sɤ^{35}tsɿ^{35}]	→	[sɤ^{35}tsɿ^{52}]	‘fingers’
[ɕjɑɒ^{35}ts^{h}ɑɒ^{35}]	→	[ɕjɑɒ^{35}ts^{h}ɑɒ^{52}]	‘small grass’

b. [13.13] → [13.31] ([L.lh][L.lh] → [L.lh][L.hl])

[dɑɒ^{13}li^{13}]	→	[dɑɒ^{13}li^{31}]	‘reason’
[mo^{13}zɒŋ^{13}]	→	[mo^{13}zɒŋ^{31}]	‘immediate’

The examples in (46) show that when two adjacent rising tones (either in high register or low register) occur in one lexical compound, the second tone of the lexical item changes to a falling contour. The tone change from [35] to [52] and from [13] to [31] is the same change in terms of features ([lh] to [hl]) for both high-register tones and low-register tones, as presented in (45). The rising contour dissimilation in (46) can be formalized as a feature-spreading operation, as shown in (47):

(47)



In (47), the [h] feature of the first tone spreads progressively to the tonal node of the following vowel, which delinks its original [h], making the rising tone a falling tone, and thus dissimilating the contour. According to the feature geometry in (47), contour dissimilation involves the tone features, disregarding register features. Therefore, not only do two identical contours of the same register such as [13.13] or [35.35] dissimilate, but two rising contours in different registers also will undergo dissimilation, e.g. [13.35], as shown in (48):

(48) [13.35] → [13.52] ([L.lh][H.lh] → [L.lh][H.hl])

[lɑɒ^{13}fu^{35}]	→	[lɑɒ^{13}fu^{52}]	‘tiger’
[dɑɒ^{13}ts^{h}ɑɒ^{35}]	→	[dɑɒ^{13}ts^{h}ɑɒ^{52}]	‘straw’

The examples in (48) show that two adjacent contours with identical tone features will be dissimilated. The analysis above shows that contour dissimilation is phonologically realized by tone feature delinking and spreading, which is in fact a phonological process of assimilation between the two adjacent tone features, and phonetically realized by way of different contours in the sandhi forms. In this sense, contour dissimilation in SX is phonetically listener-orientated and phonologically speaker-orientated. In brief, two adjacent identical contours are not allowed in surface representation: contour dissimilation involves dissimilation between adjacent identical tone features. This can be formalized as an OCP constraint against contours, as in (49):

(49) OCP(contour)
Adjacent identical contour features (disregarding register feature) are prohibited in the same phonological phrase.

OCP(contour) in (49) is inviolable in SX, so that a disyllabic word or phrase which violates OCP(contour) must undergo contour dissimilation. This can be formulated in the following rule:

(50) [lh] → [hl] / [lh] __

Rule (50) says that the contour feature [lh] changes into [hl] when following another identical [lh]. However, if two identical falling contours occur in a disyllabic unit, the avoidance of OCP(contour) is realised by contour simplification, in which the first contour becomes a low level tone. I will discuss contour simplification in the following subsection.

5.8.2 Contour simplification

Contour simplification, which is so called because a contour becomes a level tone, is another way to avoid violating OCP(contour). This depends on whether the rule applies lexically or post-lexically. In SX, when the adjacent identical falling contours occur in disyllabic compounds or phrases, contour simplification is applied, as shown in (51):

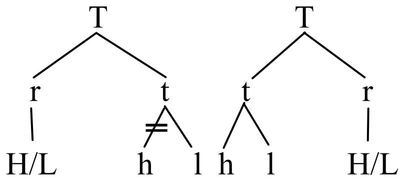
(51) a. [52.52] → [33.52] ([H.hl][H.hl] → [H.l][H.hl])

[ku^{52}ȵjaŋ^{52}]	→	[ku^{33}ȵjaŋ^{52}]	‘girl’
[ɕjaŋ^{52}ʔjẽ^{52}]	→	[ɕjaŋ^{33}ʔjẽ^{52}]	‘cigarettes’

b. [31.31] → [22.31] ([L.hl][L.hl] → [L.l][L.hl])

[dɒŋ^{31}lɒŋ^{31}]	→	[dɒŋ^{22}lɒŋ^{31}]	‘mantis’
[dʑjoŋ^{31}ȵɪŋ^{31}]	→	[dʑjoŋ^{22}ȵɪŋ^{31}]	‘poor people’

In the compound nouns in (51), the first contour feature [hl] changes into a level [l] when it is adjacent to an identical falling contour, thus simplifying one contour. This can be formalized in (52):

(52)


The contour simplification is realized by delinking an [h] feature, as shown in (52), while register is unaffected. Contour simplification is also applied to adjacent identical contour features of different registers. Consider the following examples of disyllabic compound nouns:

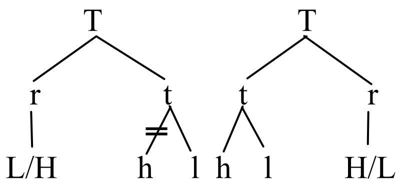
(53) a. [31.52] → [22.52] ([L.hl][H.hl] → [L.l][H.hl])

[dɑɒ^{31}hwo^{52}]	→	[dɑɒ^{22}hwo^{52}]	‘peach blossom’
[dɒŋ^{31}tɕɪŋ^{52}]	→	[dɒŋ^{22}tɕɪŋ^{52}]	‘glucide’

b. [52.31] → [33.31] ([H.hl][L.hl] → [H.l][L.hl])

[ɕjaŋ^{52}ɦjɤ^{31}]	→	[ɕjaŋ^{33}ɦjɤ^{31}]	‘sesame oil’
[kwɒŋ^{52}dɤ^{31}]	→	[kwɒŋ^{33}dɤ^{31}]	‘bald’

The examples in (53) show that in two identical contours in different registers, the first contour is simplified to a low level tone. This is also achieved by way of [h] feature delinking, as shown in (54):

(54)


So far, what is interesting is that two adjacent falling contours result in contour simplification, as shown in (51) and (53), while two adjacent rising contours induce contour assimilation to occur, as shown in (46) and (48). This phenomenon suggests that in SX, when two rising contours constitute the foot, it prefers keeping two (different) contours as its tone type. To distinguish the sandhi forms between two adjacent falling contours and two adjacent rising contours, I propose a constraint of identity tone type for rising contours as in (55):

(55) IDENT-TT(R)
Two adjacent rising contours in the input should also have two contours as the same tone type in the output (no level tone).

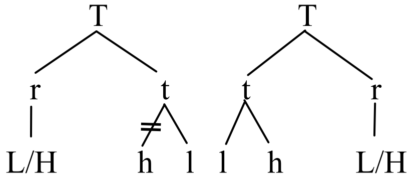
The constraint IDENT-TT(R) in (55) requires that two adjacent rising contours in the input should be still of two contours in the output, but OCP (contour) demands that the two contours assimilate in terms of tone features in sandhi forms. An OT analysis of the different sandhi forms between two falling contours and two rising contours will be presented later in this chapter.

The analysis I have presented so far suggests that when OCP(contour) is violated, the two identical contour features may be dissimilated to two different contours (one rising and the other falling) or be simplified into a sequence of a level tone and a contour to avoid violation of OCP(contour). However, there are examples in which contour simplification may occur even when there is no violation of OCP (contour). For example, in the case of [31]+[13] or [52]+[35], i.e. combinations of a falling and a rising contour, the first contour undergoes contour simplification, as shown in (56):

(56) a. [31.13] → [22.13] ([L.hl][L.lh] → [L.l][L.lh])
[nuə̃^{31}ȵy^{13}] → [nuə̃^{22}ȵy^{13}] 'men and women'
[bɪŋ^{31}dɛ̃^{13}] → [bɪŋ^{22}dɛ̃^{13}] 'ordinary'

b. [52.35] → [33.35] ([H.hl][H.lh] → [H.l][H.lh])
[ts^{h}ɤ^{52}tɤ^{35}] → [ts^{h}ɤ^{33}tɤ^{35}] 'drawer'
[kẽ^{52}ts^{h}ɑɒ^{35}] → [kẽ^{33}ts^{h}ɑɒ^{35}] 'dry straw'

Contour simplification in (56) is also realized by [h] feature delinking of the first contour, as shown in (57):

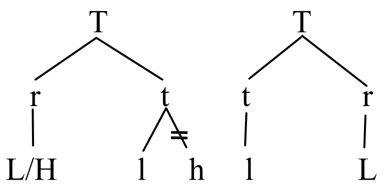
(57)


In the contour simplification process in (57), the [h] feature is delinked from the first contour, although the two contours are not identical. Yip (1989) proposes a partial OCP'' constraint, as in (58):

(58) OCP″
No adjacent partially identical tones (*l.lh, *h.hl, *hl.lh, etc.).

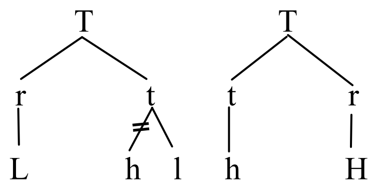
However, contour simplification in SX does not necessarily result from an OCP violation, because the [h] delinking in (57) does not change the [l][l] identical feature adjacency. Sometimes the application of contour simplification will in fact result in what is called a partial OCP violation (Yip 1989; Chen 2000). For example:

(59) [13.22] → [22.22] ([L.lh][L.l] → [L.l][L.l])
[lɑɒ¹³do²²] → [lɑɒ²²do²²] ‘eldest brother/sister’

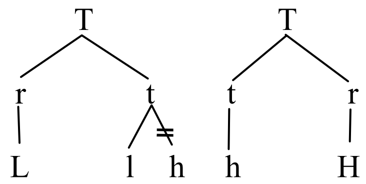


The example in (59) shows that the delinking of the [h] feature has caused a violation of partial OCP″ (58). This suggests that the partial OCP″ constraint does not account for the data regarding tone sandhi in SX. However, there are many other examples of contour simplification in SX, all of which are realized by delinking the [h] feature of the first contour in a disyllabic unit. Here follows a non-exhaustive listing of such examples:

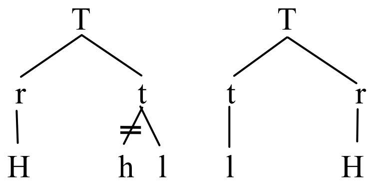
(60) [31][5] → [22][5] ([L.hl][H.h] → [L.l][H.h])
[dʑjaŋ³¹pɪʔ⁵] → [dʑjaŋ²²pɪʔ⁵] ‘wall’
[ɦɥẽ³¹tsoʔ⁵] → [ɦɥẽ²²tsoʔ⁵] ‘round table’



(61) [13.5] → [22.5] ([L.lh][H.h] → [L.l][H.h])
[lɑɒ^{13}tɕ^{h}ɪʔ^{5}] → [lɑɒ^{22}tɕ^{h}ɪʔ^{5}] ‘the seventh child’



(62) [52][33] → [33][33] ([H.hl][H.l] → [H.l][H.l])
[po^{52}laŋ^{33}] → [po^{33}laŋ^{33}] ‘wave’
[fẽ^{52}fu^{33}] → [fẽ^{33}fu^{33}] ‘instruct’



Contour simplification, as shown by the examples from (56) to (62), strongly suggests that the falling contour is not allowed in the left-hand syllable, as is clearly presented in (45). All the forms of contour simplification can be expressed by the following rules:

(63) Contour simplification rules

$$\left.\begin{array}{l}\text{a. [hl]}\\ \text{b. [lh]}\end{array}\right\} \rightarrow \text{[l]} \quad \begin{array}{l}/\ \_\_\ \text{[l]/[h]/[lh]/[hl]}\\ /\ \_\_\ \text{[l]/[h]/[lh]}\end{array}$$

The rules in (63) show that a falling contour and a rising contour will become a low level tone when followed by any tone except in the case of [lh.hl]. This parallels ‘s(trong)-w(eak)’ foot structure in terms of stress. The rules in (63) also show that contour simplification occurs with the first tone or the left syllable (not with the second tone or the right syllable) in a disyllabic compound or phrase and that contour simplification is realized just by delinking of the [h] feature of the first tone. This strongly suggests that in disyllabic compounds or phrases, the [h] feature is not preferred in the left-hand syllable, which indicates that the rhythmic type of feet is iambic in SX. Extensive cross-linguistic research into stress system reveals that the ‘best’ quantitative shapes of disyllabic iambs are (LH)

(Kager 1999: 173). For this cross-linguistic phenomenon, de Lacy (2002: 2) proposes a foot non-head constraint (after Prince & Smolensky 1993):

(64) *NON-HD/H
The foot non-heads do not prefer high tones.

The constraint *NON-HD/H in (64) plays an important role in disyllabic tone sandhi in SX. The metrical structure of the iambic feet will be discussed in detail later (see §5.9). *NON-HD/H stipulates that [h] is not preferred by the foot non-heads, which causes contour simplification to occur if the foot non-head syllable has a contour. However, if the non-head syllable has the *ru* tone which is specified as [h], the [h] feature of the *ru* tone is not deleted or changed into a low tone. The high-level (*ru*) tone always remains the same in tone sandhi because the high-level tone only occurs in checked syllables (ending with a glottal stop). If it changed, syllable structure would also have to be changed, which is not usual in tone sandhi.[24] This phenomenon is captured by the following constraint:

(65) IDENT-*ru*
If the high-level tone [h] occurs in the input, it may also appear in the output and vice versa (ru tone cannot be changed and cannot be formed).

IDENT-*ru* in (65) protects the *ru* tone from changing in any phonological environment. This constraint is inviolable in SX.

5.8.3 Contour formation

In SX tone sandhi, two adjacent identical contours have to be avoided by either contour dissimilation or contour simplification, as was discussed in the previous two subsections. Contour simplification also occurs between two different contours or even between a contour and a level tone. Moreover, a level tone may become a contour in certain phonetic or phonological environments. We will refer to this phenomenon as contour formation. Consider the following examples:

---

[24] Phonetically, *ru* syllables might become mono-moraic when unstressed, e.g. as a prefix, so that the syllable may become toneless or have a neutral tone [l] in surface representation, as was discussed in §5.9.1, in that phonology often creates new objects that are not possible underlyingly.

(66) a. [35.33] → [35.52] ([H.lh][H.l] → [H.lh][H.hl])

[tsɿ^{35}ts^{h}e^{33}]	→	[tsɿ^{35}ts^{h}e^{52}]	'laver (a water plant)'
[sɒŋ^{35}tɕjaŋ^{33}]	→	[sɒŋ^{35}tɕjaŋ^{52}]	'ginger'

b. [13.33] → [13.52] ([L.lh][H.l] → [L.lh][H.hl])

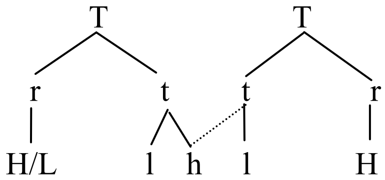
[doŋ^{13}tɕi^{33}]	→	[doŋ^{13}tɕi^{52}]	'intention'
[lɛ̃^{13}hẽ^{33}]	→	[lɛ̃^{13}hẽ^{52}]	'lazy man'

c. [13.22] → [13.31] ([L.lh][L.l] → [L.lh][L.hl])

[me^{13}li^{22}]	→	[me^{13}li^{31}]	'beautiful'
[dɑɒ^{13}lu^{22}]	→	[dɑɒ^{13}lu^{31}]	'road'

The contour formations in (66) all occur on the second tone or the right-hand syllable and are made possible by spreading an [h] feature from the preceding rising contour, as shown in (67):

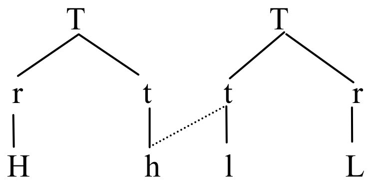
(67)



The high level tone also turns the following low level tone into a falling contour by spreading its [h] feature. For example:

(68) [5][22] → [5][31] ([H.h][L.l] → [H.h][L.hl])

[sɪʔ^{5}ba^{22}]	→	[sɪʔ^{5}ba^{31}]	'fail'
[pɪʔ^{5}ɦwo^{22}]	→	[pɪʔ^{5}ɦwo^{31}]	'mural'



The examples from (66) to (68) show that in contour formation, the [h] feature of the first tone spreads progressively to the t node of the following syllable, changing the level tone into a contour. This can be formulated in a contour formation rule, as shown in (69):

(69) Contour formation rule
[l] → [hl] / [lh/h][_]#

The rule in (69) says that a low level tone will become a falling contour when it is at a right-hand lexical or phrasal boundary and is preceded by a rising contour or a high level tone (i.e. preceded by a high pitch). The contour formation rule in (69) excludes a situation in which [l] is preceded by another [l]: neither tone will become a contour because there is no [h] feature to spread. This suggests that contour formation is realized by [h] feature spreading rather than [h] feature insertion, as shown by the examples in (67) and (68). The rule in (69) also shows that contour formation occurs when there is no [h] feature in the second syllable, which suggests that this feature is in fact required by the second tone or the right-hand syllable in a disyllabic lexical compound or a phrasal structure. For this, I introduce de Lacy's (2002) constraint (after Prince & Smolensky 1993):

(70) *HD/L
The low tone [l] is not preferred by the foot head.

In the previous analysis, I have discussed three forms of tone sandhi in disyllabic lexical compounds and phrasal expressions. The disyllabic tone sandhi data in SX are cross-tabulated in (45) above. However, tone sandhi rules may differ from different lexical conditions. For example, an affix-like syllable can never be the foot head whether it is left-hand or right-hand in spite of the fact that in general, SX is metrically iambic.

5.8.4 Default tone

As was shown in (19g), an unstressed syllable may bear a default tone that is a low [l] tone. A default [l] is different from a stressed [l] phonologically. A default [l] is borne on a monomoraic syllable while a stressed [l] is borne on a bimoraic syllable, as presented in the geometry structures (71a) and (71b), respectively:

(71) a.
σ
|
μ
|
[l]

b.
σ
μ μ
[l]

The default tone is only assigned to unstressed syllables such as affixes, lexically meaningless syllables and some padding devices for word-formation, which may be also toneless. Lexically conditioned, the default tone can be assigned to the left-hand syllable or the right-hand syllable in the stress foot. For example, the prefix [lɑɒ^{22}] in [lɑɒ^{22}.ɦwɒŋ^{31}] 'Mr Wang' is assigned the default tone that is on the left-hand syllable; whereas the padding device [ɦjɤ^{22}] in [tɕjaŋ^{33}.ɦjɤ^{22}] 'soy sauce' is also assigned the default tone that is on the right-hand syllable. Literally speaking, [tɕjaŋ^{33}] means *sauce* and [ɦjɤ^{22}] means *oil*. But [tɕjaŋ^{33}.ɦjɤ^{22}] is not an oil at all. Thus, the right-hand syllable [ɦjɤ^{22}] is not the head lexically, but only a padding device for a disyllabic word. To explain these morphophonological phenomena, I propose a default tone constraint:

(72) DEFAULT-T(R/L)
A default tone [l] is assigned to the right/left-hand syllable which is unstressed for lexical or grammatical reason.

A syllable that is assigned the default tone is absolutely not the foot head. However, the constraint DEFAULT-T(R/L) is different from the constraint *NON-HD/H, because an affix cannot be the foot head, but a non-head syllable is not necessarily an affix. Any affix-like, lexically meaningless, or padding-device syllable must be assigned the default tone [l] unless the syllable underlyingly has the *ru* tone which cannot be changed, as required by IDENT-*ru*. DEFAULT-T(R/L) may also occur on either syllable of [lh.lh] if, for example, one of them is an affix. Therefore, we have such a constraint ranking as IDENT-*ru* ≫ DEFAULT-T(R/L) ≫ IDENT-TT(R) ≫ *NON-HD/H. Analysis in tableaux will be presented later in this chapter.

There are still some other tone sandhi rules concerning disyllabic units, e.g. reduplication, which may require different grammar.

5.8.5 Reduplication

Reduplication is found in a wide range of languages and language groups, though its level of linguistic productivity varies. Reduplication is often described phonologically in one of two different ways: either (1) as reduplicated segments (sequences of consonants/vowels) or (2) as reduplicated prosodic units (syllables or moras). Reduplication in SX is invariably a reduplication of the whole syllable, including the tone. Consider the following examples:

(73) a. [$k^h\tilde{ɐ}^{33}$] 'look' → [$k^h\tilde{ɐ}^{33}k^h\tilde{ɐ}^{33}$] 'have a look'
b. [$kɑɒ^{52}$] 'knock' → [$kɑɒ^{35}kɑɒ^{52}$] 'make a knock'
c. [$dʑjɤ^{13}$] 'uncle' → [$dʑjɤ^{13}dʑjɤ^{31}$] 'uncle'[25]

The reduplication in (73a) has the same tone of the two syllables; in (73b), the tone of the first syllable is changed when in (73c), the tone of the second syllable is changed. I assume that reduplication in SX also copies the same tone from the base, but either tone will change in tone sandhi, according to the metrical structure of the iambic feet. Since reduplication in SX is a full reduplication, it is almost impossible to tell which part is the reduplicant and which is the base. However, reduplication is by its very nature a phenomenon involving phonological *identity* between 'reduplicant' and the 'base'. No matter which part is the base, the optimal reduplication should be the complete AA form ('reduplicant' is the exact copy of the 'base') unless some highly ranked constraints are violated. For example, the reduplication of (73a) best satisfies IDENT-BR (McCarthy & Prince 1995), while those of (73b) and (73c) dissimilate in contours between the two syllables, which is required by OCP(contour), as discussed above. Thus, OCP(contour) dominates IDENT-BR in SX, which is strongly supported by the examples in (73), as shown in the following tableaux:

(74)

Input [$kɑɒ^{52}$]	OCP (contour)	IDENT -BR
a. [$kɑɒ^{52}kɑɒ^{52}$]	*!	
b. ☞ [$kɑɒ^{35}kɑɒ^{52}$]		*

[25] There is no lexical difference between monosyllabic word and disyllabic word. Disyllabification of a minimal lexical word is a tendency of modern Chinese, including SX, as is required by MINWD constraint (Yip 1993; Zhang 2003).

(75)

Input [dʑjɤ^{13}]	OCP (contour)	IDENT -BR
a. [dʑjɤ^{13}dʑjɤ^{13}]	*!	
b. ☞ [dʑjɤ^{13}dʑjɤ^{31}]		*

(76)

Input [kʰẽ^{33}]	OCP (contour)	IDENT -BR
a. ☞ [kʰẽ^{33}kʰẽ^{33}]		
b. [kʰẽ^{33}kʰẽ^{35}]		*!

In tableaux (74) and (75), both candidate (a)s violate OCP(contour), so that they are ruled out, and candidate (b)s are the winner, though they violate IDENT-BR. In tableau (76), the base syllable has a low level tone, so that OCP(contour) is irrelevant and candidate (a) is thus the winner. There are many other examples of the tone sandhi forms of reduplication in SX, as shown below:

(77) a. [31.31] → [13.31] ([L.hl][L.hl] → [L.lh][L.hl])

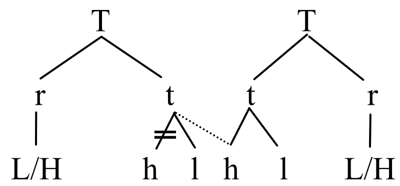
[ɦja^{31}ɦja^{31}]	→	[ɦja^{13}ɦja^{31}]	'grandfather'
[ȵjaŋ^{31}ȵjaŋ^{31}]	→	[ȵjaŋ^{13}ȵjaŋ^{31}]	'grandmother'
[bo^{31}bo^{31}]	→	[bo^{13}bo^{31}]	'grandmother'

b. [52.52] → [35.52] ([H.hl][H.hl] → [H.lh][H.hl])

[koŋ^{52}koŋ^{52}]	→	[koŋ^{35}koŋ^{52}]	'father-in-law'
[ko^{52}ko^{52}]	→	[ko^{35}ko^{52}]	'elder brother'
[la^{52}la^{52}]	→	[la^{35}la^{52}]	'make a pull'

The tone sandhi which occurs in reduplication in (77) is also realized by feature spreading and delinking, as shown in (78):

(78)



The examples in (77) show that when the base syllable has a falling contour, the falling contour of the first syllable in the reduplication form becomes a rising contour, which is different from the sandhi form of the two falling contours other than those involving reduplication where the first falling contour becomes a low level tone, as shown in (51). This phenomenon suggest that IDENT-BR should at least preserve a contour in output, though a different contour, rather than a level tone, if the base has a contour. I would propose a specific IDENT-BR[Contour] constraint:

(79) IDENT-BR[C] (after Kager 1999: 208)
Let α be a tone type in B, and β be a correspondent of α in R.
If α is [γC], then β is [γC].

IDENT-BR[C] allows a different contour in R from that in B as long as it is the same tone type,[26] as shown in (77). The same is true with the reduplication of the base that has a rising contour, as shown in (80):

(80) a. [35.35] → [35.52] ([H.lh][H.lh] → [H.lh][H.hl])

[$\text{tɕi}^{35}\text{tɕi}^{35}$]	→	[$\text{tɕi}^{35}\text{tɕi}^{52}$]	'elder sister'
[$\text{pɑɒ}^{35}\text{pɑɒ}^{35}$]	→	[$\text{pɑɒ}^{35}\text{pɑɒ}^{52}$]	'baby'

b. [13.13] → [13.31] ([L.lh][L.lh] → [L.lh][L.hl])

[$\text{dʑjɤ}^{13}\text{dʑjɤ}^{13}$]	→	[$\text{dʑjɤ}^{13}\text{dʑjɤ}^{31}$]	'uncle'
[$\text{di}^{13}\text{di}^{13}$]	→	[$\text{di}^{13}\text{di}^{31}$]	'younger brother'

The similarity between the reduplication of a falling contour and that of a rising contour is that the right-hand syllable always has a falling contour whether the tone in the base is a falling or rising contour, as shown in (77) and (80). I assume that falling contours attract more stress than rising contours phonetically, which satisfies the metrical structure of the iambic feet. According to IDENT-BR[C], the optimal sandhi form of the reduplication of [35] is [35.52], rather than [35.35] or [33.35]. This can be explained by the following tableau:

(81)

Input [35]	OCP (contour)	IDENT -BR[C]
a. [35.35]	*!	
b. [33.35]		*!
c. ☞ [35.52]		

[26] There are two tone types in SX, viz. contour tone and level tone, so that [γC] indicates either [+contour] or [–contour] (level).

Of the surface reduplication form of [35], there can be another possibility that is [52.35] which also satisfies the constraints in (81). But [52.35] violates the metrical structure of the iambic feet in SX, which will be discussed later in this chapter.

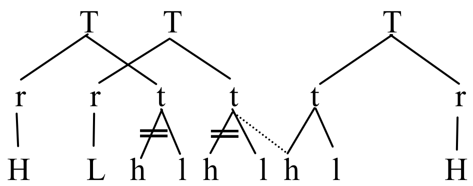
As was discussed previously, IDENT-TT(R) in (55) serves the same aim as IDENT-BR[C] in (79), viz. it preserves two contours as tone type in sandhi forms for the underlying identical contours, which means that two adjacent identical rising contours have the same sandhi rule whether they are in the form of reduplication or non-reduplication. However, we need both constraints, because these two constraints are somewhat different: IDENT-BR[C] is inviolable in SX while IDENT-TT(R) is violable and dominated by DEFAULT-T(R/L). Thus neither constraint can replace the other.

The analyses I presented above all concern disyllabic tone sandhi. However, tone sandhi also occurs in trisyllabic lexical compounds and phrasal expressions in SX, to which we will turn first.

5.8.6 Tone sandhi in trisyllables

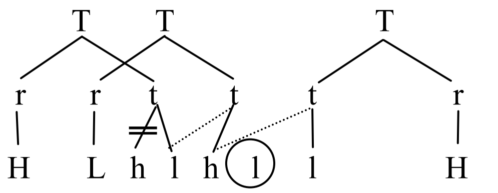
When tone sandhi occurs in disyllabic lexical compounds or phrases, only one tone in the disyllabic unit undergoes some change, either by contour dissimilation, contour simplification or contour formation. These changes are phonologically realized by spreading or delinking the [h] feature and they are usually lexically or syntactically conditioned. While in disyllabic tone sandhi at most one tone changes as a result of sandhi, in trisyllabic compounds or phrasal expressions, one, two or even all three tones may be subject to tone sandhi. The changes in trisyllabic tone sandhi take place in a way similar to feature spreading or delinking in disyllabic tone sandhi constructions. Consider the following examples:

(82) [ɕjaŋ^{52}lu^{31}foŋ^{52}] → [ɕjaŋ^{33}lu^{13}foŋ^{52}] 'Xianglu Summit'
[H.hl][L.hl][H.hl] → [H.l][L.lh][H.hl]



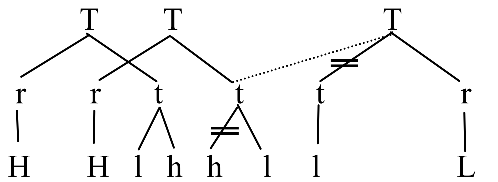
In (82), first, the leftmost contour delinks the [h] feature and becomes a low level tone; secondly, the middle tone delinks its original [h] feature when receiving the [h] feature from the following tone, resulting in the sandhi form: [l.lh.hl].[27]

(83) [ɕi^{52}ɦjaŋ^{31}tɕɪŋ^{33}] → [ɕi^{33}ɦjaŋ^{13}tɕɪŋ^{52}] ‘peep show’
[H.hl][L.hl][H.l] → [H.l][L.lh][H.hl]



In (83), first, the leftmost tone delinks its [h] feature and spreads its [l] to the following tone; secondly, the middle tone delinks its [l] feature when receiving [l] from the preceding tone, changing from a falling contour to a rising contour; finally, the middle rising tone spreads its [h] feature to the rightmost tone, resulting in a falling contour. The sandhi form is [l.lh.hl].

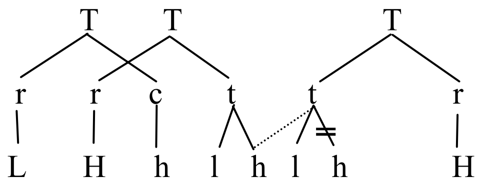
(84) [fu^{35}ts^{h}o^{52}lu^{22}] → [fu^{35}ts^{h}o^{33}lu^{31}] ‘railway’
[H.lh][H.hl][L.l] → [H.lh][H.l][L.hl]



In (84), first, the whole contour of the middle tone spreads to the Tonal node of the following syllable, which delinks its t node and becomes a falling contour; then the middle contour delinks its [h] and becomes a low level tone, resulting in the sandhi form [lh.l.hl].

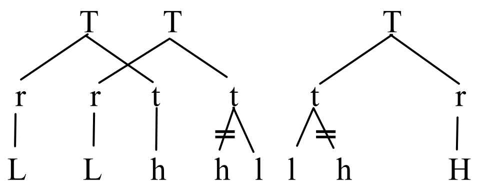
[27] Since the register feature is irrelevant in tone sandhi in SX, it is omitted in the sandhi form, in which a dot ‘.’ indicates a syllable boundary.

(85) [maʔ^{3}kẽ^{35}ts^{h}ɑɒ^{35}] → [maʔ^{5}kẽ^{35}ts^{h}ɑɒ^{52}] ‘wheat stem’
[L.h][H.lh][H.lh]→[L.h][H.lh][H.hl]



In (85), only the rightmost tone delinks its original [h] feature when receiving the [h] feature from the preceding contour, changing from a rising contour to a falling contour. The sandhi form is [h.lh.hl].

(86) [baʔ^{3}bu^{31}tsɑɒ^{35}] → [baʔ^{3}bu^{22}tsɑɒ^{33}] ‘fresh dates’
[L.h][L.hl][H.lh]→[L.h][L.l][H.l]



In (86), only contour simplification occurs in both the middle tone and the rightmost tone, resulting in the sandhi form [h.l.l].

There are many more examples of different forms of tone sandhi in trisyllabic expressions. However, the examples (82)-(86) show that the forms of trisyllabic tone sandhi again involve contour dissimilation, contour simplification and contour formation, just like in disyllabic tone sandhi. In trisyllabic tone sandhi, the leftmost and the middle tone cannot have a falling contour, just like the first tone in disyllabic tone sandhi constructions. In both cases, a falling tone only occurs in the right(most) syllable, which is always the (most) prominent in either disyllabic or trisyllabic compounds or phrases (except when the high-level (*ru*) tone is involved). What is different is that in trisyllabic tone sandhi it is not only the [h] feature that spreads and delinks but also the [l] feature.

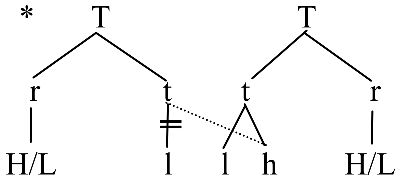
According to the data above and the analysis I have presented, some phonological principles of the processes of the tone sandhi in SX can be summarized as follows:

(87) Principles of tone sandhi in SX:

a. Tone sandhi in SX is phonologically realized by tone feature spreading or/and delinking; no new features are ever inserted.
b. Feature spreading can be progressive or regressive.
c. The sequence of application of feature spreading or delinking in trisyllabic tone sandhi takes place from left to right and from top to bottom.
d. Feature spreading cannot cross any association lines within the sandhi domain.
e. Register features are irrelevant for tone sandhi in SX; only tone features spread or delink.

Based on the principles in (87), some tone changes never occur in tone sandhi in SX such as *[l][lh] → [h][lh] and *[h][hl] → [h][lh] which would lead to crossing association lines (on the assumption that the [l] and [h] tone features are on the same tier), as shown in (88):

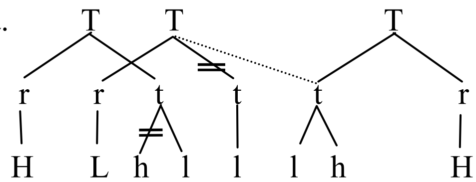
(88) *



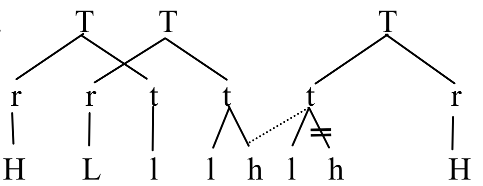
In (88), if the [h] feature spreads regressively to the t node of the preceding tone, it would cross the association line of [l], which is not allowed. Thus, a tone sandhi process such as in (88) is impossible in SX. In some cases, the phonological processes of tone sandhi in trisyllabic phrases are complicated and the feature spreading may take two steps, in a feeding order, making the second step possible. For example, tone sandhi in [huo^{52}lu^{22}sʅ^{35}] 'toilet water' changing into [huo^{33}lu^{13}sʅ^{52}] is realized in two steps, as shown in (89):

(89) [H.hl][L.l][H.lh] → [H.l][L.lh][H.hl]

a.



b.



In the first step in (89a), the [h] feature of the first tone is delinked, making the contour a level tone, while the rising contour of the last syllable spreads to the preceding Tonal node, which delinks its original level tone, resulting in a contour. The first step involves contour simplification and contour formation, producing a [H.l][L.lh][H.lh] tone pattern which violates OCP (contour). In the second step in (89b), contour dissimilation occurs, as a result of spreading the [h] feature of the middle tone to the t node of the following contour, which loses its original [h] feature. These two steps make the tone sandhi process illustrated in (89) possible, changing [hl.l.lh] to [l.lh.hl].

No matter how complicated tone sandhi may be, the five phonological principles summarized above must always be respected. But the questions *why* tone features spread or delink, and *which* feature spreads or delinks still remain unanswered. I will answer these questions and explain the phonological motivation for tone sandhi in SX in the next subsection.

However, we will leave the complicated issues involved in trisyllabic sandhi for future studies.

5.9 The Stress-foot as Sandhi Domain

As was discussed in chapters 2 and 4, 'stress' in tone languages like Chinese has different characteristics from stress in stress languages like English. Stress in Chinese has proven to be frustratingly elusive, both acoustically and perceptually (Chen 2000). When we say that every stressed syllable is bimoraic in SX, 'stress' refers to the phonetic realization of a full tone on a syllable; an unstressed syllable, by implication, is toneless. When we talk about metrical feet, the difference between trochee (sw) and iamb (ws) is made by the specific sequence of strong (s) and weak (w) syllables in the foot domain. Here a stressed syllable is a strong syllable, which can be realized by syllable lengthening, a contour tone or a high-pitched tone. There is cross-linguistic evidence that tone sandhi is intimately related to stress and that metrical prominence attracts the H tone, regardless of what this H is originally associated with (see Odden

1988, 1995). In this sense, Yip (1980: 57, 84) maintains that tone determines stress, rather than vice versa. Many authors (e.g. Shih 1986; Chang 1992; Chan 1995; Duanmu 1992, 1995; Chen 2000; de Lacy 2002) claim that tonal stability and the domain of tone association are related to the metrical structure. In this subsection, I will present my analysis of metrical prominence and tones in the foot domain. I assume that the stress foot is the tone sandhi domain in SX.

5.9.1 Right-prominence in SX

Although left-prominence has been commonly assumed for Shanghai, which is one of the northern Wu dialects (Yip 1980; Wright 1983; Duanmu 1993, 1994), words and phrases are basically right-prominent in Mandarin, Min and some southern Wu dialects (cf. Chao 1968; Hashimoto 1987). SX, one of the southern Wu dialects, is also right-prominent. Following Chen (2000: 238), who argues that a foot must be at least disyllabic, I assume that the stress foot in SX is iambic. There is independent evidence that SX is a right-prominent language.

Right prominence is iambic and has the right-headed foot structure. De Lacy (2002: 3) assumes that 'the foot's head is the head mora of the head syllable of the foot'. Head is always a stressed syllable, which is realized as long vowels or/and high tones. A weak (unstressed) syllable cannot be the head of the foot. There is independent evidence that in disyllabic phonological units, the first syllable is always unstressed, resulting in not only low tone but also shorter duration in terms of moras. For example, [lɑɒ^{13}] 'old' becomes [lɑɒ^{22}] in [lɑɒ^{22}sɿ^{33}] 'the fourth child', in which [lɑɒ^{22}] is a prefix that cannot be stressed. Underlying [lɑɒ^{13}] which has a contour ([lh]) must be bimoraic, while the prefix [lɑɒ^{22}], which has a low-level tone ([l]), can be monomoraic in surface representation, as presented in (90):

```
(90)  stressed: σ        →      unstressed: σ
               / \                          |
              μ   μ                         μ
              |   |                         |
              l   h                         l
```

The moraic structure in (90) shows that the non-head syllable is phonologically shorter. Sometimes, the non-head syllable may also change a long vowel into a short vowel. For example, the prefix [lɑɒ^{22}]

could also become a toneless [lɒ] in [lɒ.sɿ^{33}] 'the fourth child', which is phonetically and phonologically different from [lɑɒ^{13}sɿ^{52}] 'teacher'. That may be the reason, I assume, that Yang & Yang (2000) believe that the single [–tense] [ɒ] is possible in an open syllable. But [ɒ] is possible in an open syllable only when it is unstressed. However, such changes from bimoraic syllables to monomoraic syllables or from long vowels to short vowels only occur to the foot non-head syllables which are usually left-hand in disyllabic phonological units, because in general, SX is a right-prominent language, except for some special cases where the foot is trochaic if the second syllable is unstressed for some lexical reason, e.g. as a suffix.

There is another piece of cross-linguistic evidence that foot non-heads prefer lower tone and heads prefer high tone (de Lacy 2002). This is also true in SX. As was discussed previously, the tone feature [h] is not preferred in the left-hand syllable but preferred in the right-hand syllable. This is captured by the constraints *NON-HD/H in (64) and *HD/L in (70). Even if the tone of a foot non-head, the left-hand syllable, has the feature [h], the foot head, the rightmost syllable, always has a higher-pitched tone which attracts more stress or more prominence. There are more examples in different disyllabic units, as shown in (91):

(91)	a. Noun reduplication:			
	[ɦja.ɦja]	13.31	w.s[28]	'grandfather'
	[tɕi.tɕi]	35.52	w.s	'elder sister'
	b. Verb reduplication:			
	[k^{h}ɑɒ.k^{h}ɑɒ]	35.52	w.s	'give a knock'
	[tsɤ.tsɤ]	33.35	w.s	'take a walk'
	c. Compounds:			
	[ku.ȵjaŋ]	33.52	w.s	'girl'
	[sɤ.tsɿ]	35.52	w.s	'finger'
	d. Phrases:			
	[ɦjɑɒ.zẽ]	22.31	w.s	'row a boat'
	[ma.mi]	22.13	w.s	'buy rice'

[28] A dot '.' between two tones, or between 'w' and 'S', or between tone features, as shown in (88) below, indicates a syllable boundary.

The examples in (91) show that SX disyllabic lexical compounds and phrases are right-prominent and that strong syllables (compared with the metrically weaker ones) have the [h] tone feature. The most striking evidence comes from noun reduplication. In many Chinese dialects, the reduplicated version of the second syllable in noun reduplication is always unstressed and toneless, e.g. [tɕjɛ^{35}tɕjɛ^{0}] 'elder sister' in Mandarin, which has zero tone for the copied second syllable. In SX, however, the second syllable in reduplication of [tɕi^{35}tɕi^{52}] 'elder sister' has a falling tone, or a high level tone, as in [soʔ^{5}soʔ^{5}] 'uncle'. All these phenomena and the analysis I presented above strongly suggest that SX has a robust right-prominent metrical structure in the foot domain. In principle, the foot is binary, which fits in with the binary minimal word formation in modern Chinese (see Yip 1993; Zhang 2003) and the fact that 85% of all nouns in a survey of 3,000 high-frequency expressions are disyllabic or longer (Chen 2000: 366). In SX the situation is similar. In order to meet the disyllabic requirement, various "padding" devices are employed, including otherwise pleonastic expressions (repetition), e.g. each of the following words [me.li] 'beautiful', [ŋɛ̃.tɕɪŋ] 'eye', [tɕjaŋ.pɪʔ] 'wall', and countless others. In addition, noun and verb reduplications are very frequent. Thus, a minimal lexical word is usually a disyllabic foot.

Both in disyllabic and trisyllabic constructions the rightmost syllable always has a higher-pitched tone. In this way, stress demarcates a strong syllable in the foot domain, as was seen in examples (82)-(86). In trisyllabic or polysyllabic constructions, the stress-foot domain is largely decided by the lexical unit, within which rightmost prominence is defined lexically or syntactically. Thus, the middle tone in a trisyllabic construction may have different sandhi forms, depending on lexical or syntactic information. Consider the following examples:

(92)	base tones		sandhi forms
a.	[(tsoŋ^{52}nɵ̃^{31})he^{35}]	→	[(tsoŋ^{33}nɵ̃^{13})he^{52}]
	middle south sea		[(l.lh).hl]
	'Middle South Sea'		
b.	[fi^{52}(nɵ̃^{31}he^{35})]	→	[fi^{33}(nɵ̃^{22}he^{35})]
	fly south sea		[l.(l.lh)]
	'fly (over) South Sea'		
	(…) = disyllabic foot		
	[…] = compound or phrase		

The examples in (92) show that the tone of the syllable [nə̃^{31}] has different sandhi forms: [22] ([l]) as the first tone in [nə̃^{22}he^{35}], in which the prominent syllable is [he^{35}]; or [13] ([lh]) as the second tone in [tsoŋ^{33}nə̃^{13}], in which [nə̃^{13}] is the secondary prominent syllable, retaining [h] to enforce the right-prominent foot domain, although the most prominent is always the rightmost syllable in the phonological phrase. To account for this, Chen (2000: 366), following Shih (1986, 1997), proposes a Minimal rhythmic unit constraint (MRU) for Mandarin T3 sandhi (TS),[29] given in (93):

(93) Minimal rhythmic units (MRU)
Connected speech is broken up into "Minimal rhythmic units" and MRUs are binary.

As was mentioned above, minimal lexical words are mostly binary. As a result, the MRU is at least disyllabic and may also be ideally disyllabic as well (Chen 2000). However, MRUs are mostly constructed morphophonologically, as shown in (92). The MRU plays an important role in trisyllabic and polysyllabic tone sandhi in SX.

5.9.2 The role of [h] in the foot domain

Cross-linguistically, prosodically prominent syllables attract the H tone if there is one (see Odden 1988; Hyman 1989; Chen 2000; de Lacy 2002). On the assumption that SX is right-prominent in its metrical unit,[30] the rightmost syllable in the stress foot always has a higher-pitched tone than the leftmost one, so that, as it is, [h] is always required by the second tone and it is not preferred by the first tone as the constraints of *HD/L in (70) and *NON-HD/H in (64) stipulate. These two constraints respect the right-prominent metrical structure, except for some special cases where the right-hand syllable is a suffix or a lexical repetition of the first syllable which is unstressed, e.g. [ŋe^{22}tsɪʔ][31] 'idiot'. As discussed above, the [h] feature plays a crucial role in disyllabic tone sandhi. In the foot domain,

---

[29] T3 sandhi (TS) is well attested in Mandarin, in which [214.214] invariably changes to [35.214].

[30] There are some special exceptions when the second syllables are unstressed as a suffix or some other lexical reason, so that the foot may be left-prominent, which will be discussed later. However, the general metrical pattern of SX is iambic.

[31] In [ŋe^{22}tsɪʔ], [tsɪʔ] acts as a suffix to form a disyllabic word, lexically meaningless.

the [h] feature also plays an important role, because [h] assigns high-pitched tones, which always attract syllable stress. This is also strongly supported by [h] feature delinking in contour simplification on the first tone and [h] feature spreading in contour formation on the second tone, as discussed above, so as to satisfy the right-prominent foot structure in disyllabic tone sandhi.

However, the most interesting aspect of tone sandhi in SX is that whatever changes, phonetic dissimilation always occurs in tone pitches and phonological assimilation in tone features, as was exemplified in (47), (68), and (78). By the right-prominent metrical principle, the right syllable is metrically strong and the left syllable is metrically weak, which is only indicated by tone features, disregarding the surface tone pitches in SX. For example, [he^{35}li^{31}] 'sea mile' has higher pitch on the first tone than on the second tone; yet it is the proper sandhi form because its tone feature [lh.hl] satisfies the metrical structure. Whereas the underlying [lẽ^{13}hẽ^{33}] 'lazy man' has 'increase-right-forward' tone pitches, its tone feature sequence is [lh.l], which is not acceptable in the sandhi domain, so that the correct sandhi form is [lẽ^{13}hẽ^{31}] with the feature [lh.hl]. Thus, it is the tone features, the arrangement of which constitutes four tones: falling ([hl] *ping*), rising ([lh] *shang*), low level ([l] *qu*) and high level ([h] *ru* for checked syllables), that decide the syllable stress. Cross-linguistically, there is a special relationship between tone and metrically prominent positions: metrically prominent positions can attract high tone (Goldsmith 1987; Downing 1990; Bickmore 1995; de Lacy 1999, 2002; and many others), and high-toned moras can attract metrical prominence (de Lacy 1999). De Lacy (1999, 2002) proposes the three-degree tonal prominence scale in (94):

(94) Tonal Prominence Scale: H > M > L

In (94), H, M, and L stand for high, mid, and low tone, respectively. Only three degrees of height are shown for the sake of brevity. To be more precise, the Tonal prominence scale states that the higher tone is more prominent than lower tone (de Lacy 1999: 5).

Tones in SX are not classified or specified for H, M, or L, but [H/L] for registers and [h/l] for tones. If it is true that a contour is a sequence of two level tones, the first tone is always longer and more prominent than the second one of a contour,[32] just like a diphthong, of which the first element is always longer or more prominent. Acoustically, a falling contour ([hl]) usually gives more stress to [h] while a rising contour ([lh]) gives more stress to [l]. Thus, falling contours can attract more metrical prominence than a rising contour. A *ru* tone is a high tone, but it only occurs on a checked syllable, which is characterized with shorter sound than a contour. According to the sandhi data and phonetic properties of tones, I propose a different three-degree tonal prominence scale of SX:

(95) Tonal Prominence Scale in SX:[33]

Tones	Prominence
[hl]	S
[lh]/[h]	s
[l]	w

The tonal prominence scale in (95) shows that syllable prominence reduces from [hl] to [lh]/[h] to [l]. The falling tone attracts the strongest metrical prominence (marked S); the low level tone attracts the weakest prominence (marked w); the rising tone or the high-level tone is intermediate level (marked s). Prosodically speaking, if the right syllable has an [lh] tone, the left syllable will not have an [hl] tone in the right-prominent foot domain, while the reverse is true for the left-prominent foot. As it is, [hl] is never allowed on the left-hand syllable in a foot domain of SX. Accordingly, in SX there can be such sandhi forms as [35.52] and [22.13], but not *[52.35] and *[13.22], which are not well-formed feet. Based on the tonal prominence scale in (95), I propose that there is a well-formedness condition on the foot as follows:

(96) Well-formedness condition on feet (WFC(F))
The prominence ranking may not decrease from left to right in the foot domain (two equally strong syllables are well-formed).

In SX, WFC(F) in (96) is always respected in all possible tone sandhi processes. The constraints *NON-HD/H in (64), *HD/L in (70), and

[32] e.g. the third tone in Mandarin is [214], but the last [4] is too short to have any phonological property (Woo 1969; Yip 2002).
[33] Since in tone sandhi in SX only tone features are involved, the stress degree is ranked only according to tone features, disregarding register features and tone pitches.

WFC(F) in (96) play a similar role to ALLFTR[34] (McCarthy & Prince 1993; Kager 1999), which requires feet to be aligned as much as possible to the right edge of the prosodic word in right-prominent languages like SX. These foot-form constraints share the universal principle of metrical structure in which the directionality of foot parsing is always asymmetrical. Based on the examples of tone sandhi discussed in the previous subsection, the disyllabic tone sandhi rules in SX are formalized by way of the configurations in (45), which are based on a right-prominent metrical structure.

The table of sandhi rule configurations in (45) shows all the possible disyllabic sandhi forms of the 16 tone combinations in SX, each including the possibilities of two high-register tones, two low-register tones or two different-register (H-L or L-H) tones. It is very interesting that register features have no effect on tone sandhi in SX, so that two-tone combinations, whether they are in the high register, low register or different registers, follow the same tone-sandhi rules. The configurations in (45) respect the stress ranking of tone features in (95) (except for the special sandhi tone [h.l] in high-level+low-level (on a prefix syllable) combinations, which was discussed previously and will be discussed more later). They also show such a foot pattern that is accentual. With the sole exception of [h.lh] which is special because the *ru* tone [h] cannot change, although it is unstressed, no disyllable has more than one [h] in spite of the fact that [h] may be shared by two syllables, which is realized by the [h] feature spreading and delinking, as presented in (47), (68), and (78). Through the analyses I presented above, the following generalizations can be made:

(97) Generalizations regarding disyllabic tone sandhi in SX:
  a. When there is [h] on the first tone, there must be [h] on the second (except [h.l]);
  b. When the second tone is [l], the first must also be [l] (except [h.l]);
  c. A falling contour on the prominent syllable always remains unchanged (except when it is assigned a default tone);
  d. No new identical contours can result from tone sandhi;
  e. A high-level (*ru*) tone always remains the same in tone sandhi;
  f. When a low-level tone is the first tone in the foot domain, the base tones always remains unchanged in the sandhi forms;

[34] ALLFTR: Align (Ft, Right, PrWd, Right): Every foot stands at the right edge of the PrWd. (McCarthy & Prince 1993; Kager 1999).

g. No falling contour can occur on the left-hand syllable.

The generalization in (f) says that when the first tone is the low-level tone in the foot domain, the second tone always remains the same in tone sandhi. The low-level tone has just the [l] feature. When the first (or left) syllable is [l], any tone is acceptable for the second (or right-hand) tone, because the combination of any tone with the first [l] satisfies WFC(F). The optimal output in OT always requires as little violation as possible, which is captured by the constraint of IDENT-IO(T):

(98) IDENT-IO(T)
A tone in the input should have the same value in the output ("no change of tone").

The constraint in (98) is supported by the tone sandhi process in [l]+T[35] combinations, as shown below:

(99)	[po^{33}dʑɤ^{33}]	→	[po^{33}dʑɤ^{33}]	[l.l] →[l.l]	'broken'
	[ma^{22}mi^{13}]	→	[ma^{22}mi^{13}]	[l.lh]→[l.lh]	'buy rice'
	[k^{h}wa^{33}loʔ^{5}]	→	[k^{h}wa^{33}loʔ^{5}]	[l.h] →[l.h]	'happy'
	[k^{h}ẽ^{33}ɕy^{52}]	→	[k^{h}ẽ^{33}ɕy^{52}]	[l.hl]→[l.hl]	'read books'

The tones in the disyllabic words or phrases in (99) do not change in sandhi forms, because they satisfy WFC(F) and are obviously preferred by IDENT-IO(T). The examples in (99) also show that the two adjacent identical level tones ([l.l] as well as [h.h], as shown in (45)) are acceptable, while two adjacent identical contours are never allowed in surface representation. I assume that the reason for the possibility of [l.l] and [h.h] is that tone sandhi in SX is phonologically realized by tone feature spreading or delinking and no new tone feature insertion is allowed in tone sandhi in SX so that [l.l] or [h.h] can never change, in addition to the constraint IDENT-*ru*.

Maddieson (1978: 341) observes that tones associated with stressed syllables have a special status (and could be regarded as 'heads') and play a dominant role in assimilatory processes, in the sense that tone under stress generally remains unchanged, while tones in metrically weak positions tend to assimilate to a more prominent tone. In right-prominent feet, the left-hand syllable is metrically weak and the right-hand syllable is

[35] 'T' refers to any tone in SX.

strong(er). Maddieson's observation also holds true for tone sandhi in SX, which does not affect the high tone of the prominent syllable, as shown in (99), while the high tone of the metrically weak syllable is always changed to a low tone if it can, as shown in (45). When [lh] or [hl] occur on the second (prominent) syllable, it always remains unchanged, unless it violates OCP(contour); when [lh] or [hl] occur on the first (metrically weak) syllable, it is always changed either to a low level tone or it spreads its [h] to the second to satisfy WFC(F). Both the configurations in (45) and the generalization in (97) are evidence that SX has a right-prominent metrical structure and that the stress foot is the tone-sandhi domain. All the changes in disyllabic tone sandhi, including contour dissimilation, contour simplification and contour formation, are sensitive to the right-prominent metrical structure in the foot domain. It is clear that tone sandhi in SX is a matter of [h] spreading or delinking, and plays a role exactly like a stressed syllable.

It has been observed that tone sandhi often takes place between not just any two words, but only between words that are in the same domain. The domain is usually taken to be prosodic, for example, involving a phonological phrase. This prosodic phrasing may be at least partly syntactically determined, so that the syntax indirectly conditions the tonal rule (Yip 2002). Thus, the same tones in different foot positions can have different sandhi forms, as shown in (92), and even when they occur in the same foot position but have a different lexical or syntactic structure, they may also have different sandhi forms. Consider the following examples:

(100)	Phrasal constructions (PC)		Lexical compounds (LC)	
a.	kẽ.k^{h}wɒŋ	'look at the light'	ljaŋ.k^{h}wɒŋ	'bright light'
	[l.hl]	base tone	[l.hl]	base tone
	[l.hl]	sandhi form	[l.l]	sandhi form
b.	fu.dɤ	'wash head'	la.dɤ	'favused head'[36]
	[l.hl]	base tone	[l.hl]	base tone
	[l.hl]	sandhi form	[l.l]	sandhi form
c.	lɤ.ɦjɤ	'leak oil'	tɕjaŋ.ɦjɤ	'soy'
	[l.hl]	base tone	[l.hl]	base tone
	[l.hl]	sandhi form	[l.l]	sandhi form

[36] Favus: contagious fungal infection of the scalp; occurs mainly in Africa and the Middle East.

In (100), the base tones of the second syllable (in the disyllabic expressions) are the same no matter whether they occur in PC or in LC, but they have different sandhi outputs, which are therefore lexically or syntactically conditioned. In the VP structure (PC), the object (the second syllable) always has a strong stress (viz. it can never be assigned a default tone). If, on the other hand, the disyllabic expression is a lexical compound (LC), and the two lexical syllables are synonyms, the second syllable may be assigned a default tone and do *not* have strong stress. In this case, the second syllable is not the foot head, so that *NON-HD/H will rule out [l.hl] in tone sandhi, as shown in the lexical compounds in (100), although in most cases the stress foot is iambic and [l.hl] is always the best sandhi form even for lexical compounds, as stated in (97f). For example:

(101)	ze.zẽ	'god of wealth'	ȵɪŋ.ko	'family'
	[l.hl]	base tone	[l.hl]	base tone
	[l.hl]	sandhi form	[l.hl]	sandhi form

*NON-HD/H plays an important rule in tone sandhi in SX. The difference between lexical combinations in (100) and those in (101) is the head status whether it occurs on the left-hand syllable or the right-hand syllable. For example, affixes cannot be stressed and cannot be the foot head cross-linguistically (Spencer 1991; Katamba 1993; Trask 1996; among others). Neither prefixes (left-hand syllable) nor suffixes (right-hand syllable) can be the foot head, as exemplified in (102):

(102) a.	prefix		b.	suffix	
	[lɑɒ^{22}sɛ̃^{52}]	'the third child'		[tɕjɑɒ^{35}tsɿ^{33}]	'dumpling'

In (102a), the left-hand syllable is a prefix and unstressed, and in (102b), the right-hand syllable is a suffix and also unstressed, so that either the prefix or the suffix bears a default tone [l]. The compound in (102b) is obviously not iambic, but rather trochaic. Thus, sometimes, whether disyllabic compounds are right-headed or left-headed is also lexically conditioned, although the general metrical structure of SX is iambic. The fact that affixes are never stressed is also captured by the constraint *NON-HD/H. I assume that all affix-like syllables are assigned default tone in SX except when segmental constraints like IDENT-*ru* come into play, as in

[ʔɛʔ^{5}] ‘ah’, a prefix used before monosyllabic surnames (or numbers) to form terms of endearment. Thus, [h] on the left-hand syllable remains unchanged because of the IDENT-*ru* constraint while assigned a default tone, which allows any tone to be possible on the right-hand syllable in a foot domain because the left-hand high-level tone [h] in this case bears the same value as a default tone and thus the [h]+[T] combination survives the WFC(F) constraint. This phenomenon means that if the first syllable with an [h] tone is a prefix and should be assigned the value of a default tone, the tone of the second syllable remains unchanged to satisfy IDENT-IO(T). This is confirmed by the data. For example:

(103) [h]+[T] ([h] is a prefix):

[ʔɛʔ^{5}sɛ̃^{52}]	→	[ʔɛʔ^{5}sɛ̃^{52}]	[h.hl]→[h.hl]	‘third child’
[ʔɛʔ^{5}ŋ̍^{13}]	→	[ʔɛʔ^{5}ŋ̍^{13}]	[h.lh]→[h.lh]	‘fifth child’
[ʔɛʔ^{5}do^{22}]	→	[ʔɛʔ^{5}do^{22}]	[h.l] →[h.l]	‘eldest child’
[ʔɛʔ^{5}tɕ^{h}ɪʔ^{5}]	→	[ʔɛʔ^{5}tɕ^{h}ɪʔ^{5}]	[h.h] →[h.h]	‘seventh child’

The examples in (103) show that all tones after an [h] (*ru*) tone which is on an affix syllable, remain unchanged, even though the second tone is [l]. This requires a special constraint dominating WFC(F), *NON-HD/H and *HD/L, for which I propose a tone in-situ constraint:

(104) Tone in-Situ (T-IN-SITU)
An affix-like syllable with [h] tone on the left-hand should allow the right-hand tone to stay in-situ (no change for the right-hand tone).

T-IN-SITU in (104) is required by the fact that the left-hand [h] (*ru*) tone is on an unstressed affix-like syllable so that the right-hand syllable, whether it has a high tone or low tone, is the foot head. T-IN-SITU in (104) permits an ‘unnatural’ output in terms of tonal prominence scale, as exemplified in (103). However, all the sandhi rules and sandhi constraints in SX take effect in the foot domain, confirmed by all the data we have discussed. In the next subsection, I will present an OT analysis incorporating these constraints for all disyllabic tone sandhi rules in SX.

5.10 An OT Approach to Tone Sandhi

In the data I have presented above, it is clear that in general, SX has a right-prominent metrical structure and that the stress foot is the tone-sandhi domain. Tone sandhi behaves differently under different lexical or syntactical conditions. I have also claimed that tone sandhi in SX is phonologically motivated in terms of metrical structure. However, different languages have different rhythmic patterns, and will therefore have different tone sandhi rules. Tone sandhi, therefore, serves as an "effective diagnostic probe into the anatomy of the complex entity we call tone" (Chen 2000).

Sandhi forms represent the optimal rhythmic melodies in a tone language. Any investigation of tone sandhi rules sheds light on the constraints of syllable structure, consonant-tone interaction, metrical structure, and word formation. Although tone sandhi rules can be different from language to language, the general principles will be cross-linguistically consistent and comprise phonological constraints on optimal outputs. Following the data and the analysis presented previously, I propose the following hierarchical constraint ranking to account for tone sandhi in SX:

(105) Constraint ranking of disyllabic tone sandhi in SX:
T, OCP(c), IDENT-BR[C], IDENT-*ru*, T-IN-SITU ≫ DEFAULT-T(R/L) ≫ IDENT-TT(R) ≫ WFC(F) ≫ IDENT-IO(T) ≫ *HD/L ≫ *NON-HD/H.

All the constraints mentioned in (105) have been discussed previously in this chapter. To bear in mind what they are, I present the statements of these constraints again as follows:

(106) a. T: Every stressed syllable must have a tone.
b. OCP(c): Adjacent identical contour features (disregarding register feature) are prohibited in the same phonological phrase.
c. IDENT-BR[C]: Let α be a tone type in B, and β be a correspondent of α in R.
If α is [γC], then β is [γC].

d. IDENT-*ru*: If the high-level tone [h] occurs in the input, it may also appear in the output and vice versa (ru tone cannot be changed and cannot be formed).
e. T-IN-SITU: An affix-like syllable with [h] tone on the left-hand should allow the right-hand tone to stay in-situ (no change for the right-hand tone).
f. DEFAULT-T(R/L): The right/left-hand syllable is assigned a default tone [l] which is unstressed for lexical or grammatical reason.
g. IDENT-TT(R): The two adjacent rising contours in the input should also have two contours as the same tone type in the output (no level tone).
h. WFC(F): The prominence ranking may not decrease from left to right in the foot domain (two equally strong syllables are well-formed).
i. IDENT-IO(T): A tone in the input should have the same value in the output ("no change of tone").
j. *HD/L: The low tone [l] is not preferred by the foot head.
k. *NON-HD/H: The foot non-heads do not prefer high tones.

These constraints are so ranked as in (105) to account for the data and all the disyllabic sandhi rules of SX. However, not all the constraints above are relevant to every sandhi form. I rank T, OCP (c), IDENT-BR[C], IDENT-*ru* and T-IN-SITU at the top, for these five constraints are all inviolable in SX so that there is no hierarchical difference between them. Among the five constraints, T and OCP(c) are markedness constraints and IDENT-BR[C], IDENT-*ru* and T-IN-SITU are faithfulness constraints. All inviolable constraints in a given language are undominated while all violable constraints are dominated. The other six constraints are all violable in one way or another.

DEFAULT-T(R/L) is ranked above all other violable constraints because it is almost inviolable except for one case that the affix-like syllable has the *ru* tone which cannot change to a low tone because of the dominating constraint IDENT-*ru*.

IDENT-TT(R) is dominated by T, OCP(c), IDENT-*ru*, or DEFAULT-T(R/L) because of the fact that the identical rising contours [lh.lh] cannot have such sandhi forms as *[∅.lh], *[lh.lh], *[h.lh] and *[lh.h], but may have [l.lh], as indicated by the tone sandhi configurations in (45).

WFC(F) is a general principle of metrical structure in the foot domain. It is always respected, except in some special cases concerning those constraints like IDENT-*ru*, T-IN-SITU and DEFAULT-T(R/L) which allow some optimal sandhi forms to be against WFC(F). Therefore, WFC(F) should be ranked just next to IDENT-TT(R) and dominate the other three constraints.

IDENT-IO(T) is a faithfulness constraint which says that any optimal output should retain as many features from the input as possible. Thus, it is an important violable faithfulness constraint and should be ranked above the rest two violable markedness constraints. It is always true that an optimal output is the most faithful to the input unless some more important markedness constraints are violated.

*HD/L and *NON-HD/H are both markedness constraints, which require an optimal sandhi form to be faithful to WFC(F) and request that an ideal disyllabic sandhi form should be [l]+[h]/[lh]/[hl] for the iambic foot structure. A high-toned head syllable is more necessary than a low-toned non-head syllable, so that *HD/L dominates *NON-HD/H. However, both are very often violated and thus ranked very low, as presented in (105).

5.10.1 Identical contours

Bearing in the mind all the constraints stated in (106), let us examine how tone sandhi works in SX. With the constraint ranking in (105), the two adjacent identical contours have different sandhi forms between reduplication and non-reduplication. For example, the reduplicated compound [ɦia^{31}ɦia^{31}] 'grandfather' and the non-reduplicated compound [dɒŋ^{31}lɒŋ^{31}] 'mantis' have the same base tones [hl.hl], but are realised differently in sandhi forms as [lh.hl] in (107) and [l.hl] in (108), respectively. This can be explained in the following two tableaux:

(107) Reduplication

Input: hl.hl ([ɦia^{31}ɦia^{31}])	OCP (c)	IDENT -BR[C]	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. hl.hl	*!					*
b. hl.lh			*!	*		*
c. ☞ lh.hl				*		*
d. h.hl		*!		*		*
e. l.h		*!		**		

(108) Non-reduplication

Input: hl.hl ([dɒŋ^{31}lɒŋ^{31}])	OCP (c)	IDENT -BR[C]	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. hl.hl	*!					*
b. hl.lh			*!	*		*
c. lh.hl				*		*!
d. h.hl				*		*!
e. ☞ l.hl				*		

In both (107) and (108), T, IDENT-TT(R), IDENT-*ru*, T-IN-SITU, DEFAULT-T(R/L) and *HD/L are not listed to save space, because they are irrelevant to the output candidates in the two tableaux above. However, the ranking must be the same. In (107), the input [hl.hl] ([ɦia^{31}ɦia^{31}] 'grandfather') is a reduplication form, so that candidates (d) and (e) violate IDENT-BR[C] and are also ruled out after candidates (a) and (b) are ruled out. As a result, dandidate (c) is the optimal output. In (108), the input [hl.hl] ([dɒŋ^{31}lɒŋ^{31}] 'mantis') is not a reduplication form, so that the constraint IDENT-BR[C] is irrelevant to all the candidates. Thus, candidate (e) is the winner according to the constraint ranking in (105), as presented in tableau (108). When it comes to adjacent identical rising contours, the constraint IDENT-TT(R) is involved. However, with the same constraint ranking in (105), two adjacent identical rising contours (whether they are in forms of reduplication or non-reduplication) have the same tone sandhi, as exemplified with the reduplicated [tɕi^{35}tɕi^{35}] 'elder sister' and non-reduplicated [dɑɒ^{13}ts^{h}ɑɒ^{35}] 'straw' in tableaux (109) and (110), respectively:

(109) Reduplication

Input: lh.lh ([tɕi^{35}tɕi^{35}])	OCP (c)	IDENT -BR[C]	IDENT -TT(R)	WFC (F)	IDENT -IO(T)	*HD /L	*NON -HD/H
a. lh.lh	*!						*
b. hl.lh				*!	*		*
d. ☞ lh.hl					*		*
d. h.lh		*!	*		*		*
e. l.lh		*!	*		*		

(110) Non-reduplication

Input: lh.lh ([dɑɒ^{13}ts^{h}ɑɒ^{35}])	OCP (c)	IDENT -BR[C]	IDENT -TT(R)	WFC (F)	IDENT -IO(T)	*NON-HD/H
a. lh.lh	*!					*
b. hl.lh				*!	*	*
e. ☞ lh.hl					*	*
d. h.lh			*!		*	*
e. l.lh			*!		*	

The tableaux in (109) and (110) show that two rising contours in reduplication and non-reduplication forms have the same sandhi forms. In (109), candidates (d) and (e) are ruled out by the constraint IDENT-BR[C], and in (110), candidates (d) and (e) are ruled out by the constraint IDENT-TT(R). In (110), I omit *HD/L to save space, because none of the candidates violate *HD/L. However, all different forms of tone sandhi are analysed through the same hierarchical constraint ranking in (105). This constraint ranking determines the different sandhi forms between a pair of rising contours and a pair of falling contours and between contours in reduplication and those in non-reduplication.

5.10.2 Identical level tones

The output sandhi forms of reduplications with identical level tones can also be predicted in the OT analysis with the same hierarchical constraint ranking that was presented in (105). For identical level tones, OCP(c) is vacuously satisfied and IDENT-TT(R) is irrelevant. To the analysis of identical level tones, I add the constraints T and *HD/L, which become relevant here. For example, the reduplicated [k^{h}ẽ^{33}k^{h}ẽ^{33}] 'have a look' has a sandhi form which is identical to the input, as shown in (111): [37]

(111) Reduplication

Input l.l	T	OCP (c)	IDENT -BR[C]	WFC (F)	IDENT -IO(T)	*HD /L	*NON HD/H
a. ☞ l.l						*	
b. ∅.l	*!		*		*	*	
c. hl.l			*!	*	*	*	*
d. l.lh			*!		*		
e. l.h			*!		*		

[37] ∅ refers to the absence of tone.

Since [k^{h}ẽ^{33}k^{h}ẽ^{33}] is in reduplication form and [l.l] is of non-contour tone type, candidates (b), (c), (d) and (e) in (111) all violate IDENT-BR[C], so that candidate (a) is obviously the winner. As was discussed previously, no new tone feature can be inserted in tone sandhi rules in SX. The identical level tones [l.l] only involve two [l] features, so that neither tone can change to any other tone, because there is no [h] feature available. The same is true with the identical level tones of non-reduplication, as exemplified with [tɕi^{33}ȵjɛ̃^{22}] ‘memorize’ in (112):

(112) Non-reduplication

Input l.l	T	OCP (c)	IDENT -BR[C]	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. ☞ l.l						*	
b. ∅.l	*!				*	*	
c. hl.l				*!	*	*	*
d. l.lh					*!		
e. l.h					*!		

The reduplication structure with identical high level tones can be analysed in the same way, e.g. [pɛʔ^{5}pɛʔ^{5}] ‘pat’, whose sandhi form is also [pɛʔ^{5}pɛʔ^{5}], as shown in (113):

(113) Reduplication

Input h.h	T	OCP (c)	IDENT -BR[C]	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. ☞ h.h							*
b. ∅.h	*!				*		
c. lh.h			*!		*		*
d. l.h					*!		

The tableau in (113) shows that candidate (b) is the first to be ruled out, because it violates T; candidate (c) violates IDENT-BR[C] and IDENT-IO(T), so that it is also ruled out; candidate (d) also violates IDENT-IO(T), because the first [h] changes into [l]. Thus, candidate (a) is the optimal output. In fact, [h.h] cannot change in tone sandhi, because no [l] feature can be inserted. Therefore, the identical [h.h] of non-reduplication still have the same sandhi forms as those of the identical [l.l] in (111) and (112).

5.10.3 Different contours

There are two types of different-contour combinations in a disyllabic compound or a phrasal expression: falling + rising [hl.lh] and rising + falling [lh.hl]. The former violates WFC(F) and will end up with different tone(s) in the sandhi form; the latter satisfies WFC(F) and would likely remain unchanged in the sandhi form. Consider the example [bɪŋ^{31}dɛ̃^{13}] 'common', which can be analysed as in (114):

(114)

Input hl.lh	OCP (c)	IDENT -*ru*	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. hl.lh			*!			*
b. hl.hl	*!			*		*
c. lh.hl				**!		*
d. ☞ l.lh				*		
e. h.lh		*!		*		*
f. hl.h		*!	*	*		*

In tableau (114), I do not list constraints like T, IDENT-BR[C], IDENT-TT(R), T-IN-SITU and DEFAULT-T(R/L) to save space, because these constraints are irrelevant to the candidates. If listed, these constraints have no violation from the candidates in (114). This tableau shows that candidate (b) violates OCP(c) and is ruled out as a result; candidate (a) violates WFC(F) and is also ruled out; candidates (e) and (f) violate IDENT-*ru*, because no new *ru* tone can be formed in tone sandhi; candidate (c) violates IDENT-IO(T) twice and is finally ruled out; candidate (d) is the winner, so that [bɪŋ^{22}dɛ̃^{13}] is the sandhi output. In the same way, we can work out the optimal output of the sandhi form [lh.hl], which usually remains unchanged, because it satisfies WFC(F). Taking [ŋɛ̃^{13}tɕɪŋ^{52}] 'eye' as an example, the optimal output is still [ŋɛ̃^{13}tɕɪŋ^{52}], as shown in tableau (115):

(115)

Input lh.hl	OCP (c)	IDENT -*ru*	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. ☞ lh.hl						*
b. h.hl		*!		*		*
c. l.hl				*!		
d. l.h		*!		**		

The tableau in (115) has the same constraints as those in (114), the ranking of which follows that in (105). In (115), candidates (b) and (d) violate IDENT-*ru* and are therefore ruled out; candidate (c) is also ruled out because it violates IDENT-IO(T); candidate (a) violates the constraint that is ranked lowest and [lh.hl] is obviously the winner.

However, as was discussed above, the tones in the foot domain can be lexically or syntactically conditioned. If in a [lh.hl] tone pattern, the first rising contour occurs on an affix-like syllable[38] (which is assigned the default tone), the result will be different, because the constraint DEFAULT-T(L) (left tone should have default tone) comes into play. For example, [lɑɒ^{13}ɦwɒŋ^{31}] 'Old Wang', in which [lɑɒ^{13}], as a prefix, is assigned the default tone [l] in tone sandhi, as shown in (116):

(116)

Input lh.hl	OCP (c)	IDENT -*ru*	DEFAULT -T(L)	WFC (F)	IDENT -IO(T)	*HD /L	*NON -HD/H
a. lh.hl			*!				*
b.☞l.hl					*		
c. l.l					**!	*	
d. l.h		*!			**		

According to the constraint ranking in (105), the relevant constraint, DEFAULT-T(L), is dominated by IDENT-*ru* and dominates WFC(F), as shown in (116). The tableau in (116) shows that candidate (a) violates DEFAULT-T(L) and is ruled out; candidates (b), (c) and (d) all have [l] tone on the left-hand syllable, satisfying DEFAULT-T(L), but (d) is ruled out because it violates IDENT-*ru*; (c) violates IDENT-IO(T) twice and is also ruled out; (b) is the winner and thus [lɑɒ^{22}ɦwɒŋ^{31}] is the output sandhi form. Whatever type the input contour is, if it is assigned the default tone for any lexical or morphological reason, it has [l] in its sandhi form. One more example is [lɑɒ^{13}toŋ^{13}] 'Old Dong', in which [lɑɒ^{13}] is also a prefix. The sandhi form is presented in (117):

---

[38] In SX, which is in principle a monosyllabic language, every possible affix syllable can also be a lexical stem syllable and the underlying tone is fixed on a syllable, viz. every syllable has the same underlying tone, whether it is a lexical stem or an affix. The latter is assigned a default tone only in surface representation, i.e. sandhi form. That is why every syllable should be potentially bimoraic for purposes of stress.

(117)

Input lh.lh	OCP (c)	DEFAULT -T(L)	IDENT -TT(R)	WFC (F)	IDENT -IO(T)	*HD /L	*NON -HD/H
a. lh.lh	*!		*		*		*
b.☞l.lh			*		*		
c. l.l			*		**!	*	
d. lh.hl		*!			*		*

In tableau (117), I omit the constraint IDENT-*ru*, which is irrelevant to all the output candidates as well as the input. IDENT-TT(R) plays a role, because the input is a pair of identical rising contours which usually have two different contours in sandhi forms, required by the contraint. However, IDENT-TT(R) is dominated by DEFAULT-T(L), as shown in (105). Thus, the correct sandhi form of [lɑɒ^{13}toŋ^{13}] is [lɑɒ^{22}toŋ^{13}], as explained by the tableau in (117).

Under certain lexical conditions, the default tone is sometimes assigned to the right-hand syllable, which violates WFC(F). This shows that DEFAULT-T(R/L) dominates WFC(F). Consider [dʑjẽ^{22}kɒŋ^{52}] ‘health’, in which the two lexical syllables are in fact synonyms and have one meaning. In this case, the second syllable is assigned the default tone, as shown in (118):

(118)

Input l.hl	OCP (c)	DEFAULT -T(R)	IDENT -TT(R)	WFC (F)	IDENT -IO(T)	*HD /L	*NON -HD/H
a. l.hl		*!				*	*
b. hl.hl	*!	*			*		*
c. lh.l				*!	**		
d. ☞l.l					*	*	

The tableau in (118) shows that candidate (b) is the first to be ruled out because it violates OCP(c); candidate (a) violates DEFAULT-T(R) and is also ruled out; candidate (c) is ruled out because it violates WFC(F); candidate (d) is thus the winner, so that [dʑjẽ^{22}kɒŋ^{33}] is the optimal sandhi form. For the sandhi form [l.l] in (118), the right-hand syllable is assigned the default tone, so it cannot be the foot head in that lexically unstressed syllables, e.g. affixes, cannot be the head. Another example is [zoʔ^{3}ɕjɑɒ^{35}] (lit. ‘weak small’) ‘weak’, in which the lexical meaning is in the first syllable so that the second syllable is assigned the default tone [l]. As a result, the sandhi form is [h.l] ([zoʔ[3]ɕjɑɒ[33]]) rather than [h.lh], as explained by the tableau in (119):

(119)

Input h.lh	OCP (c)	IDENT -*ru*	DEFAULT -T(R)	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. h.lh			*!				*
b. lh.lh	*!	*			*		*
c. ☞h.l				*	*		
d. l.lh		*!	*		*	*	
e. l.l		*!			**	*	

In (119), IDENT-*ru* is inserted, for the [h] is a *ru* tone on checked syllables and can never change while IDENT-TT(R) becomes irrelevant because the input is not a pair of rising contours. However, the constraint ranking is the same as that in (105). The tableau in (119) shows that candidate (b) is the first to be ruled out, because it violates OCP(c); both (d) and (e) are ruled out for their violating IDENT-*ru*; (a) is also ruled out because it violates DEFAULT-T(R); the candidate [h.l] in (b) is the optimal output for the sandhi form of [h.lh], the result of which, though, violates WFC(F), so that [h.l], as an exception, is trochaic, so that the left-hand syllable is the foot head, as shown by the tableau in (119). This phenomenon illustrates the interplay between lexical information and prosodic structure in SX tone sandhi.

5.10.4 Combination of a contour and a low tone

Any combination of contour + low level tone, either [hl.l] or [lh.l], violates WFC(F), so that contour simplification will usually occur on the first tone, or contour formation will occur on the second tone. For example, [ʔi[52]ko[33]] 'clothes rack' has the base tones [hl.l], in which contour simplification is preferred to contour formation, as the OT analysis predicts, as shown in (120):

(120)

Input hl.l	OCP (c)	IDENT -ru	DEFAULT -T(R/L)	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. hl.l				*!		*	*
b. hl.hl	*!				*		*
c. ☞ l.l					*	*	
d. l.hl					**!		
e. h.hl		*!					*

In (120), the constraint DEFAULT-T(R/L) is in fact irrelevant, because no syllable is assigned the default tone for any lexical reason. The tableau in (120) shows that candidate (c) is the optimal output, although [l] is not preferred in the prominent syllable. Candidate (d) is impossible because not only it violates IDENT-IO(T) twice, but also the [h] feature cannot spread to the right-hand syllable across the [l] feature, which is not allowed by the tone sandhi principle in (87d). Thus, the observed sandhi form is [ʔi^{33}ko^{33}]. This shows that [hl] cannot occur in the left-hand syllable in the right-prominent foot domain. However, the combination of [lh.l] will have a different sandhi form, as shown by [me^{13}li^{22}] 'beautiful' in the following tableau:

(121)

Input lh.l	OCP (c)	IDENT -ru	DEFAULT -T(R/L)	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. lh.l				*!		*	*
b.☞lh.hl					*		*
c. l.l					*	*!	
d. lh.lh	*!				*		*
e. l.hl					**!		

In (121), candidate (b) wins over candidate (c) because of the constraint ranking that *HD/L dominates *NON-HD/H, which strongly suggests that the foot-head syllable with a high tone is more important than the foot-non-head syllable with a low tone. Since [l] is the lowest tone type, it is the least preferred by the right-hand syllable in iambic structure. To amplify the stress of the right prominence, [l] in the right-hand syllable always receives an [h] feature from the left tone if there is one and becomes a falling tone [hl], as shown in (121) unless the left-hand contour is assigned the default tone. For example, in [lɑɒ^{13}do^{22}] 'the eldest child', the first syllable [lɑɒ^{13}] is a prefix and is assigned an [l] tone in the sandhi form. In this case, the right-hand [l] will not change, as is shown in (122):

(122)

Input lh.l	OCP (c)	IDENT -*ru*	DEFAULT -T(L)	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. lh.l			*!	*		*	*
b. lh.hl			*!		*		*
c. ☞ l.l					*	*	
d. l.hl					**!	*	
e. lh.lh	*!		*		*		*

In (122), DEFAULT-T(L) comes into play because the left-hand syllable is a prefix for the lexical formation. The tableau in (122) shows that candidates (a) and (b) violate DEFAULT-T(L) and are therefore ruled out; (d) is ruled out because this candidate violates IDENT-IO(T) twice, though it satisfies WFC(F); (e) is the first to be ruled out because it violates OCP(c); candidate (c) is the winner, so that [lɑɒ^{22}do^{22}] is the resulting sandhi form. However, the simple reason why /hl.l/ simplifies to [l.l], but /lh.l/ becomes [lh.hl], is that spreading the adjacent tone feature in /hl.l/ would produce *[hl.ll], which does not solve anything, while spreading in /lh.l/ would just produce [lh.hl], which satisfies the right-prominent foot structure. These tone sandhi rules further prove that tone sandhi in SX is phonologically realized by tone feature spreading or delinking.

5.10.5 Combination of a low tone and a contour

Any combination of low tone with a contour satisfies the WFC(F), so that the tone on the right syllable remains unchanged. Prosodically, [l.hl] or [l.lh] are ideal patterns for the right-prominent metrical structure, as evidenced by many tone-sandhi data in SX. For example, [k^{h}ẽ^{33}ɕy^{52}] 'read a book' has the same sandhi form as its base tones, expressed by the OT tableau in (123):

(123)

Input l.hl	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. ☞l.hl				
b. lh.hl		*!		*
c. l.l		*!	*	
d. l.lh		*!		

In (123), I only list the last four of the eleven constraints in (105) because all the others are irrelevant to the sequences of a low level tone + a contour. Tableau (123) shows that candidates (b), (c) and (d) are all ruled out on account of their violation of IDENT-IO(T); (a) violates nothing and is the absolute winner. Thus, the optimal sandhi form is the fully faithful candidate [kʰẽ³³ɕy⁵²]. The same holds for [l.lh]. For example, [ma²²mi²³] 'buy rice' also has the same sandhi form as the input, as shown in (124):

(124)

Input l.lh	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. ☞l.lh				
b. lh.lh		*!		*
c. l.l		*!	*	

This formalizes the generalization we uncovered above, viz. that when the low-level tone is the first tone in the foot domain, the base tones always remain unchanged in the sandhi forms. The OT analysis always chooses the candidate with the least change from the input.

5.10.6 The *ru* tone in tone sandhi

Although Maddieson (1978) finds that tones under stress generally remain unchanged, while tones in metrically weak positions tend to assimilate to more prominent tones, the unchangeable [h] tone in SX has different properties, because it does not change even in metrically weak position. As was discussed earlier, [h] is the high-level (*ru*) tone, used exclusively in checked syllables ending with the glottal stop [ʔ]. Thus, the high-level tone can never change and no other tone can change into the high-level tone either, because the high-level tone and the checked syllable structure must go hand-in-hand. Nevertheless, the syllable structure can never change unless the lexical meaning is changed.

Because of the special properties of the [h] tone, in [h.T] combinations, the [h] feature always spreads to a following tone which has to delink its original [h], if it is a rising contour, to change into a falling contour, making the right-prominent syllable properly stressed. In any case, the [h] tone remains unchanged whether it is in a prominent position or in a metrically weak position. For example, [sɪʔ⁵ba²²] 'defeat' has the base tones [h.l] and its sandhi form is [h.hl], as shown in (125):

(125)

Input h.l	IDENT-*ru*	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. h.l		*!		*	*
b.☞ h.hl			*		*
c. l.l	*!		*	*	
d. h.h	*!		*		*

Since the [h] tone is an unchangeable tone, when one tone in any tone sandhi is [h], the constraint IDENT-*ru* must be applied and ranked high, as assumed in the constraint ranking in (105). The tableau in (125) shows that candidates (c) and (d) violate IDENT-*ru* (which also stipulates that formation of a new *ru* tone is not allowed) and are therefore ruled out; (a) is also ruled out since it violates WFC(F); (b) is then the winner, yielding [sɪʔ^{5}ba^{31}] as the sandhi form. When the [h] tone is in the prominent position, the first tone remains unchanged if WFC(F) is not violated. For example, [fẽ^{35}pɪʔ^{5}] 'chalk' has the same sandhi form as its input, as shown in (126):

(126)

Input lh.h	IDENT-*ru*	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. ☞ lh.h					*
b. lh.lh	*!		*	*	
c. l.h			*!		

If the first tone has a higher stress value than the second [h] tone, viz. [hl], the first tone has to lose its [h] feature to satisfy WFC(F) during which contour simplification occurs. For example, in [dʑjaŋ^{31}pɪʔ^{3}] 'wall', the first [hl] has to change to [l] in the sandhi form, as shown in (127):

(127)

Input hl.h	IDENT-*ru*	WFC (F)	IDENT -IO(T)	*HD /L	*NON-HD/H
a. hl.h		*!			*
b. hl.hl	*!		*		*
c. ☞ l.h			*		
d. lh.h			*		*!

The tableau in (127) shows that candidates (a) and (b) violate WFC(F) and IDENT-*ru*, respectively, and are ruled out; (d) is also ruled out because it violates *NON-HD/H; (c) is the optimal output since it satisfies

*HD/L and *NON-HD/H and is preferred by the iambic foot, although the candidate also violates IDENT-IO(T). However, a problem occurs when the left, metrically weak syllable has the [h] tone underlyingly and it should be assigned the default tone when it is unstressed for lexical reason, e.g. as a prefix, in surface representation. As was discussed above, lexical information dominates prosodic information in tone sandhi in SX. Does the [h] tone change to [l] in this case? Consider the case of [ʔɛʔ^{5}do^{22}] 'the eldest child', in which the first syllable [ʔɛʔ^{5}] with the *ru* tone plays a role as a prefix and it should be assigned the default tone. But, as a *ru* tone with a special condition on syllable structure, viz. ending in the glottal stop [ʔ], the tone [5] cannot change, even though it is assigned the default tone, because IDENT-*ru* is inviolable and dominates DEFAULT-T(R/L) in SX tone sandhi, as shown in (105). In this case, the left-hand [h], possessing the value of default tone, assigns the right-hand tone the status of tone-in-situ, allowing the right-hand tone to remain unchanged. This shows that T-IN-SITU is as important as IDENT-*ru* in the SX tone sandhi. T-IN-SITU makes it possible that the optimal output may violate WFC(F). The sandhi form of [ʔɛʔ^{5}do^{22}] is presented through the same hierarchical constraint ranking, as shown in (128):

(128)

Input h.l	IDENT-*ru*	T-IN-SITU	DEFAULT-T(L)	WFC (F)	IDENT-IO(T)	*HD/L	*NON-HD/H
a. ☞ h.l			*	*		*	*
b. h.hl		*!	*		*		*
c. hl.l	*!		*	*	*	*	*
d. l.l	*!				*	*	

Since [ʔɛʔ^{5}] in [ʔɛʔ^{5}do^{22}] is a prefix, it should be assigned a default tone. However, DEFAULT-T(L) is dominated by IDENT-*ru* and T-IN-SITU. The tableau in (128) shows that candidates (c) and (d) are ruled out because they violate IDENT-*ru*; (b) is ruled out because it violates T-IN-SITU; (a) is the winner so that the optimal sandhi form is still [ʔɛʔ^{5}do^{22}], even though [h.l] violates DEFAULT-T(L) and WFC(C).

All in all, the constraint ranking in (105) captures the facts of all the disyllabic tone sandhi in SX and accounts for the phonological behaviour of tones in the foot domain.

5.11 Summary

In this chapter I attempted to account for four aspects of the tonal system of SX. First, I formalized the specifications of tones and assumed that tones in SX are classified into a high and a low register, specified as [H] and [L], respectively, and divided into four types of tones, specified as [hl], [lh], [l] and [h]. Secondly, I scrutinized the consonant-tone correlation in SX and claimed that neither voiced initial obstruents nor low-register tones are allophones or allotones, respectively. I showed how the voiced obstruent-L and voiceless obstruent-H correlation can be captured. Thirdly, I formalized the complexities of tone sandhi in SX, assuming that tone sandhi is phonologically realized by tone feature spreading and delinking. I also proposed that tone sandhi is phonologically motivated by metrical structure, in which the foot domain and the stress foot coincide and constitute the tone-sandhi domain in SX. Finally, I presented my OT analysis of tone sandhi in SX, proposing a tone-sandhi constraint ranking which captures the behaviour of tones in the metrical structure and accounts for all the possibilities of disyllabic sandhi forms in the foot domain.