The Phonology of Shaoxing Chinese

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
http://wwwlot.let.uu.nl

Cover illustration: A mural painting of Emperor *Gou Jian* of the Yue Kingdom (497-465 B.C.) (present-day Shaoxing). The photo was taken by Xiaonan Zhang in Shaoxing.

ISBN 90-76864-90-X
NUR 632

The Phonology of Shaoxing Chinese

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D.D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 31 januari 2006
klokke 15.15 uur

door

JISHENG ZHANG

geboren te Shaoxing, China
in 1955

Promotiecommissie

promotor:	prof. dr. V.J.J.P. van Heuven
co-promotor:	dr. J.M. van de Weijer
referent:	prof. dr. M. Yip (University College London)
overige leden:	prof. dr. C.J. Ewen
	dr. M. van Oostendorp (Meertens Instituut)
	dr. N.S.H. Smith (University of Amsterdam)

Dedicated to my mother who gave me my life and brought me up on this ancient land — Shaoxing.

Contents

Acknowledgements

Writing this thesis was my life-long ambition, but would never have been accomplished and brought to a favourable conclusion without the help of those to whom I would express my heart-felt thanks.

Leiden regulations forbid me to thank the members of my doctoral committee for their guidance and encouragement from the initial to the final stages of this work and for their critical reading of the manuscript, detailed and extensive comments and instructive and valuable advice.

Prof. Iggy Roca has been an influential scholar in my phonological studies. My visit to him in Colchester in 2003 benefited me a lot, as a result of his profound knowledge and his insightful opinions about phonological theories. Prof. Wang Futang, an expert at dialectology in Beijing University, China, and also a Shaoxing native, helped me through his letters to me, discussing the phonetics of the onset consonants in Shaoxing.

I have profited greatly from discussions in the ULCL phonology group, held every week, which provided everyone with the opportunity to freely and extensively exchange phonological ideas and learn from each other.

I wish to thank also all my colleagues in LUCL (previously ULCL), not only in the phonological and phonetic circles, but also in syntax. Talks with everyone during my stay in Leiden have enriched my linguistic knowledge. Besides, their friendly and considerate help, and good humour, always made me feel at home.

Although I am a native speaker of Shaoxing myself, I have also learned great deal from my family members in Shaoxing and neighbours who helped with data that I could record and collect. I thank them for allowing me to consult them.

My PhD study in Leiden University was privately funded, for which I feel deeply indebted to my family, my wife and my son, who fully understand me and who have given me enormous support during these years.

Many thanks should go to my friends at home and abroad whose constant care and concern helped me to go through these years.

Thanks should also go to my former home university, Ningbo University, China, which made it possible for me to work on my PhD project in Leiden between the intervals of my teaching in Ningbo during these years.

Finally, I wish to thank the China Scholarship Council (CSC), which financially supported my first visit to Leiden University for half a year as a visiting scholar in 2001, which gave me the opportunity to pursue my PhD study in phonology in LUCL.

1 Background

1.1 Introduction

There are seven large language families in China: the Northern family (which includes Mandarin, which is considered to be Standard Chinese), the Wu family (which includes Shaoxing, the topic of this dissertation), the Gan family, the Min family, the Xiang family, the Yue family (which includes Cantonese, which is the other most prestigious variety of Chinese), and the Hakka family (see e.g. Chao 1967; Zhan 1991; Simmons 1999; Campbell 2003). Wu and the Northern family are the largest Chinese language families. The Wu languages, which have about 75 million speakers, are spoken in the area around Shanghai, about one fourth of the Jiangsu Province, almost the whole of Zhejiang Province (where Shaoxing is located) and some northern parts of the Jiangxi and Fujiang provinces. Currently, 94 Wu languages (or dialects) are spoken in this area (Campbell 2003). Of the seven Chinese language families, the Wu languages show the largest amount of variation. Even within the Zhejiang Province, some Wu languages are too different to be mutually intelligible (Zhan 1991; Yan 1994; Cao 2002; among others). For example, Yangzhou and Wenzhou, which are both Wu languages, are mutually unintelligible (Cheng 1992).[1] As a Shaoxing native speaker, I am completely unable to communicate with Wenzhou native speakers in our own dialects.

This dissertation focuses on Shaoxing Chinese, which is spoken by the inhabitants of the city of Shaoxing. In 2003, the number of speakers was estimated at about 633,000 (Campbell 2003).

It is believed that Shaoxing originally dates back to the ancient dialect of Chu, which was spoken in the Zhanguo ("Warring States") Period (475–221 B.C.). During this period, there were seven kingdoms, including Yue (present-day Shaoxing and some other parts of Zhejiang), Wu (present-day Suzhou, the centre of the modern Wu languages), Chu

[1] Cheng (1992) presents a syllable/constituent-based analysis of mutual intelligibility between 17 Chinese dialects which are grouped into 136 pairs and finds that the Yangzhou-Wenzhou dialects (which are both Wu) rank among the lowest of these pairs, indicating very low mutual intelligibility.

(present-day Hubei), Qin (present-day Shanxi) and others. In 473 B.C. Yue conquered Wu. Later, in 333 B.C., Chu conquered Yue. Then Wu and Yue became part of the Chu Kingdom and the Chu language began to be used in both the Wu and the Yue areas. This probably represented the very beginning of the Wu language family, which underwent numerous changes through the next centuries. The Chu Kingdom rose to great power in Middle China before it was finally defeated by Qin in 223 B.C. Emperor Qin (221–210 BC) united China in 221 B.C. after conquering all other kingdoms and divided the country into 36 counties. He made Kuiji (present-day Shaoxing City) the capital of the Wu-Yue county. However, the Wu dialects began to diverge after the Tang Dynasty (618–907 A.D.) and developed as the second largest language family in China, which is called Wu because it historically centred on the “Three-Wu” area, which includes present-day Suzhou, Huzhou and Shaoxing.

Shaoxing is therefore among the earliest Wu dialects and began to spread to many other parts in China, and even abroad, in early times. It is believed that Japanese [ȵi] ‘two’ was borrowed from the ancient Shaoxing dialect (Chen 1994) and many other Shaoxing expressions, such as [naʔgoʔ naʔgoʔ] ‘very’, can be found in Japanese with similar meaning. The typical features of Shaoxing include an eight-tone system and an opposition between voiceless and voiced obstruents, which are regarded as characteristics of Middle Chinese (6^{th}–10^{th} centuries A.D.) (Chao 1928). There are some other remarkable linguistic phenomena in Shaoxing, such as a rich obstruent inventory, which is rare among Asian languages, clitic morphemes, which are generally regarded as impossible in strictly isolating languages like Chinese, and ablaut (which plays a role in the pronoun system, e.g. /ŋo/ ‘I, me’ vs. /ŋa/ ‘we, us’ and /noʔ/ ‘you-SING’ vs. /na/ ‘you-PLUR’), which is never possible in Mandarin.

Much research has been done on Chinese languages, including Wu. Most efforts, however, have been directed at the description of the linguistic representations of different dialects from the perspective of phonetics, phonology, lexicology or syntax, diachronically or synchronically. Previous studies include, for instance, the classification of languages or dialects (e.g. Simmons 1999), comparison between Chinese languages (e.g. Zhan 1991), the history of the Chinese languages (e.g. Chen 1994), and the diachronic description of language change (Ding 1992; Cao 1998). There are also some phonological studies focusing on the Wu languages, mainly concerned with Shanghai, the Suzhou dialect or the Wenzhou dialect, by Chao (1928), Fu (1986) and Cao (2002, 2004). Very little work has been done on Shaoxing, however, especially with regard to phonological aspects. This dissertation seeks to present an overall analysis of the phonology of Shaoxing.

1.2 Methodology

Besides a description of the overall features of Shaoxing phonology, this dissertation will explore analyses of many phonological questions that have remained unanswered so far, focusing on the phonological realizations of segments, the syllable structure, the tonal system, and the phonotactics of Shaoxing. As a result of our discussions, we also hope to cast some light on general phonological issues in Chinese. For instance, we will discuss the question whether Chinese has triphthongs (cf. Wiese 1997), and the issues of Chinese syllable structure, particularly the status of the Chinese medials (/j/ and /w/) in syllable structure (cf. Bao 1990; Wang 1999; Duanmu 2000a; Wang & Chang 2001). Data for this dissertation come from published sources, my own fieldwork and my intuitions as a native speaker.

Theories of phonetics and phonology, especially generative phonology, have sought to provide an explanatory approach to many linguistic issues and will be applied to my analyses of the phonology of Shaoxing. Segmental theories (e.g. Clements 1985; Sagey 1986a, 1986b; Harris 1994) provide representations of segmental structure and offer possible approaches to the representation of complexity in segmental phonology. Segmental theory also helps to postulate the underlying vowel inventory of Shaoxing through an analysis of the distribution of different allophones and the behaviour of vowels in processes of assimilation or dissimilation.

Experimental phonetics and phonology (van Heuven 1994, 2004; Lass 1995) help to shed light on the phonetic properties of consonants, the status of prenuclear glides within the syllabic structure, and the issue of consonant-tone interaction in Shaoxing.

Many previous studies have dealt with tone in Chinese languages (e.g. Wang 1967; Woo 1969; Yip 1980, 1989, 2002; Pulleyblank 1986; Chan 1991; Bao 1999; Chen 2000). In this dissertation, too, we will investigate the relation between tones and segments. My dissertation will provide more evidence for consonant-tone interaction and present the geometrical structure of tone correlated with the syllable structure in Shaoxing. The theory of tone register, especially that proposed by Yip

(1980, 1989, 2002) and Bao (1999), will turn out to be very helpful with respect to the tonal structure of Shaoxing, especially as regards the relationship between register and tone pitch and the formalisation of tone sandhi in Shaoxing.

Theories of the syllable (e.g. Selkirk 1984; Anderson & Ewen 1987; Harris 1994; Blevins 1995; Broselow 1995) provide an explicit analysis of the syllable structure of Shaoxing through different approaches, which will especially be directed at the problem of the status of Chinese medials, which has remained largely unsolved (Wang 1999; Duanmu 2000; Yip 2003). Following Levin (1985), my dissertation offers a multiple-specifier X-bar syllable structure for the analysis of the status of prenuclear glides in Shaoxing.

Optimality theory (OT) (Prince & Smolensky 1993; Gussenhoven & Jacobs 1998; Kager 1999) is committed to the view that the difference between grammars (and languages) is a difference in constraint rankings, and that the set of constraints is universal and violable. OT shares with the *Sound Pattern of English* (*SPE*) (Chomsky & Halle 1968) the notion of an underlying form, or input, and, of course, both theories produce outputs. The difference is that *SPE* moves from input to output in a series of stages, called a derivation, whereas in OT the well-formedness constraints apply simultaneously to representations of structures and select one output. OT will be applied to the description of Shaoxing in my analysis, especially with respect to loanword phonology, vowel distribution and tone sandhi.

In the next subsections I will present the different subtheories in greater detail. In §1.6 I will present the overall organization of the dissertation.

1.3 Syllable Theory

After *SPE*[2], syllable theories developed in various directions and are used, among other things, to account for the phonotactics of languages. Levin (1985) and van der Hulst & Ritter (1999) present the most influential theories of the syllable. Their study not only gives a very good picture of the syllable structure of most European languages, but can also be employed to shed light on the syllable structure of Shaoxing. Although a

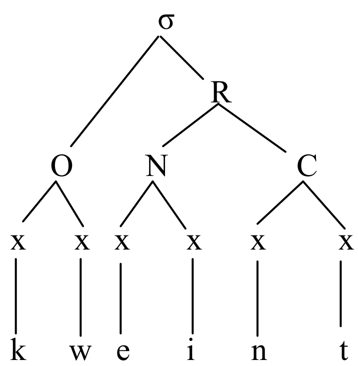
[2] *SPE* took a linear phonological approach and did not recognize syllable-internal structure.

full discussion of syllable theories is outside the scope of this thesis, I will briefly discuss some of the most important syllable theories, in particular the Onset-Rhyme model and mora theory.

1.3.1 Onset-Rhyme models

Onset-Rhyme (OR) models of syllable structure were developed in the mid twentieth century (see Pike & Pike 1947; Kuryłowicz 1948; Fudge 1969). Models like these maintain that the syllable is a prosodic unit with an internal hierarchical organization (Selkirk 1982, 1984) which maximally consists of Onset (O) and Rhyme (R). The Rhyme must have a Nucleus (N) and optionally also has a Coda (C). Since N is obligatory but the other constituents are not, possible structures are V, CV, VC, CVC, etc.: different languages permit different structures (see e.g. Blevins 1995). OR claims that the nodes can be maximally binary branching, so that the maximal syllable shape is CCVVCC (i.e. two onset consonants, a long vowel or diphthong as the nucleus and two coda consonants). Fudge (1969, 1987), Selkirk (1982) and many others analyze English syllable structure in terms of an OR model. In this model, the syllable structure of [kweint] ‘quaint’ is represented as in (1):

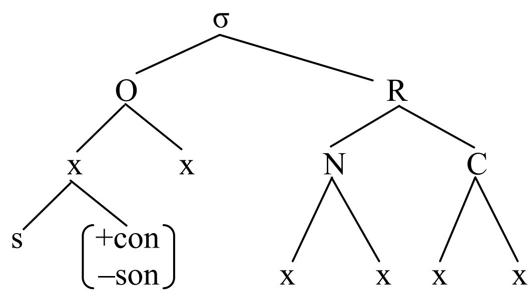
(1)



The structure in (1) represents a syllable with a branching onset and a branching rhyme, in which the R constituent dominates two further branching constituents: the nucleus and the coda. The terminals in this model are formed by the skeletal points (represented with x), i.e. the anchor point for root nodes or phonological features (Clements & Keyser 1983; van der Hulst & Ritter 1999). OR is adequate for most English syllables, except when it comes to three-member initial clusters such as [spr-], [str-], and [skr-], for which different solutions have been proposed.

Selkirk (1982) tried to solve this problem with an auxiliary syllable template, as shown in (2):

(2)



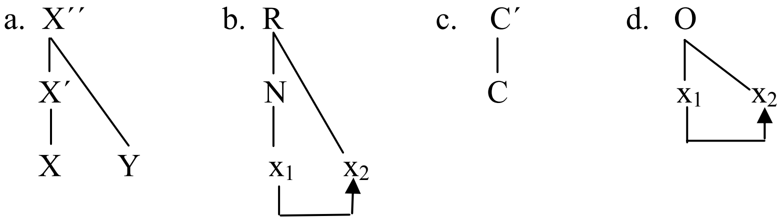
The template in (2) claims that [sp], [st], and [sk] in the sequences “spr-”, “str-”, and “scr-”, respectively, all take up only one consonant slot. Note that words like /nekst/ ‘next’ remain a problem in such an approach. Selkirk (1982) did not apply the notion of initial [st-] to the same segment sequences in coda position because [-st] in the coda is well-formed according to the SSP[3] and is phonologically different from the initial [st-], but [kst] in [nekst] is still against SSP. However, the Chinese syllable structure cannot be fully parsed in OR models. Yip (2003) presents a critique of OR models in her discussion of the status of prenuclear glides in Chinese syllable structure. This issue will also be discussed in detail in chapter 4.

1.3.2 Dependency and Government Phonology

The fundamental contribution of Dependency Phonology is the claim that phonological structure involves head/dependency relations at all levels of organization (including intrasegmental organization), just as notions like government and headedness play a role in syntax and morphology (Anderson & Ewen 1987). This is referred to as the Structural Analogy Hypothesis. Government Phonology (GP) takes a very similar perspective regarding both this hypothesis and the organization of phonology proper (Kaye, Lowenstamm & Vergnaud 1985, 1990; van der Hulst & Ritter 1999). The GP approach is different from the OR approach, because in GP syllable structure is represented in the form of an X-bar syntactic structure, with the nucleus (X′) as a projection of V(X) (its head

[3] SSP, short for the Sonority Sequencing Principle (Hooper 1976; Kiparsky 1979; Clements 1990, among others), says that sonority should increase monotonically the closer one gets to the sonority peak, i.e. the vowel.

position), and the rhyme as the second projection (X′′) of a nuclear head position; if an onset head position is C, then the onset category is C′. In GP, there is no single constituent which corresponds to the traditional (OR) notion of 'coda'. The different structures that GP recognizes are illustrated in (3):[4]

(3)


The structure in (3a) shows some kind of X-bar-like structure disregarding 'coda' as a constituent, which presents the same notion of GP in (3b), both treating the coda position as a rhymal adjunct directly dominated by the rhymal node and the rhyme as a projection of the nucleus. The axiom of binary branchingness adopted in GP also allows just a limited array of options for the onset which is treated as a constituent in GP, as shown in (3c) and (3d). The direction of the government (head/dependent) relation is universally left to right (left-headed) within the syllable of constituent domain, as shown in (3b) and (3d).

GP captures the special character of the onset-rhyme relation by postulating a 'government relation', which means the nucleus is the head of the rhyme and their relations are manifested in terms of a licensing mechanism, which serves to authorize the units that comprise phonological representations. For the details of the licensing principles in GP, see Kaye, Lowenstamm and Vergnaud (1985), Anderson & Ewen (1987) and Harris (1994).

GP has been applied to a number of languages, clearly stating where a vowel should be inserted between consonants and where consonant gemination is possible and well licensed. Words such as *next* [nekst] may remain an unsolved problem. 'Superheavy' syllables with rhymes such as VVC, VCC, VVCC, or VCCC are all excluded in the government

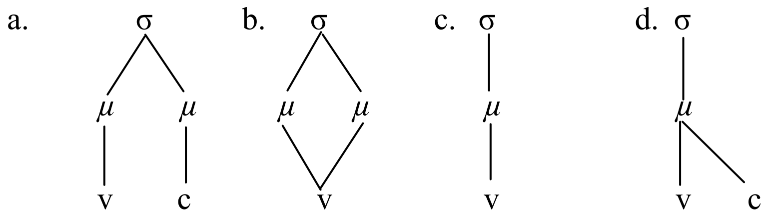
---

[4] In (3a), X′′ refers to a maximal projection (i.e. a constituent) in X-bar structure, in which Y is not a constituent but a terminal slot; in (3b), R refers to Rhyme in the syllable structure, sharing the same constituent structure as (3a). In (3c), C′ refers to the onset category when Consonant (C) is the head of an onset. In this case, an onset may have a second consonant as a compliment of Onset (O), as shown in (3d).

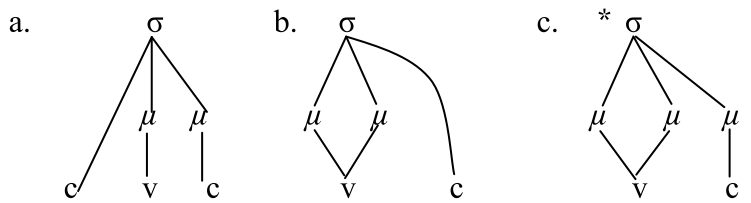
phonology approach, so that Harris (1994) introduced the concept of 'degenerate syllable', which means that an audible nucleus is absent after a word-final onset.

1.3.3 Moraic models

Another important and influential syllable theory is the mora model, in which, basically, some x-slots (in the nucleus and/or coda) but not others (in the onset) are mora positions (referred to as 'weight units' by Hyman 1985). In moraic theory, all vowels are moraic; a short vowel is monomoraic and a long vowel is bimoraic. Onset consonants are always non-moraic but a coda consonant can be assigned a mora according to the rule of Weight-by-Position (Hayes 1989). It is often argued that syllables are maximally bimoraic in moraic models (Hyman 1985; Davis 1999). As a result, the following structures are allowed for a moraic ("heavy") rhyme, a long vowel, a short vowel and a monomoraic rhyme, respectively:

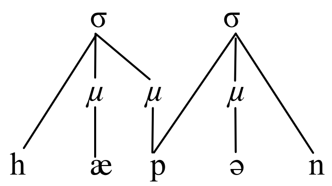
(4)


Any non-moraic consonant is directly attached to the syllable node, as in the case of an onset consonant and a consonant after a long vowel:

(5)


Davis (1999) claims that geminate consonants are inherently moraic and long vowels are bimoraic, so that true geminates can only be preceded by a short vowel. If the medial consonant in English word *happen* is regarded as a geminate, this explains why the vowel preceding this consonant must be short:

(6)



The moraic approach has been claimed to be superior to other syllable theories in some aspects (see van der Hulst & Ritter 1999 for discussion):

(a) Moras are better integrated into the prosodic hierarchy;
(b) The model expresses the weight-irrelevance of the onset;
(c) It expresses the variable nature of coda-weight;
(d) It offers an account of short vs. long vowels and singletons vs. geminates;
(e) It offers a way of expressing light, heavy and superheavy syllables.

However, problematic areas also exist in the moraic approach. The GP theory is better suited to account for phonotactics and to express the phonotactic independence of the onset with respect to its irrelevance for the prosodic hierarchy. Mora theory, since it does not recognize the onset as a constituent, cannot express the notion of onset well-formedness in a straightforward manner in that all onset consonants are non-moraic and directly connected to syllable node, ignoring the sonority difference between them. Another problem concerns the representation of glides, which is carried out exclusively by reference to the moraic level: vowels and glides are the same segment, the former being moraic and the latter being non-moraic (note that in Hyman's approach, by contrast, glides must be marked [–consonantal] at the melodic level) (Roca 1994). In this mirror, [wu] could be in fact a combination of *[uu] (differing in non-moraic and moraic status), which violates the OCP[5].

1.4 Tone Theory

It is estimated that about 60 to 70% of the world's languages are tone languages (Yip 2002). Theories of tone, however, have not developed to such a degree as theories of segmental or syllable structure. However, in

[5] Obligatory Contour Principle (OCP) (Roca 1994; Goldsmith 1990; Kenstrowicz 1994) will be discussed in chapter 2.

the past decades linguists have begun to pay more attention to tones with respect to their function, their phonetic realization and phonological behaviour, and the tonal structure of different languages. In both tone and non-tone languages, substituting a phoneme in a syllable or a lexical item may be responsible for a change in lexical meaning. In a tonal language, changing the tone on a syllable has the same effect. This is illustrated by the following examples from Shaoxing:

(7)	[tuŋ^{52}]	'east'	[duŋ^{31}]	'same'
	[tuŋ^{35}]	'understand'	[duŋ^{13}]	'pail'
	[tuŋ^{33}]	'freeze'	[duŋ^{22}]	'cave'

The examples in (7) show that tone plays as important a role in the representation of a lexical unit. The difference between phonemes can be described in terms of the distinctive features of which each phoneme consists. Tones can also be analysed in terms of features, which makes it possible that tones, like segments, can undergo processes like assimilation or dissimilation. There are two main issues concerning the internal structure of tone: the representation of tone features and that of register features.

1.4.1 Tone features

There are different traditions for marking tones in different languages. Chao (1928) marked the tone pitches of Chinese in a numerical system and divided them into five levels, with 1 as the lowest and 5 as the highest, as indicated in the examples in (7). Five seems to be the maximum number of levels used by any language (Yip 2002). Each syllable is marked by zero to three digits. Most syllables are given two digits (as in (7)), one for the starting pitch and one for the ending pitch. This is true even for level tones, unless the syllable is very short, in which case only one digit is usually used. Generally, there are at most three level tones in a language such as [55], [33] or [11], which are usually marked as H (High), M (Mid) and L (Low), respectively. [35] is a rising tone and [52] and [31] are falling tones; any non-level tone is referred to as a contour tone. A level tone is simpler than a contour, so that if a language has a contour tone, it must also have a level tone but the converse is not true. Three digits are used for tones which change direction in the middle of the syllable, such as [314], which are referred to as concave. A tone with a rise and a fall (e.g. [351]) is referred to as convex. As a rule of thumb, any contour with a two-digit difference between starting and ending points, such as 13 or 53, is probably phonologically a contour, but tones with only a one-digit difference, like 21 or 45, should be analysed with a degree of caution (Yip 2002). It has been argued in the literature that contour tones are sequences of level tones (Leben 1973; Yip 1989, 2002; Chen 2000).

The unit that bears the tone is called the tone-bearing unit (TBU). It is usually accepted that TBU cannot be the onset, which is weightless in the syllable and therefore non-moraic (Yip 1980; Pulleyblank 1986; Bao 1990). In some cases it is hard to decide what the exact TBU is in a particular language; it could be the vowel, the mora or the syllable in different languages, as Yip (2002: 73) illustrates (where T refers to a tone or set of tone features):

```
(8)  C V — T        or  CV             or  CV
       |                 |                  |
       μ                 μ — T              μ
       |                 |                  |
       σ                 σ                  σ — T
```

The minimal constituent which contains the TBU is the tone domain. We assume that the tone domain can be different from language to language. If a tone language has a moraic coda, the tone domain can be the rhyme constituent; if a tone language has a non-moraic coda, the tone domain can be the nucleus constituent. Whether tones are associated to a vowel, mora or syllable, tonal features can spread, assimilating or dissimilating, depending on the circumstances. Tonal assimilation or dissimilation is referred to as tone sandhi, which is a common phenomenon in Chinese languages. For example, when two rising tones occur in a lexical phrase in Shaoxing, the first rising tone is dissimilated to a level tone through a dissimilation rule, as illustrated in the example in (9):

(9) [tsɤ^{35}tsɤ^{35}] → [tsɤ^{33}tsɤ^{35}]
walk walk
'Take a walk.'
Rule: [lh] →[l] / __ [lh]

There are many different theories with respect to tone features, the behaviour of tone, their characteristics, their internal structure, and their phonological realizations, such as Halle & Stevens (1971), Yip (1980, 1989, 2002), Clements (1983) and Hyman (1986, 1993), Pulleyblank 1986), (Bao 1990, 1999), and Snider (1999), and these theories form a topic of controversy.

1.4.2 Register features

The relation between laryngeal features and the voice quality of consonants is a topic which has also inspired some debate in the literature (Halle & Stevens 1971; Bao 1999; Lombardi 1994; Snider 1999; Jansen 2004). One proposal holds that there are two laryngeal features: [stiff vocal cords] and [slack vocal cords] (Halle & Stevens 1971): [+stiff] specifies voicelessness in consonants and relatively high pitch in vowels, and [+slack] specifies voicing in consonants and relatively low pitch in vowels (Bao 1999). This captures directly the correlation between tone registers (H or L) and the voice qualities of obstruents (recall the data in (7), where this same relation is seen and which will be explored in more detail below). Yip (1980, 1989) formalizes this relation between vocal fold tension and voicelessness in obstruents on the one hand, and pitch in tone on the other by proposing that tone is not an indivisible entity in phonological representation. Rather, it consists of two parts, 'register' and 'tone', in which the two features play different roles. 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) (marked by *l* and *h*), as illustrated in the geometrical structure below (Yip 1989):

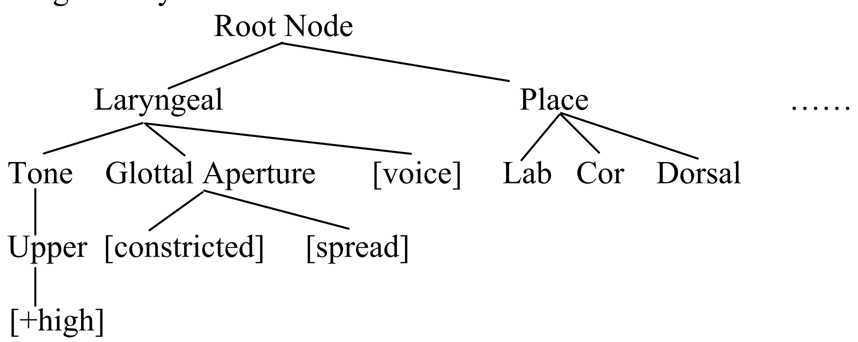
(10)
σ
|
H   Tonal Node   *register*
/\
l *h*   *tone*

Feature geometry provides the means to deal with some generalizations about laryngeal phonology that were impossible to capture in previous systems. In Yip's feature geometry in (10), Tone[6] features are differently represented compared to a model using H, M, and L features (Woo 1969; Clements 1985; Chen 2000) as mentioned above. In the

[6] The capitalized Tone refers to the whole tone feature structure, including register features and tone features.

feature geometry in (10), tones are dominated by [±Upper], which takes two values: high register [+Upper] or low register [–Upper]. In the low register, the highest pitch is maximally [3]. Thus if a tone [3] occurs in a L (low) register, the tone is *h* (high); if it occurs in a H (high) register, it is *l* (low). For example, the tone features of [35] in a H register are [H.lh]; and [31] in a L register is [L.hl].[7] The register feature [+Upper] (or H/L) is believed to be dominated, in turn, by laryngeal features. Yip (2002: 60) incorporates the Tone features in the feature geometry, as represented in (11):

(11) The geometry structure:



The geometry structure in (11) shows that tonal pitch is dominated by Upper (register); the tone features themselves are dominated by the laryngeal node. Just like segmental features, the tone feature(s) can spread, delete, assimilate or dissimilate since they are located within the segment domain in the geometry. This representation raises the question whether the TBU is necessarily the segment. However, some general points can be made regarding the behaviour of tone features. Yip (2002) lists the following properties:

(a) Mobility: Movement away from the original tone-bearing unit.[8]
(b) Stability: Survival after loss of original host segment.
(c) One-to-many: A single tonal feature shared by two or more segments.

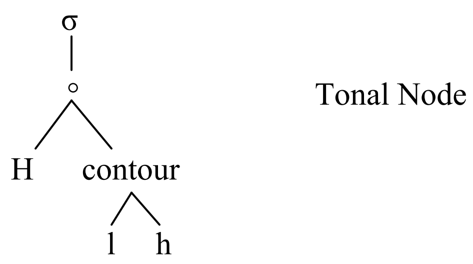
---

[7] H/L are used for the register feature; l/h for tone feature in Yip's model (2002). Thus, for example, [H.lh] represents a rising tone in a high register; while [L.hl] represents a falling tone in a low register.

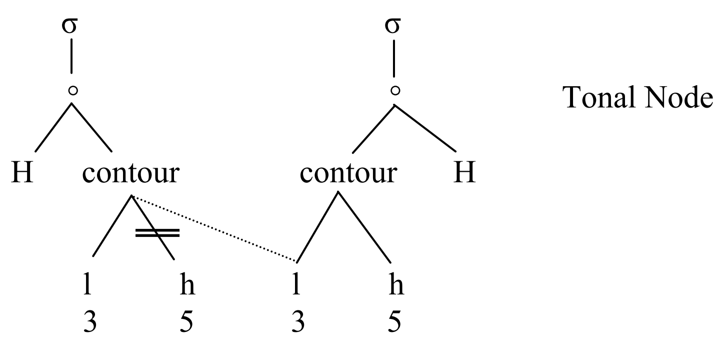
[8] It is claimed that tone(s) can move away from the original TBU to an adjacent syllable progressively or regressively, as in the case of tone sandhi in Chinese (Yip 1980, 2002; Chen 2000, among others).

(d) Many-to-one: Multiple tonal features surfacing on a single host segment.
(e) Toneless segments: Potentially tone-bearing segments that never acquire phonological tone.

Although tone features are in one way or another similar to segmenal features, Yip's Tone theory offers a systematic, explanatory approach to Chinese tones. One disadvantage of Yip's tone model in (10) and (11) may be that it predicts that when a register feature spreads, it carries tone feature(s) with it, so that they spread together. The data of some languages, e.g. Chaozhou (Bao 1999), Shaoxing (Yang & Yang 2000), etc. suggest that register feature can spread in isolation, i.e. without carrying the tone feature along. For this reason, Bao (1999) proposes a different model to account for tone data in the Min language group, which can express register spreading without tone spreading, as shown in (12):

(12)


In (12), the register feature (H) and contour (tone) feature are in a sister relationship, so that register and tone can spread independently, or one tone feature can spread but the other does not, as in tone contour sandhi. For example, the tone sandhi rule in (9) can be expressed as follows:

(13)


The illustration in (13) expresses the rule in (9) and shows that it is a possible tone rule. In chapter 5 we will discuss the Shaoxing data with respect to tone in greater detail.

1.5 Data

Most of my data come from spoken Shaoxing, of which I am a native speaker. Some of the data were collected or used by other researchers or scholars, such as Chao (1928), Wang (1959), Chen (1982), Zhan (1991), and Yang & Yang (2000). Some of the data come from my own observations. I have also collected data from other native speakers who were deemed (by me) to have undergone little influence from Mandarin or other dialects in their pronunciation of their native language. Data from books were always checked against my own intuition and I omitted any data that did not comply with my intuitions.

Data from all other languages referred to in this dissertation are from books or journals or native speakers of the languages in question. Transcriptions follow the IPA system and are, if necessary, in narrow transcription. Tones are marked by numbers after each syllable; all tones are citation tones, except when discussing tone sandhi in chapter 5.

1.6 Organization

This dissertation presents an analysis of the overall phonology of Shaoxing Chinese from a synchronic perspective, with the secondary goal of casting some light on current issues in Modern Chinese (Mandarin).

Chapter 1 gives a general background, including a brief introduction to the historical profile of Shaoxing, a brief idea of the methodology adopted, some theoretical approaches applied in my dissertation, including syllable theory and tonal theory, and an overview of the data used in the dissertation.

Chapter 2 presents a general introduction to the consonants and vowels of Shaoxing in surface representation, providing an analytic description of its stops, affricates, fricatives, sonorant consonants, vowels, and all the possible forms of its complex “finals” (see ch. 2). I assume that the surface nasal vowels in Shaoxing are phonological sequences of vowel + nasal. In this chapter, I also offer a discussion of the phonological properties of all consonants in the initial position in Shaoxing and their distribution, which will make it possible to propose some constraints on phonotactics of Shaoxing. I make the claim that there are no complex onsets, no vowel-glide (VG) combinations, and no diphthongs in Shaoxing.

Chapter 3 sketches the distribution of the 14 surface vowels of Shaoxing through a description of different phonetic and phonological environments where complementary distribution occurs, with the aim of establishing the underlying vowel inventory of Shaoxing, which has only six phonemic vowels: /i u e ɤ o a/. In this chapter, I present an analysis from the perspective of Optimality Theory and propose a suitable constraint ranking for the relations between the 14 surface vowels and six phonemic vowels. I assume that the underlying forms of the three nasalized vowels [ẽ], [ɛ̃] and [ɵ̃] in Shaoxing are /en/, /an/ and /on/, respectively.

Chapter 4 presents a comparative analysis of the different approaches to the Chinese syllable structure, focusing on the prenuclear glide in the syllable of Shaoxing and Mandarin. I assume that the phonetic onsets [ʔ] and [ɦ] serve as 'tone assigners' when there is no other onset consonant and that there is a possible coda in Shaoxing syllable structure, which only allows [ʔ] and [ŋ]. I present a multiple-specifier X-bar syllable structure, assuming that the prenuclear glides are neither in the onset nor in the rhyme, but act like the specifier of N'', suggesting that N'' is a sub-constituent dominating N' (Rhyme). I also present a cross-tabulation of all the possible combinations of the Shaoxing syllables.

Chapter 5 gives an analysis of the tone system of Shaoxing, focusing on consonant-tone interactions. I assume that the voiced/low tone correlation is related to historical tonogenesis in Shaoxing and both voiceless and voiced initial obstruents and high and low tones occur in underlying forms. I formalize the complexity of tone sandhi in Shaoxing, assuming that tone sandhi is phonologically realized by way of tone feature spreading and delinking, and motivated by metrical mechanisms. I also present an OT analysis of all possible disyllabic tone sandhi rules in Shaoxing and offer a constraint ranking which explicitly accounts for all the disyllabic tone sandhi rules in Shaoxing.

Chapter 6 gives a summary of my dissertation and proposes topics for future studies.

2 The Consonants and Vowels of Shaoxing: Surface Representations

2.1 Introduction

This chapter provides an analytic description of the surface consonants and vowels of Shaoxing Chinese (henceforth SX). Before analyzing the distinct phonemes and their phonological features in chapter 3, I will describe the overall numbers and forms of possible initials and finals of SX. I use the terms "initials" and "finals", which correspond to "shengmu" and "yunmu" in traditional Chinese accounts of syllable structure. Wang & Smith (1997: 7) introduce the traditional way of representing Chinese syllable structure, given here in the tree diagram in (1a):

(1) a. syllable

shengmu yunmu

yuntou yunfu yunwei

b. σ

Initial Final

Glide Nucleus Coda

The diagram in (1a) shows that the Chinese syllable structure has two main constituents: "shengmu", corresponding to what current phonological theory would call onset, and "yunmu" (which literally means 'rhyme'), referring to all that follows "shengmu". "Yunmu", in turn, consists of three parts: "yuntou", which is the position reserved for prenuclear glides, "yunfu", which is the nucleus, and "yunwei", which is the coda. The diagram in (1b) expresses the traditional representation of Chinese syllable structure in Western terminology. In traditional Chinese phonology,[1] "yunmu" in SX (as well as in Mandarin and all other Chinese dialects) is not equivalent to the syllable rhyme, because the Chinese "yunmu" does not count in the poetic rhyming system, which only includes "yunfu" and "yunwei", i.e. it excludes "yuntou". Thus, we cannot adopt the term "rhyme" for "yunmu" in the Chinese syllable terminology.

[1] Traditional Chinese phonology can be traced back to as early as the 'Qieyun' rhyme table (AD 601), which deals with the pronunciation of Ancient Chinese (Karlgren 1915-1926; Wang 1963; Chao 1968; Xue 1986; Yang 1996).

To avoid confusion, many modern Chinese linguists (Chan 1997; Wang & Smith 1997; Chen 2000; and many others) adopt the terms "initials" and "finals" for the traditional terms "shengmu" and "yunmu", instead of general phonological terms such as "onset" and "rhyme". In Chapter 4, I will present an analysis of SX syllable structure and also seek to cast some light on the syllable structure of Mandarin.

In this chapter, I also use the terms of "Initials" and "finals", referring to "shengmu" and "yunmu", i.e. the constituents into which the SX syllables can be divided and within which all the surface consonants and vowels of SX may occur.

As was mentioned in chapter 1, both the constraint-based theory (e.g. Prince & Smolensky 1993) and the rule-based theory (e.g. Chomsky & Halle 1968) share the notion of an underlying form, or input, and produce outputs, either having the advantage of the other in explaining the language, so that both theories are applied to the analysis in this chapter.

2.2 Initials

I will refer to syllable-initial consonants as "initials" throughout this chapter. The most remarkable characteristic of the SX initial consonants (compared with Mandarin and other Chinese dialects) is the fact that SX still retains the historical voiced obstruents, just as Middle Chinese did[2] (Chao 1928; Yip 1980; Zhan 1991). In the following sub-sections, I will discuss all five classes of possible SX initials: oral stops, the glottal stop, affricates, fricatives, and sonorant consonants.

2.2.1 Stops

In this sub-section, I will discuss voiceless unaspirated stops, voiceless aspirated stops and voiced unaspirated stops, all of which can appear as distinctive initials in SX. I list the nine stops of SX in (2):

(2)

	stops		
voiceless unaspirated	p	t	k
voiceless aspirated	p^h	t^h	k^h
voiced unaspirated	b	d	g

[2] Middle Chinese is the language of the Sui, Tang, and Song dynasties (7^{th}–10^{th} centuries AD).

SX has voiceless unaspirated, voiceless aspirated and voiced unaspirated bilabial, alveolar and velar stops. The voiced stops are typical of the Wu dialects. Modern Chinese (Mandarin) and the other five Chinese language families have lost all the voiced stops and voiced affricates (Campbell 2003). SX has retained all the voiced stops and affricates that were present in Middle Chinese. The nine stops in SX are very commonly used as the onset of the syllable. Some examples are presented in (3):

	a		b		c	
(3)	[pɪŋ^{33}]	'soldier'	[p^{h}ɪŋ^{33}]	'marry'	[bɪŋ^{22}]	'ill'
	[tɪŋ^{33}]	'book'	[t^{h}ɪŋ^{33}]	'listen'	[dɪŋ^{22}]	'decide'
	[koŋ^{33}]	'supply'	[k^{h}oŋ^{33}]	'free time'	[goŋ^{22}]	'total'

In the examples above, the syllables in column a all have voiceless unaspirated stops; those in column b have voiceless aspirated stops; column c has voiced unaspirated stops. These examples show that the tones of the words in column c are lower[3] than those of the words in columns a and b. In Middle Chinese tonogenesis, the syllables with voiced obstruents in the initial had a lower tone than those with voiceless obstruents (Chao 1928; Yip 1980, 1989; Bao 1999; Duanmu 2000b; and many others), which is consistent with the articulatory and acoustic properties of voiced and voiceless initial obstruents on the one hand and pitch on the other (see also below). The current tonal structure of SX therefore still reflects the tonal system of Middle Chinese. The Chinese tones have been philologically and linguistically classified as "*yin*" and "*yang*" registers since the 7th century (Bao 1999). The *yin* register is also referred to as high register, corresponding to high-pitched tones, while the *yang* register is also referred to as low register, corresponding to low-pitched tones. In SX, there are altogether eight tones. Among these tones, [52], [35], [33] and [5] are in the *yin* register, which mainly occur in voiceless-obstruent-initial syllables, and [31], [13], [22] and [3] are in the *yang* register, which occur only in voiced-initial syllables, including voiced obstruents and sonorants.

In SX, we can predict from the tone of the syllable whether the initial obstruent is voiced or voiceless. The reverse, however, is not the case,

[3] The pitch of the different tones is marked on a five-point pitch scale, in the same way as in Mandarin. The highest pitch is marked 5 and the lowest is marked 1 (cf. also ch. 1). See chapter 5 for a further exploration and discussion of SX tones.

because there are four low-register tones and four high-register tones with voiced-initial syllables or voiceless-initial syllables, respectively (see examples in (5) below). We can only tell from the voiced or voiceless initials that the tone of the syllable falls within the *yang* (low) register or *yin* (high) register – we cannot predict the individual tone. Moreover, it is completely unpredictable whether voiceless stops in *yin*-register syllables are aspirated or unaspirated, as the examples in (4) show:

(4)	[pu^{33}]	'cloth'	[p^{h}u^{33}]	'berth'
	[tɪŋ^{33}]	'book'	[t^{h}ɪŋ^{33}]	'listen'
	[kɛʔ^{5}]	'cut'	[k^{h}ɛʔ^{5}]	'block'
but	[bu^{22}]	'step'	*[bu^{33}]	
	[dɪŋ^{22}]	'decide'	*[dɪŋ^{33}	
	[gɛʔ^{3}]	'squeeze'	*[gɛʔ^{5}]	

Since the syllables with a voiced stop in the initial have a lower tone than those with a voiceless stop, [b], [d] and [g] cannot form exact minimal pairs with syllables with voiceless aspirated or unaspirated initial stops, as is shown in (4). Thus, the question arises whether the voiceless stops and voiced stops in SX are allophones of each other. I claim that the voiceless unaspirated and aspirated stops and voiced unaspirated stops in SX are not allophones of each other, i.e. that they are distinctive phonemes. There are two main points that bear on this issue. First, allophones are variants of a distinctive phoneme which are usually in complementary distribution. For two phones to be classified as allophones of a single phoneme, they must exhibit phonetic similarity, and they must not be in contrastive distribution (e.g. Trask 1996). For example, [k] and [k^{h}] are allophones of the same phoneme /k/ in English.

In SX, however, the tones on the syllable are not predictable from the syllable-initial consonant. There are four tones that can occur in each case. Take the four tones of a syllable with a voiced initial stop in SX. Rhymes ending in a glottal stop have tone [3]. The other three tones, [22], [13] and [31], are unpredictable. For example:

(5)	a. [bɪŋ^{22}]	'ill'	[bɪŋ^{13}]	'judge'	[bɪŋ^{31}]	'level'
	b. [doŋ^{22}]	'cave'	[doŋ^{13}]	'pail'	[doŋ^{31}]	'same'
	c. [gɑɒ^{22}]	'a tool'	[gɑɒ^{13}]	'confuse'	[gɑɒ^{31}]	'finish'

The examples in (5) show that the tone pattern cannot be predicted from the initial. However, the voicing quality of the initial consonant could still be predictable from the tone registers. For articulatory and acoustic reasons, a syllable beginning with a voiced obstruent intrinsically has a lower tone than one beginning with a voiceless obstruent (Haudricourt 1954; Lehiste 1970; Matisoff 1973). This relation between voiceless obstruents and high register, and between voiced obstruents and low register, is widely attested in natural languages (Yip 2002). The consonant-tone correlation is also well documented in the tonogenesis literature cross-linguistically (Hombert 1978; Hyman 1978; Hombert, Ohala & Ewan 1979).

From an acoustic point of view, if the initials of the syllables in column a and column c or those in column b and column c in (3) were in complementary distribution, the tones on these syllables could be analyzed as 'allotones', rather than allophones. Yip (1980: 138) also realizes that in such cases the voiced stops are always accompanied by low-tone allotones. I will assume that tones in high register and those in low register occur both in underlying forms in SX. However, the exact distribution of tones and their relations with initial consonants will be left for more detailed discussion in chapter 5.

2.2.2 Glottal stop

Besides the bilabial, alveolar and velar stops discussed above, there is another stop in SX, viz. the glottal stop [ʔ]. Although about 46% of the world's languages have a glottal stop, according to Maddieson (1984a), in many languages glottal stops serve to demarcate the boundaries of phrases or other prosodic units (Ladefoged & Maddieson 1996). However, in some languages like Arabic, Thai, Hebrew and Hawaiian, the glottal stop is a contrastive consonant. In (6) some examples from Hawaiian are given; the form in (a) comes from Ladefoged & Maddieson (1996) and those in (b) and (c) from Gussenhoven & Jacobs (1998):

(6) a. [ʔaʔa] 'dare'
b. [aa] 'jaw'
c. [ʔaa] 'fiery'

The examples in (6) suggest that the glottal stop in Hawaiian is a phonemic stop, although perhaps the issue is not decided conclusively, lacking more data and a formal analysis. In Middle Chinese, the glottal stop [ʔ] served as an initial and was used before [u], [a], [i] and any of the three glides, [j], [w] and [ɥ][4], when there was no other consonantal onset in the syllable (Chao 1928). The glottal stop in SX seems to play the same role in the syllable as it did in Middle Chinese. Some examples from SX are listed in (7):

(7)	[ʔe^{33}]	'love'	[ʔɛʔ^{5}]	'duck'
	[ʔjẽ^{52}]	'smoke'	[ʔɥə̃^{52}]	'complain'

The examples in (7) show that in SX [ʔ] can serve as an onset preceding a vowel or glide in the syllable. Here the question arises whether the glottal stop in SX is a phonemic stop, as in Hawaiian, or only a phonetic form in words where no other consonant is present, as in German (see (8)) or Dutch.

The status of the initial glottal stop in SX as well as in all other Wu dialects has been a topic of debate, also with respect to Middle Chinese (Karlgren 1915–1926; Chao 1928; Dragunov 1930; Hope 1953; and many others). Dragunov (1930, following Karlgren) states that "in the ancient Chinese languages, as is well known, words with zero initials were differentiated from words beginning with a glottal stop. We discover this distinction again, in certain Northern Chinese dialects of the 13th and 14th centuries, as is shown by the hPhags-pa script, and in nearly all the modern Wu dialects" (cited by Hope 1953).[5] Hope (1953: 2) strongly argues against the existence of phonemic initial glottal stop in Middle Chinese and argues that "Karlgren presents little or no evidence; all he really does is to make an assumption for the purpose of filling a psychologically created lacuna". Hope (1953: 13) further claims that "in all modern Chinese dialects without exception the glottal stop, where it exists, is not only of no phonemic significance but is not even heard by the speakers of the language". It is true that Karlgren (1915–1926) only reconstructed an initial glottal stop for ancient Chinese (because the phonetics of Middle Chinese are uncertain). Chao (1928) did fieldwork on

---

[4] In traditional Chinese phonology, the three glides were transcribed as [i], [u] and [y], i.e. using the same symbols as the counterpart vowels (Chao 1928; Wang 1963, 2003, among many others). Through my dissertation, I transcribe them as [j], [w] and [ɥ], respectively, following the general linguistic transcription. I will claim that the prenuclear glides in SX are not in the onset. This will be discussed in detail in chapter 4.

[5] "HPhags-pa" or the "hPhags-pa script" is the language created by hPhags-pa (1235–1280) in the Yuan Dynasty (1207–1367). It is sometimes referred to as New Mongolian, which had over 40 letters used to spell Mongolian, Chinese and Tibetan.

the phonetic realization of the initial glottal stop in modern Wu dialects, and states (in his Table 3) that SX has a stable [ʔ] as an initial preceding [u], [e], [jẽ] and [ɥõ]. I also claim that the initial glottal stop in SX exists, but its status is phonetic rather than phonological. There are two main reasons for this: First, in SX there is no minimal pair which would show that [ʔ] is a contrastive phoneme as was the case for Hawaiian in (6). The syllables in (7) can optionally be pronounced without the initial glottal stop and with the same tones and would then have the same lexical meaning: thus, the glottal stop plays no contrastive role.

It is true that syllables in SX are strictly classified into two categories: high-register syllables with high-pitch tones and low-register syllables with low-pitch tones, which is determined by the voicing status of the initial obstruents (c.f. above). As expected, the glottal stop [ʔ] as the initial of a syllable is accompanied by a high-pitch vowel so that the initial glottal stop [ʔ] is predictable in SX. However, underlyingly, there is no initial glottal stop for the syllables in (7). A similar phonetic phenomenon occurs in German as Hall (1992:58) describes "the glottal stop is completely predictable in German, since it only occurs before a vowel-initial stressed syllable and then only optionally". For example (Hall 1992:58):

(8)	arm	[ʔáʀm]	[áʀm]	'poor'
	elf	[ʔɛ́lf]	[ɛ́lf]	'eleven'
	oft	[ʔɔ́ft]	[ɔ́ft]	'often'
	Uhr	[ʔú:ʌ̯]	[ú:ʌ̯]	'clock'

Hall formulates this phenomenon in German by way of a Glottal Stop Insertion rule, which is presented in (9):

(9) Glottal Stop Insertion (GSI: optional)
$\emptyset \rightarrow$ [ʔ] / $_{F}$[__ [−cons]

Jongenburger & van Heuven (1991) carried out a phonetic experiment in Dutch with the example of [dɑt ən ʔɑ:ntɑl] 'that a number' and found that there is a striking acoustic difference with a smooth vowel onset in [ən] and an abrupt onset (glottal stop) in [ʔɑ:ntɑl]. They claim that the phonetic glottal stop is inserted only in hiatus position, and nowhere else in Dutch (Jongenburger & van Heuven 1991, 1993). The only difference between German and Dutch and SX is that the GSI rule in

German cannot apply to the syllable which is not foot-initial, while in Dutch it only operates in hiatus position whilst GSI in SX applies to every vowel-initial syllable with a high-register tone, due to the fact that SX is a monosyllabic language. However, neither the initial glottal stop in German or Dutch nor that in SX is an underlying phoneme. To briefly summarize, SX, like many other languages, inserts a glottal stop in syllables beginning with a vowel. Does this mean that there is no phonemic /ʔ/ in SX? I assume that there *is* a phonemic glottal stop /ʔ/ in SX, but it is in the final rather than in the initial. I will discuss the phonological behavior of final /ʔ/ in SX in §2.4.

2.2.3 Affricates

Phonetically, affricates are consonants whose articulation involves a complete oral closure followed by a comparatively slow release, yielding perceptible friction noise which is clearly longer than the noise burst of a plosive. Affricates are not uncommon phonemes across languages (Maddieson 1984a) in terms of their occurrence in languages. In SX, there are six affricates, including three dental affricates and three alveolo-palatal affricates, as shown in (10):

(10)		dental	alveolo-palatal
	voiceless unaspirated	ts	tɕ
	voiceless aspirated	ts^h	$tɕ^h$
	voiced unaspirated	dz	dʑ

Like the SX stops in (2), there are voiceless unaspirated, voiceless aspirated and historically voiced unaspirated affricates in SX[6]. Phonetically and phonologically, these affricates are phonemic consonants, although voiced affricates and voiceless affricates are accompanied by different tones in the syllable, because of the association between stiff and slack vocal cords and high and low tone register, respectively, as was discussed above. For example:

[6] Historically speaking, Middle Chinese had voiceless unaspirated, voiceless aspirated and voiced unaspirated stops and affricates (Karlgren 1954; Wang 1985; Baxter 1992), which are claimed to still retain in the Wu Chinese, like SX (Wang 1959; Zhan 1991; Cao 2002) in spite of the fact that some phonetic experiments show that the historical voiced stops and affricates in the Wu Chinese do not have VOT (Shryock 1995).

(11)	[tsoŋ[52]]	'end'	[tɕi[33]]	'post'
	[tsʰoŋ[52]]	'dash'	[tɕʰi[33]]	'air'
	[dzoŋ[31]]	'worm'	[dʑi[31]]	'ride'

Middle Chinese had nine affricates in initial position, including [ts], [tsʰ], [dz], [tʃ], [tʃʰ], [dʒ], [tɕ], [tɕʰ] and [dʑ], according to the *Guangyun*[7] (Ding 1984). Of these, SX has retained six. SX does not distinguish between alveolar and palatal affricates, lacking post-alveolar and retroflex affricates such as /tʃ/, /tʃʰ/, /dʒ/, /tʂ/, /tʂʰ/, and /dʐ/.

The status of affricates as one or two segments has been a topic of some debate. Some have argued that affricates are combinations of stops and fricatives and should be treated as two segments (Brooks 1965; Szigetvári 1997; and others). Durand (1990) cites some examples of segment sequencing in English, which show that in English either /ʃ/ or /t/ can be followed by /r/; but /tʃ/ cannot. For example:

(12) a. [ʃrimp] 'shrimp'
b. [trai] 'try'
c. *[tʃr-]

The examples in (12) show that (c) is ruled out, possibly because there are already two consonants before /r/ and the maximal onset cluster is CC (except [s]+CC) in English. This kind of evidence can be used to suggest that affricates form a consonant cluster in a particular language. Harris (1994) claims that affricates are qualitatively complex but quantitatively simple. Theoretically speaking, affricates are combinations of the features of the two constituent phonemes rather than two separate phonemes (see also Clements 1985; Anderson & Ewen 1987; McCarthy 1988; van de Weijer 1996 for discussion).

Evidence from SX strongly suggests that affricates are single segments. SX is a language that has no complex onsets.[8] No combinations of two consonants preceding a glide or a vowel are possible in SX (*[sp], *[sh], *[sl], *[pl], *[kl], *[sn], *[ph], etc). The SX syllable structure maximally has a CGVC, or CGVV pattern (where C is a consonant, G is a glide, and V is a vowel). For example:

---

[7] *Guangyun* (1008) is an ancient book of traditional Chinese phonology that describes the pronunciation of the Chinese syllables and rhyming system.

[8] It is also my claim that the prenuclear glides are not part of the onset in SX. This will be discussed in detail in chapter 4 about the syllable structure.

(13) [pjɑɒ^{35}] 'watch'
[ɕjɑɒ^{35}] 'small'
[kwɒŋ^{52}] 'light'
[ɦwoŋ^{31}] 'red'

There are no *CCV, *CCGV, or *CCGVC syllables in SX. However, affricates may fill the C position in the syllable template; this is in fact very common in SX. For example:

(14)	[dʑjɑɒ^{31}]	'bridge'	[tswo^{33}]	'do'
	[tɕ^{h}joʔ^{5}]	'lack'	[ts^{h}wo^{33}]	'wrong'
	[dʑjoŋ^{31}]	'poor'	[dzwo^{31}]	'tea'

The syllables in (14) are well-formed and acceptable. This is consistent with the assumption that affricates are single segments, such that the SX maximal syllable template of CGVC/CGVV can be maintained.

Another piece of evidence for the idea that affricates are single segments in this language comes from SX loanwords. SX does not allow complex onsets, as discussed above. This could be captured by an Optimality-theoretic constraint on the syllable onset in SX:

(15) *COMPLEX(ONS) (Itô 1986; Blevins 1995)
*[$_{\sigma}$ CC: Onsets must be simple.

The constraint *COMPLEX(ONS) is highly ranked in SX, so that complex onsets are not acceptable in SX, even in loanwords (Zhang 2003). For example, the English noun [k^{h}loun] 'clone' will be realized as [k^{h}ɤloŋ], without the original CC cluster in the onset; rather, an epenthetic vowel is inserted. SX has a very restricted coda. Only two consonants, [ŋ] and [ʔ] are allowed in the coda position, which can be stipulated by a coda-condition constraint, as stated below:

(16) CODA-COND (Itô 1989; Zhang 2003)
Coda can only be [ʔ] or [ŋ][9].

[9]Coda condition in SX will be discussed in the next section and in more detail in chapter 4.

Another important phenomenon in loanword phonology of SX is the tendency of disyllabification for a minimal word, which is also well documented in many other Chinese languages (Yip 1993; Chen 2000; Zhang (2003). There is a Minimal-word constraint as follows:

(17) MINWD (Yip 1993)
A loanword is minimally disyllabic.

Among the three constraints explained above, *COMPLEX(ONS) and CODA-COND are inviolable in SX while there are occasionally some exceptions for MINWD.[10] For example, the English *clone* and *tank* are [k^{h}ɤloŋ] and [t^{h}ɛ̃k^{h}ɤ], respectively, in SX loan words. The tableau deriving the optimal candidate is given in (18):

(18)

input k^{h}loun	*COMPLEX(ONS)	CODA-COND	MINWD
a. k^{h}loun	*!	*	*
b. k^{h}loŋ	*!		*
c. ☞ k^{h}ɤloŋ			
d. k^{h}oŋ			*!
e. loŋ			*!

In tableau (18), candidate (a) is the closest to the input, but violates all the three constraints, so it is the worst candidate. Candidate (b) is also ruled out first because it violates *COMPLEX(ONS). Candidates (d) and (e) violate MINWD and are also ruled out. Candidate (c) is the winner because it does not violate any constraints in (18) (it violates another constraint, DEP-IO, which militates against insertion of material, so this constraint must be lower-ranked).

If we compare a loanword with a cluster to a loanword with an initial affricate, it turns out that (English) affricates are well-formed onsets in SX. For example, English [dʒi:p] ‘jeep’ is adapted as [tɕɪʔpu] in SX. The selection of this candidate is represented in the following OT tableau:

[10] Some technical terms may be still monosyllabic, e.g. chemical elements. This involves more constraints, which falls outside of the scope of my discussion.

(19)

input /dʒi:p/	*COMPLEX(ONS)	CODA-COND	MINWD
a. tɕɪʔp		*!	*
b. ☞ tɕɪʔpu			
c. tɕɪʔ			*!

In tableau (19), candidates (a) and (c) are ruled out because they violate CODA-COND and/or MINWD. Candidate (b) does not violate any constraint and is the winner. This shows that affricates behave as single segments and are acceptable in SX phonology: a candidate with epenthesis would be treated as violate the constraint DEP-IO mentioned above. In short, for any foreign word with a CC or CCC cluster in the onset, the SX loanword phonology exceptionlessly has to insert a vowel between CC or two vowels between CCC of the output loanwords. More examples (disregarding tones) are given in (20):

(20)

English		Loanword in SX	
a. [sprɪŋ]	*spring*	[sɿbɪʔlɪŋ]	'spring lock'
b. [brændi]	*brandy*	[paʔlɛ̃di]	'Brandy wine'
c. [fræns]	*France*	[faʔlɛ̃ɕi]	'France'
d. [tʃɔkəlit]	*chocolate*	[tɕʰjaɒkʰəʔlɪʔ]	'chocolate'
e. [tʃə:tʃɪl]	*Churchill*	[tɕʰjɤtɕɪʔɬ]	(person's name)

In the examples in (20), vowels are inserted between the consonants in the initial clusters of the words in (a), (b) and (c) to form an acceptable SX syllable (other constraints in SX loanword phonology will be left aside until later: see chapter 4). In the words for *chocolate* and *Churchill*, however, the affricates are treated as single segments, so that no vowel needs to be inserted to break up affricates.

2.2.4 Fricatives

Fricatives are sounds that are produced by a narrow approximation of two articulators so as to produce a turbulent airstream (Ladefoged 1971: 46). While it is true that there are fricatives in nearly all of the world's languages, there have been relatively few studies of their precise distribution or of the patterns of occurrence which they show, according to Maddieson (1984a). However, some research has been done on the description of the articulation, classification of fricative sounds and their acoustic properties cross-linguistically (e.g. Jakobson, Fant & Halle 1952;

Nartey 1982; Ladefoged 1983). Nartey (1982) presents a cross-linguistic phonetic analysis of fricatives in 14 languages[11], measuring the phonetic differences within and between languages. In this sub-section, I will examine the behaviour of fricatives in SX. I provide the surface fricatives of SX in (21):

(21)	labiodental	dental	alveolo-palatal	glottal
voiceless	f	s	ɕ	h
voiced	v	z	ʑ	ɦ

SX has four pairs of voiceless and voiced fricatives. In most languages, there is a tendency to prefer voiceless fricatives and to avoid voiced and voiceless pairs of fricatives at the same place of articulation (Maddieson 1984a). Although this tendency holds for obstruents in general, fricatives appear to be more asymmetric. Most Southeast Asian languages have relatively few fricatives; Mandarin has five fricatives, which are all voiceless; Cantonese has four fricatives; Thai, Korean and Taba[12] (Abramson 1999) have only three fricatives. Indo-European languages usually have eight or more fricatives. The SX voiceless and voiced fricatives are contrastive phonemes in spite of their being accompanied by different tones in the syllable, which follows the same register division as with the stops and affricates (cf. above). For example:

(22)	voiceless		voiced	
	[fu^{33}]	'pay'	[vu^{22}]	'attach /enclose'
	[sɒ^{35}]	'little'	[zɑɒ^{13}]	'make/invent'
	[ɕy^{33}]	'need'	[ʑy^{22}]	'tree'
	[ho^{52}]	'shrimp'	[ɦo^{31}]	'river'

Most languages in the world have no contrast between voiced and voiceless glottal fricatives (/h/ vs. /ɦ/). In fact, Maddieson (1984a) lists only two languages in the world which have a contrast between /h/ and /ɦ/, and one of these is from the Wu language family, just like SX. Although there has been some disagreement on the classification of /h/ and /ɦ/ (e.g. whether they are fricatives or laryngeals, vowels or approximants; see Maddieson 1984a), /h/ and /ɦ/ in SX are undoubtedly glottal fricatives

[11] Nartey's (1982) analysis includes Amharic, Arabic, English, Hebrew, Hopi, Japanese, Korean, Navajo, Papago, Pima, Polish, Swedish, Yoruba and Zuni.
[12] Taba is an Austronesian language spoken in the northern Maluku province, Indonesia.

and both are phonemic consonants which may occur in the syllable onset. For example:

(23)	voiceless		voiced	
	[ho^{52}]	‘shrimp’	[ɦo^{31}]	‘river’
	[hɛʔ^{5}]	‘blind’	[ɦɛʔ^{3}]	‘narrow’

According to Maddieson (1984a), about 63.7% of the languages in UPSID[13] have voiceless /h/, while only 4.1% of the languages have voiced /ɦ/. In SX, voiced /ɦ/ is much more frequent than voiceless /h/ in syllable onsets. Besides the words mentioned above, I list some more examples in (24):

(24)	Voiced		Voiceless	
	[ɦɑɒ^{31}]	‘roar’	[hɑɒ^{52}]	‘spend (time)’
	[ɦəŋ^{31}]	‘stable’	[həŋ^{52}]	‘groan’
	[ɦoʔ^{3}]	‘learn’	[hoʔ^{5}]	(surname)
	[ɦwo^{22}]	‘speech’	[hwo^{33}]	‘spend’
	[ɦwe^{31}]	‘return’	[hwe^{52}]	‘ash’
	[ɦwɛ̃^{31}]	‘play’	[hwɛ̃^{52}]	‘well-behaved’
	[ɦwɒŋ^{31}]	‘king’	[hwɒŋ^{52}]	‘nervous’
	[ɦu^{22}]	‘unclear’	[hu^{33}]	‘call’
	[ɦɒŋ^{31}]	‘line’	[hɑŋ^{52}]	‘ram’
	[ɦwa^{22}]	‘bad’	?	(undecided possibility)
	[ɦjẽ^{22}]	‘hate’	*	(systematic
	[ɦjɤ^{31}]	‘oil’	*	impossibility)
	[ɦjoʔ^{3}]	‘moon’	*	
	[ɦja^{22}]	‘night’	*	
	[ɦɥə̃^{31}]	‘cloud’	*	

The examples in (24) show that where there is an acceptable syllable beginning with voiceless [h] in SX, there is also a word with [ɦ] followed by an identical final vowel or combination. The reverse is not true, however. In short, [ɦ] is more often found in this position in SX than [h]. In fact, [h] plus a front high vowel is systematically ruled out in SX (*[h][+high, −back]), while [ɦ] can be followed by all different vowels, as

---

[13] The UCLA Phonological Segment Inventory Database.

shown in (24). There are no constraints on the distribution of [ɦ] as an onset. This will be discussed in detail in chapter 4.

In some languages [h] and [ɦ] have been described as voiceless versus breathy-voiced counterparts of the vowels that follow them (Ladefoged 1971). In SX, vowel-initial syllables invariably receive a [ɦ] if the tone is low-register, but [ʔ] appears (not [h]) if it is high-register, as discussed in §2.2.2 above. More examples are given in (25):

(25) [ɦ]:		[ʔ]:		[h]:	
[ɦi^{31}]	'move'	[ʔi^{52}]	'clothes'	*[hi^{52}]	
[ɦjɤ^{31}]	'oil'	[ʔjɤ^{52}]	'low voice'	*[hjɤ^{52}]	
[ɦjẽ^{31}]	'salt'	[ʔjẽ^{33}]	'smoke'	*[hjẽ^{33}]	
[ɦɥɵ̃^{31}]	'cloud'	[ʔɥɵ̃^{33}]	'complain'	*[hɥɵ̃^{33}]	

Generally speaking, there are more voiceless obstruents in a language than voiced ones; if a language has a voiced obstruent, it usually also has the voiceless counterpart (Maddieson 1984a). However, the asymmetric distribution of [h] and [ɦ] in SX is inconsistent with this general situation. I assume that the wider distribution of [ɦ] compared to [h] in SX is caused by the tonal system, which has a voiceless-obstruent/high-register and voiced-obstruent/low-register correlation (see further discussion about it in chapter 5). This means that it is impossible for a syllable to have a low-register tone and a voiceless initial obstruent or a high-register tone and a voiced initial obstruent. There is cross-linguistic evidence that the glottal stop [ʔ] is often inserted before a vowel-initial syllable, as discussed in §2.2.2. The situation in SX is similar in that [ʔ] is always inserted before a high-register syllable when there is no other initial consonant. To satisfy the tonal system of SX, a voiced glottal obstruent is required before a low-pitched rhyme when there is no other voiced consonant (the tonal system will be discussed in detail in chapter 5). But there is no voiced counterpart of the glottal stop [ʔ] in the consonant inventory of the world's languages. However, [ʔ] and [ɦ] form a natural class in certain aspects. Both can be regarded as laryngeals and the laryngeals are always considered as placeless-component, lacking the complexity when compared with other consonants (Harris & Lindsey 1995; Humbert 1995; Botma 2004). Many Chinese scholars assume that [h] is a phonemic consonant while [ʔ] and [ɦ] are a pair of phonetic onsets in the Wu Chinese (Xu & Tang 1988). The optimal choice of [ʔ] and [ɦ] for the onset of a high-register syllable and low-register syllable, respectively, when there is no other onset consonant, can be stipulated by the following four onset-condition constraints:

(26) Onset-condition Constraints:
a. ONSET(Itô 1989)
Syllables must have an onset.
b. Voiced-L
Voiced initial obstruents must have low-register tones on the following vowels.
c. Voiceless-H
Voiceless initial obstruents must have high-register tones on the following vowels.

The constraint ONSET in (26a) will rule out full-tone syllables without an onset such as *[e^{52}], *[a^{33}], *[u^{31}] and *[o^{22}] in surface representation and the constraints of (26b) and (26c) will also rule out any syllable that has voiced initial obstruent for high register or voiceless initial obstruent for low register. Such a consonant-tone correlation is well explained by Halle & Stevens' (1971) laryngeal feature specifications ([stiff] & [slack], which will be discussed in chapter 5).

In short, I assume that the onset [ɦ] in SX is both a phonological onset when it is contrastive with [h], as shown in (24), and a phonetic onset just like [ʔ] in surface representation, as shown in (25). Such is the case decided by the onset-condition constraints in (26). This is why [ɦ] is more frequent than [h] in the formation of syllables in SX.

2.2.5 The sonorants

Besides the 24 obstruents discussed above, there are also five sonorant consonants that can appear in syllable-initial position in SX. In this subsection I present a brief analysis of the five sonorant initials of SX. In (27), I list the four nasal initials:

(27)	bilabial	alveolar	alveolo-palatal	velar
	m	n	ȵ	ŋ

SX has a bilabial, alveolar, alveolo-palatal and velar nasal in surface representation, all of which can appear in the initial position of the syllable, but not all are in contrastive distribution. Consider the three groups of syllables in (28):

(28)		[o]		[i]		[u]	
	[m]:	[mo^{22}]	'grind'	[mi^{13}]	'rice'	[mu^{22}]	'grave'
	[n]:	[no^{22}]	'glutinous'	*[ni]	—	[nu^{22}]	'anger'
	[ȵ]:	*[ȵo]	—	[ȵi^{22}]	'two'	*[ȵu]	—
	[ŋ]:	[ŋo^{22}]	'hungry'	*[ŋi]	—	*[ŋu]	—

The examples in (28) show some syllables with the four initial nasals and three different vowels, [o], [i] and [u], and also show that not all the four nasals can surface with the same vowel. This means that the four nasals cannot be all contrastive with each other. In fact, only [m], [n] and [ŋ] can all occur before [o]. I assume that [m], [n] and [ŋ] contrast with each other, as attested by more examples such as [me^{22}] 'younger sister', [ne^{22}] 'patient' and [ŋe^{31}] 'fool', and [ma^{13}] 'buy', [na^{13}] 'milk' and [ŋa^{13}] 'we, us'. Thus, [m], [n] and [ŋ] are phonemic nasals in SX. The velar nasal [ŋ] cannot occur before the high vowels, [i] and [u], in SX. This can be formalized by a constraint:

(29) *[ŋ][+high]
[ŋ] cannot occur before any [+high] (semi-)vowel.

The constraint *[ŋ][+high] in (29) forbids */ŋi/, */ŋja/, */ŋu/ and */ŋwa/ in SX. The alveolar nasal [n] cannot appear before the front high vowel [i] or the glide [j] while the alveolo-palatal nasal [ȵ] can only appear before [i]/[j].[14] This indicates that [ȵ] is in complementary distribution with [n]. Thus, [ȵ] is an allophone of the distinctive nasal /n/, which can be expressed by a nasal palatalization rule as in (30) or a constraint as in (31):

(30) $$[\text{n}] \rightarrow [\text{ȵ}] \;/\; \_\_ \begin{bmatrix}\text{+high} \\ \text{–back}\end{bmatrix}$$

or

(31) $$*[\text{n}] \begin{bmatrix}\text{+high} \\ \text{–back}\end{bmatrix}$$

The rule in (30) formalizes that in SX the coronal nasal becomes alveolo-palatal [ȵ] when preceding a front high (semi-)vowel, e.g. [i], [ɪ] or [j],

[14] Actually in SX syllables, [m] can appear before all vowels; [n] and [ŋ] can appear all except high front (semi-)vowels. The phonotactics of the SX segment sequences will be discussed in chapter 4.

and the constraint in (31) rules out any syllable where [n] is followed by a front high (semi-)vowel. More examples that involve this nasal palatalization rule are given in (32):

(32)	[ȵi^{22}]	'two'	*[ni]
	[ȵjɤ^{31}]	'cow/ox'	*[njɤ]
	[ȵɪŋ^{31}]	'silver'	*[nɪŋ]
	[ȵɪʔ^{3}]	'hot'	*[nɪʔ]

From the analysis above, we have observed that in SX, there are six alveolo-palatal consonants: [tɕ], [tɕ^{h}], [dʑ], [ɕ], [ʑ] and [ȵ]. All the alveolo-palatal consonants have the same contribution that they can only precede high front (semi-)vowels. Some scholars (e.g. Duanmu 1999) assume that all the alveolo-palatal consonants are only allophonic segments. The distribution of all the consonants will be discussed in chapter 4.

However, the nasal palatalization rule does not apply to the bilabial nasal in SX, presumably because the coronal in general shifts to the alveolo-palatal place of articulation more easily than the bilabial one. However, nasal asymmetric behavior is common cross-linguistically, though they are of a natural class (e.g. Bhat 1978; Botma 2004). For example, the formation of compounds in Dutch has optional place assimilation of /n/, but not of /m ŋ/, which strongly suggests asymmetric behaviour in the class of the nasals (Botma 2004), as shown in (33):

(33)	steen+bok	stee[mb]ok	'Capricorn'	(*[nb])
	tram+kaart	tra[mk]aart	'tram ticket'	
	meng+paneel	me[ŋp]aneel	'mixing panel'	

Finally, SX also permits the lateral /l/ in the initial position of a syllable, but it has no /r/. /l/ is a common onset in SX syllables. For example:

(34)	[li^{13}]	'inside'	[ləŋ^{13}]	'cold'
	[le^{31}]	'come'	[laŋ^{13}]	'cool'
	[la^{52}]	'pull'	[loŋ^{31}]	'dragon'
	[lɑɒ^{13}]	'old'	[loʔ^{3}]	'green'
	[ljɤ^{31}]	'flow'	[ljaŋ^{22}]	'bright'

Altogether, SX has 29 consonants in the surface representation, all of which can be used as the initial of a syllable. Of these, only the glottal stop /ʔ/ and the velar nasal /ŋ/ can appear in postnuclear position (cf. § 2.3.5.3). The constraints that formalize this will be discussed in more detail in the next section. I list all 29 consonants of SX in (35):

(35) SX consonant inventory:

Manner \ Place		bilabial	labio-dental	dental	alveolar	alveolo-palatal	velar	glottal
stop	−asp	p			t		k	ʔ
	+asp	p^h			t^h		k^h	
	+voice	b			d		g	
africate	−asp			ts		tɕ		
	+asp			ts^h		$tɕ^h$		
	+voice			dz		dʑ		
fric.	−voice		f	s		ɕ		h
	+voice		v	z		ʑ		ɦ
nasal		m			n	ȵ	ŋ	
lateral					l			

Next, I will present the feature specifications for the 29 consonants in SX in (35) for the sake of further phonological analysis of their exact distribution, which will be discussed in chapter 4. To specify the 29 SX consonants, I use ten distinctive features, which include:

Laryngeal features: [stiff], [slack], [spread];
Place features: [anterior], [coronal], [dorsal];
Strictures features: [nasal], [continuant], [strident], [sonorant].

Among the ten features above, following Halle & Stevens (1971), I use [stiff] and [slack] instead of [voice] to specify obstruents, [+stiff] for [–voice] and [+slack] for [+voice], although one of them is redundant for the specifications of obstruents. Sonorants have default voicing so that they are not specified for [voice] in (36). The feature [dorsal] has the same value as [back] for consonants. To distinguish dental/alveolar from alveolo-palatal consonants, I use the feature [anterior]. In the text, I also use [apical] as a feature (as assumed by Williamson (1977)), when discussing the distribution of the apical vowel [ɿ]; sometimes, I use [back] instead of [dors] for consistency with vowel, and sometimes I also use other more features such as [labial] for both [p] and [u]. In table (36), I only mark "+", leaving "–" blank for simplicity.

(36) Feature specifications for the 29 consonants in SX:

	p	pʰ	b	m	f	v	ts	tsʰ	dz	s	z	t	tʰ	d	n	l	tɕ	tɕʰ	dʑ	ɕ	ʑ	ȵ	k	kʰ	g	ŋ	ʔ	h	ɦ
stiff	+	+			+		+	+		+		+	+				+	+		+			+	+			+		
slack			+			+			+		+			+					+		+				+				+
spr		+						+					+					+						+				+	+
ant	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+													
cor							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+							
dors																							+	+	+	+			
nas				+											+							+				+			
cont					+	+				+	+									+	+							+	+
strid					+	+	+	+	+	+	+						+	+	+	+	+								
son				+											+	+						+				+			

In the feature matrix in (36), the 29 consonants in SX are all distinguished from one another, using ten features. Since SX still retains historically voiced obstruents, I still apply [±voice] to the specifications for obstruents as well as sonorants throughout this dissertation when [stiff] and [slack] are applied particularly to the analysis of consonant-tone interaction. There are controversial specifications for the feature [strident]. I assume that all the anterior and coronal fricatives and affricates in SX are [+strident], while in some languages, [±strident] is used to distinguish [f] ([+strid]) and [ɸ] ([–strid]), [s] ([+strid]) and [θ] ([–strid]), and [ʃ] ([+strid]) and [ç] ([–strid]) (see Trask 1996). All the analyses of the SX consonants are based on the feature specifications in (36) in this dissertation.

The initial consonants have been relatively stable through time and there is not much disagreement on the number or realizations of the initial consonants among field workers or researchers, with the exception of the initial glottal stop [ʔ]. Some researchers (e.g. Yang & Yang 2000; Campbell 2003) prefer the term "zero onset" with respect to syllables beginning with vowels or glides. The syllabic status of glides will be discussed in detail in chapter 4. The behaviour of vowels and glides in Finals will be discussed in the next section.

2.3 Finals

At the beginning of this chapter, I explained that "finals" in Chinese syllable structure refer to all material that follows syllable-initial consonants. However, "finals" in SX are not equivalent to "rhymes". The syllable structure of SX will be discussed in chapter 4. In this section, I will present a phonologically analytic description of all the possible surface segments in the finals of SX.

The finals in SX are much more complicated and unstable than the initials. Different native speakers may realize different phonetic vowels in the same lexical form. Another remarkable phenomenon of SX is that quite a number of lexical syllables are different in pronunciation when comparing literary and colloquial style.[15] For example:

[15] In SX, literary style refers to the syllables for the written forms, which are mostly borrowed from Mandarin or phonetically influenced by Mandarin and are always more formal; colloquial style refers to the syllables for the oral forms, which are less formal.

(37) Literary style / Colloquial style

Literary style	Colloquial style	
[vi^{31}]	[bi^{31}]	'grease'
[vi^{22}]	[bi^{22}]	'taste'
[tɕjɑɒ^{52}]	[kɑɒ^{52}]	'hand in'
[da^{22}]	[do^{22}]	'big'
[zəŋ^{22}]	[ɲɪŋ^{22}]	'recognize'
[tso^{35}]	[tɕja^{33}]	'left'

The difference between literary style and colloquial style in SX can be found in the onset, the rhyme, or the whole syllable, as shown in (37). This difference is caused mainly by the influence of Mandarin, so that the literary style is close to the Mandarin pronunciation. These two styles are equally frequent in modern SX, each usually occurring with fixed lexical collocations. Such differences are not so common in other Chinese dialects. Researchers disagree on the phonetic transcription or even the number of finals that exist in SX. In this subsection, I will first discuss the Final inventory of SX because all syllables in SX are usually split into two parts: Initial and Final, and all the surface vowels occur in fixed combinations in Finals.

2.3.1 Final inventory

The exact inventory of Finals in SX is a controversy issue, not only with regard to the number of Finals but also with regard to vowel qualities of some final combinations. In this subsection, I will introduce some different versions of the Final inventory of SX and present my proposal after comparison with other proposals. Chao (1928) made an investigation of Wu dialects and recorded all the phonetic vowels occurring in the Finals[16] of SX, which can be summarized in the following table:

[16] In traditional Chinese phonology, Finals were referred to as 'yunmu', classified into four categories (call 'sihu'), viz. *Kaikouhu, Qichihu, Hekouhu and Cuokouhu.. Kaikouhu* includes those 'yunmu' with simple vowels as the rhyme or beginning with a vowel; *Qichihu* are those with *i* as the rhyme or beginning with *i*; *Hekouhu* are those with *u* as the rhyme or beginning with *u*; *Cuokouhu* are those with *y* as the rhyme or beginning with *y*. In the following tables of Final inventory, all 'yunmu' are categorized into sihu.

(38) Chao’s Finals in SX[17]

	Kaikouhu	*Qichihu*	*Hekouhu*	*Cuokouhu*
In open syllables	ʅ	i, ij	u	yj, yɥ
	e	ie	ue	
	a	ia	ua	
	ɤ, ə, ɵ	iɤ, ɪɤ		
	o	io	uo	yo
	æ, ᴀ		uæ	
	ɑɒ	iɑɒ		
With nasal coda	en, ɘŋ, əŋ	ɪŋ, ʏŋ	uɵn, uɵɲ	
	aŋ	iaŋ	uʌŋ	
	ɒŋ	iɒŋ	uɒŋ	
	uŋ	iuŋ		
With nasalized vowels	æ̃	iæ̃	uæ̃	
	ɵ̃		uɵ̃	yɵ̃
	ẽ	ɪ̃	uẽ	
Syllabic consonants	m̩, ŋ̍, l̩			

Chao’s surface Final inventory in (38) contains altogether 53 different Finals in the form of simple vowels, syllabic consonants, and combinations of GV, VV, VG, VC, and GVC. Chao (1928) recorded these Finals according to individually investigated subjects. His recording represents the phonetic realization of Finals in SX, so that it is not phonologically systematic. Besides, the pronunciation of some vowels can be different from subject to subject. Some of his different forms of Finals are just free variations, such as [uɵ̃], [uɵn] and [uɵɲ], [i] and [ij], [en] and [ẽ], all groups or pairs of which could be used to pronounce the same lexical syllable, differently from speaker to speaker. In Chao’s recording, there are VG combinations such as [yj] and [yɥ], which, I claim, are unacceptable in the SX surface representation because they violate the OCP.

Campbell (2003) presents his inventory of 47 Finals in SX as follows:

[17] This is my summary of Chao’s (1928) Table of Yunmu (in Wu dialects), in which the prenuclear glides were originally transcribed in [i] and [u], which stay in the same symbols as the original in (38), (39) and (40). However, in my analysis through the dissertation, I present the two pre-nuclear glides as [j] and [w].

(39) Campbell's 47 Finals in SX[18]

	Kaikouhu	*Qichihu*	*Hekouhu*	*Cuokouhu*
In open syllables	ʅ	i	u	
	e	ie	ue	yɥ
	a	ia	ua	
	o	io	uo	
	ɤ	iɤ		
	ɑɒ	iɑɒ		
With nasal coda	əŋ	ɪŋ		
	ɒŋ		uɒŋ	
	aŋ	iaŋ	uaŋ	
	uŋ	iuŋ	uuŋ	
With nasalized vowels	æ̃	iæ̃	uæ̃	
	ɵ̃		uɵ̃	yɵ̃
		ĩ		
With stop coda	əʔ	ɪʔ		
	ᴀʔ	iᴀʔ	uᴀʔ	
	oʔ		uoʔ	yoʔ
	eʔ			
Syllabic consonants	m̩ n̩ ŋ̍ l̩			

Campbell's (2003) 47 Finals in SX as shown above are more systematic than Chao's (1928) except some combinations such as [əŋ] and [uuŋ] which are rather uncommon in Chinese dialects and can be hardly manifested by the present data. In Campbell's version of the SX Finals, there is one VG combination [yɥ] which sounds unusual. Yang & Yang (2000) present a similar Final inventory of SX to Campbell's in (40):

[18] Campbell's (2003) 47 Finals in SX are presented on a website: http://wu-dialect.myrice.com/shaoxing.jpg, which may be upgraded every year. The three glides are transcribed as [i], [u] and [y], following the original version.

(40) Yang & Yang's Finals in SX[19]

	Kaikouhu	*Qichihu*	*Hekouhu*	*Cuokouhu*
In open syllables	ɿ	i	u	y
	ɒ	iɒ		
	a	ia	ua	
	o	io	uo	
	ɤ	iɤ		
	e	ie	ue	
With nasal coda	əŋ	iəŋ		
	ɒŋ	iɒŋ	uɒŋ	
	aŋ	iaŋ	uaŋ	
	oŋ	ioŋ	uoŋ	
With nasalized vowels	ɛ̃	iɛ̃	uɛ̃	
	ɤ̃		uɤ̃	yɤ̃
	ẽ	iẽ		
With stop coda	əʔ	iəʔ		
	aʔ	iaʔ		
	oʔ	ioʔ	uoʔ	
	ɛʔ		uɛʔ	
Syllabic consonants	m̩ ȵ̩ ŋ̍ l̩			

In (40), Yang & Yang present 49 Finals in the SX surface representation. The Final inventory in (40) is phonologically more systematic than both Chao's (38) and Compbell's (39), leaving out free variations as well as the VG combinations in the 49 Finals in SX and not much difference in phonetic realization of the surface vowels from the data. However, some Finals in (40) remain problematic from either a phonetic or a phonological viewpoint, e.g. with regard to the question whether [ɒ] alone can be the rhyme of a syllable or whether the combination of [iəŋ] in the surface SX exists. This will be discussed in the following sub-sections. The main difference between Campbell's analysis and Yang & Yang's analysis concerns different vowel qualities. For example, [ᴀʔ] and [æ̃] in (39) refer to [aʔ] and [ɛ̃] in (40), respectively. Based on the different versions of the Final inventory of SX, including Chao's (1928), Yang & Yang's (2000), Campbell's (2003) and others, and also through Yang's

[19] In (40), Yang and Yang (2000) also transcribed the three glides in SX as [i], [u] and [y], instead of [j], [w] and [ɥ] which are used through out in this dissertation.

(2000), Campbell's (2003) and others, and also through exhaustive consultation with the SX native speakers and according to the data and my native intuition of SX, I will present my proposal of the Final inventory of SX, as shown in (41):

(41) Finals in SX[20]

	Kaikouhu	*Qichihu*	*Hekouhu*	*Cuokouhu*
In open syllables	ɿ	i	u	y
	e	je	we	
	a	ja	wa	
	o	jo	wo	
	ɤ	jɤ		
	ɑɒ	jɑɒ		
With nasal coda	əŋ	ɪŋ		
	ɒŋ		wɒŋ	
	aŋ	jaŋ	waŋ	
	oŋ	joŋ	woŋ	
With nasalized vowels	ɛ̃	jɛ̃	wɛ̃	
	ɵ̃		wɵ̃	ɥɵ̃
	ẽ	jẽ		
With stop coda	əʔ	ɪʔ		
	aʔ	jaʔ	waʔ	
	oʔ	joʔ	woʔ	
	ɛʔ			
Syllabic consonants	m̩ n̩ ŋ̍ l̩			

The Final inventory of SX I present in (41) contains 48 different forms of Finals, including 17 in open syllables, ten with a nasal coda, eight with nasalized vowels as the rhyme, and nine with final stops. The 48 Finals in (41) can also be re-classified as V, C̩,[21] GV, VV, VC, and GVC, in general linguistic terms. The difference in the Final inventory of SX between Yang & Yang's (2000) (indicated by 'Yangs'' below) and my proposal (indicated by 'Zhang's' below) can be summarized as follows:

---

[20] In (41), I transcribe the three glides in SX in [j], [w] and [ɥ] which are equivalent to [i], [u] and [y], respectively, in traditional Chinese transcription mentioned in this chapter.

[21] Here 'C̩' refers to syllabic consonants.

(42)

Yangs’	n̩̥	ɤ̃	ɒ	iəŋ	iɒŋ	iaŋ	uɤ̃	yɤ̃	iəʔ	uɛʔ
Zhang’s	n̩	ɵ̃	ɑɒ	ɪŋ	jaŋ		wɵ̃	ɥɵ̃	ɪʔ	waʔ

The difference between the Yangs’ and Zhang’s representations in the Final inventory of SX, as shown in (42), concern either phonological behavior or phonetic realization of some vowels, which will be discussed in the following sub-sections and next chapter. However, in my analysis, SX has 48 Finals in three different forms, including syllabic consonants, vowels (monophthongs) and complex finals (combinations). In this section, I will discuss consonant syllabicity, vowel nasalization, phonological behaviours of vowels in the syllable, and different patterns and combinations of complex Finals in the SX syllables, based on the Final inventory in (41).

2.3.2 Syllabic consonants

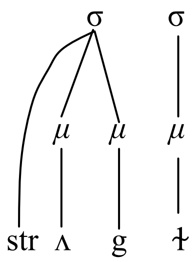
2.3.2.1 Rules of syllabicity

In this subsection, I will examine the syllabic consonants and analyze the phonological rules and constraints concerning the syllabicity of consonants in SX. A syllabic consonant is a segment which has the phonetic characteristics of a consonant but which, in a particular case, functions as a syllabic nucleus (e.g. Trask 1996). SX has four sonorant consonants that can be syllabic, which are listed in (43):

(43) m̩ n̩ ŋ̍ l̩

The consonants in SX that can be syllabic are the bilabial nasal, the alveolar nasal, the velar nasal, and the lateral liquid. These are all distinctive sonorants in the SX initial position, which means that all the phonemic sonorant consonants can be syllabic in SX. Syllabicity of consonants is found in many languages. In English, for example, syllabic sonorants appear after an obstruent in word final position, as shown in (44):

(44) [ˈstrʌgɫ̩] *struggle*



The example in (44) shows that in English, the liquid /l/, when in word-final position and preceded by an obstruent, is syllabic, making *struggle* a disyllabic word [ˈstrʌgɫ̩], in which the syllabic lateral [ɫ̩] is the nucleus of the second syllable.

However, sonorant consonants in SX are syllabic in a different environment from English. In SX, a sonorant becomes syllabic when standing alone. It is always assumed that syllabic [m̩], [n̩], [ŋ̍] and [l̩] in SX are monosyllables, without any other consonant or a schwa preceding or following, to form a lexical syllable by themselves. For example:

(45) Lexical syllable

a. [m̩^{22}]	‘yes’
b. [n̩^{33}]([~jɤ^{35}])	‘not’ (only used together with ‘have’)
c. [ŋ̍^{13}]	‘five’
d. [l̩^{22}]	‘also’

I assume that underlyingly, the lexical items in (45) consist of a single nasal or lateral for the whole lexical syllable. In (45a), the word for *yes* is a single syllabic bilabial nasal. The lexical syllable [m̩] does not combine with any other syllable to form a phrase: it is a lexical word in itself. In SX, the syllabicity of sonorants can be formulated as a phonological rule, as in (46):

(46) [+son] → [+syll] / $ _ $

Rule (46) says that any sonorant becomes syllabic when it occurs on its own between syllable boundaries underlyingly. This rule also includes vowels.

2.3.2.2 Weight-by-stress

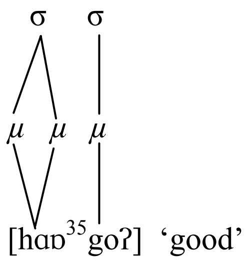
Like many other Chinese dialects, SX has no contrast between long and short vowels. Thus, a feature like [± long] or an underlying mora distinction between long and short vowels (or consonants) (cf. ch.1) plays no role. Surface syllable weight is not decided at the phonemic level or by the syllable structure (e.g. CV, CVC, or CGVC), but by stress. Stressed syllables (as opposed to unstressed ones) may be signaled acoustically with higher pitch, greater intensity, longer duration or some combination of these (Selkirk 1984). However, stress in tone languages such as SX and Mandarin plays a quite different role from that in stress languages such as English and French. In tone languages, it is often difficult or impossible for someone who is not a native speaker of the language to identify stress functioning separately from tone: syllables may sound stronger or weaker according to the tone they bear. Generally speaking, in tone languages pitch is divorced from stress and prominence, which means that various combinations of H and L (and sometimes also M=mid) tones may occur in a single word. SX is a tonal monosyllabic language. Almost every syllable is a word and every word has a stress to demarcate it as a lexical unit. Underlyingly, every syllable is bimoraic in SX, but the weight of the syllable is realized by stress. A lexical syllable must be prominent underlyjngly in SX and its bimoraic status has to be realized by way of a full tone, which means that only a bimoraic syllable is a stressed syllable. Thus, stress in tone languages plays a role of realizing syllable weight in the form of tone pitches. Accordingly, if a syllable bears a neutral or zero tone, it has no stress. In SX, as in many other languages, all stressed syllables are heavy and heavy syllables must be stressed (Prince 1990; Prince & Smolensky 1993). In SX, not only are all stressed syllables heavy, but also all unstressed syllables are light, and light syllables cannot bear a full tone. In terms of mora structure, all heavy syllables are bimoraic and all light syllables are monomoraic (Duanmu 1999, 2000; Wang 1999). The correlation between syllable weight and stress in SX is stated in (47):

(47) Weight-by-stress

Stressed syllables must be heavy and bimoraic; unstressed syllables must be light and monomoraic.

The weight-by-stress principle is largely consistent with constraints like WSP (WEIGHTTOSTRESSPRINCIPLE) [22] (Prince 1990) and SWP (STRESSTOWEIGHTPRINCIPLE) [23] (Prince 1983; Myers 1987; Prince & Smolensky 1993). Neither WSP nor SWP, however, constrains unstressed syllables. Weight-by-stress in (47) is a typical characteristic of moraic syllable structure of SX. Weight-by-stress also differs greatly from Weight-by-position (Hayes 1989). Weight-by-stress only allows stressed syllables to be bimoraic no matter whether they are CV, CVV, V or even just syllabic nasals, as in (45). In SX as well as in Mandarin, most of the lexical syllables are stressed. Only some particles for grammatical functions are always unstressed, such as [lɤ] to mark 'past tense' in Mandarin and [goʔ] to mark the status of 'adjective' in SX. For example, in [hɑɒ^{35}goʔ] 'good', the second syllable is unstressed because it is not lexically meaningful. Their moraic syllable structure can be represented as follows:

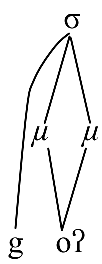
(48)



The representation in (48) shows that the first syllable in [hɑɒ^{35}goʔ] is stressed and is heavy and the second syllable is only an adjective marker grammatically and is unstressed. The unstressed syllable is short and always toneless or has a neutral tone because it is monomoraic. Thus, the unstressed syllable is always a non-TBU in Chinese (Pulleyblank 1986; Duanmu 2000a). The length of a vowel in SX is a morphophonological matter rather than a phonetic matter. If the same syllable plays a different lexical role or a sentential role, it has a different weight. For instance, [goʔ^{3}] in [goʔ^{3}lɑɒ^{31}] 'standstill' is stressed and bimoraic, because it is a lexical word. The moraic structure of the stressed [goʔ] in [goʔ^{3}lɑɒ^{31}] is represented as follows:

[22] WSP (WEIGHTTOSTRESSPRINCIPLE): Heavy syllables must be stressed.
[23] SWP (STRESSTOWEIGHTPRINCIPLE): Stressed syllables must be heavy.

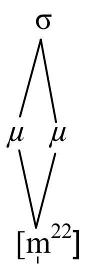
(49)



in [goʔ^{3}lɑɒ^{13}]

According to Weight-by-stress, the syllabic sonorants in (45a), (45c) and (45d) are all independent lexical words and are all stressed and bimoraic because they are all full-tone syllables. For example, the moraic syllable structure of [m̩^{22}] ‘yes’ can be illustrated in (50):

(50)



Duanmu (1999, 2000a) claims that all heavy syllables (CVV,[24] CVC, CVG) of Chinese dialects are bimoraic and that all monomoraic syllables are light (CV). This means that a syllable like CVC or CVG has to be stressed, which is not the case with SX. This issue will be discussed in chapter 4.

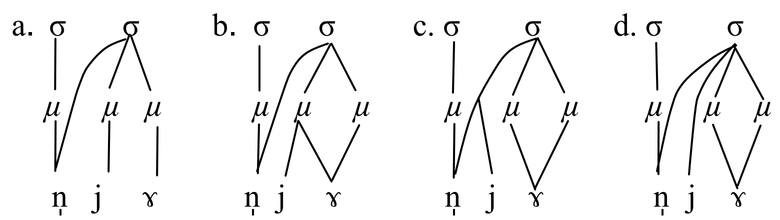
It should be noted that in the list of the syllabic sonorants in (45), all syllables have low-register tones except [n̩^{33}], which has a high-register tone. Is it possible that a sonorant initial has a high-register tone in the syllable? According to the onset-condition constraints in (26), no high-register vowel is preceded by a voiced obstruent onset. However, a voiced initial sonorant can be a high-register syllable. Yip (2002) points outs that in sonorant consonants the rate of vibration of the vocal folds is controlled by a number of factors. Rotation of the thyroid and cricoid cartilages with respect to each other can be deformed in several ways, and as a result the vocal cords may or may not be stiff. So, articulatorily, a

[24] ‘VV’ here refers to either a two-vowel sequence, or a long vowel, or a stressed bimoraic vowel (Duanmu 1999).

syllable with a sonorant initial may have a low-register tone or a high-register tone. There are some examples in SX listed in (51):

(51) a. [lɤ^{22}] ‘leak’
b. [lɤ^{52}] ‘hollow out’
c. [mɪʔ^{3}] ‘extinguish’
d. [mɪʔ^{5}] ‘screw’

In the examples in (51), the syllables of (a) and (c) have low-pitch tones and those of (b) and (d) have high-register tones.[25] This fact is also captured by the configurations of Halle & Stevens’ (1971) laryngeal feature specifications that sonorants are specified as [–stiff, –slack] and high register is specified as [+stiff] or [–slack] and low register is as [+slack] or [–stiff]. The consonant-tone correlation will be discussed in more detail in chapter 5. The example in (45b) also shows that the syllable [n̩^{33}] for ‘not’ is syllabic only when uttered in the combination [n̩jɤ^{35}], which is actually a disyllabic lexical item, meaning ‘not have/haven’t’. A question may arise: is [n̩jɤ^{35}] a CGV monosyllabic unit? If not, what is its phonological representation? In fact, the prosodic word [n̩jɤ^{35}] deserves a comment. I assume that [n̩] in [n̩jɤ^{35}] is a clitic, which is a toneless syllable by itself and is phonetically and phonologically fused with the host syllable [jɤ]. As a syllabic nasal, [n̩] is both a lexical syllable by itself and the onset of the second syllable. The moraic structure of [n̩jɤ^{35}] can be represented in (at least) four possible ways, given in (52):

(52)


The possible moraic syllable structures in (52) all show that the syllabic nasal is the nucleus of the first syllable and also the onset of the second syllable. The only difference is the location of the prenuclear glide [j]. It can be assumed that [j] has one mora independently as [ɤ] does, as in

[25] According to the tonal structure of SX, the high register carries high-pitch tones of 52, 35, 33, 5 and the low register carries low-pitch tones of 31, 13, 22 and 3. The details of the tonal system of SX will be discussed in chapter 5.

(52a), suggesting that [j] is in the Nucleus; or [j] is moraic but shares one mora with [ɤ], as in (52b), suggesting [j] is in the Nucleus but cannot bear a tone by itself; or [j] is non-moraic and is in the Onset, as in (52c); or [j] is non-moraic and is neither in Onset nor in Nucleus, as in (52d). I claim that the status of the prenuclear glide [j] in [n̩jɤ^{35}] is like in (52d), which will be discussed in detail in chapter 4. However, there still remain some mysterious questions as to how [n̩] becomes a clitic; why [n̩] in [n̩jɤ^{35}] is not palatalized as [ȵ] according to the nasal palatalization rule in (30); why [n̩jɤ^{35}] is disyllabic; and what the phonological motivation for the syllable structure in (52) is. All these issues will be discussed in chapter 5 (see also Zhang 2005: 69-79).

In (45c), the syllabic velar nasal [ŋ̍] has three different meanings with different tones, representing three different lexical words, as listed in (53):

(53) a. [ŋ̍^{31}] ‘fish’
b. [ŋ̍^{13}] ‘five’
c. [ŋ̍^{22}] ([tə̃^{33}~^{22}]) ‘a dragon boat festival’

In (53c), the lexical syllable [ŋ̍^{22}] only exists in a fixed phrase for the meaning of ‘dragon boat festival’, a traditional Chinese festival. All the three forms of the syllabic velar nasal in (53) are stressed syllables because they all bear full tones, so each can constitute a bimoraic syllable by itself, according to Weight-by-stress, as illustrated in (54):

(54)
σ
μ μ
[ŋ̍^{31}] ‘fish’

Nasals can be syllabic in many languages. However, it is not very common for nasals to be syllabic in a monosyllable and by themselves as in SX.

The liquid [l̩] in (45d) is syllabic in much the same way as it is in English. When it is syllabically articulated, the middle of the tongue is excessively lowered so that the sound is pronounced with no intervening vocoid. There is a big difference in articulation and acoustics between the syllabic lateral and the non-syllabic lateral. In SX, the initial [l] and syllabic [l̩] cannot appear in the same world because SX is a monosyllabic language and has no coda [l]. Like English, the initial [l] and syllabic [l̩] in SX are allophones of the phoneme /l/ because the syllabic [l̩] is only possible when it occurs alone. It is in complementary distribution with the initial [l]. For example:

(55) [la^{52}] 'pull'
[lɑɒ^{13}] 'old'
[lu^{22}] 'road'
[ljaŋ^{22}] 'two (people)'
[loʔ^{3}] 'green'
[l̩^{22}] 'and/also'

The examples from (43) to (55) all show that sonorant consonants in SX can be syllabic, as is stipulated by the rule in (46). Actually, all consonants can be syllabic in one way or another, but only if there is no other better peak available (Laver 1994; Ladefoged & Maddieson 1996; among others) because every syllable must have a peak. I propose the following peak principle:

(56) Peak Principle:
Segment α can be the syllable peak iff α is the most sonorant segment in the syllable.

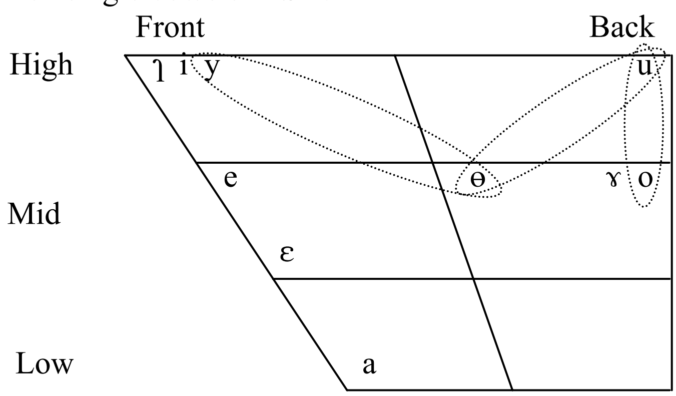
The Peak Principle in (56) says that any segment can be the peak of a syllable in a certain environment. Vowels are always good peaks because they are more sonorant than any consonant; sonorant consonants are likely to be the peak because they are [+son]; obstruents can also be syllabic, but very rarely so because they are the least sonorant. However, some languages, such as Imdlawn Tashlhiyt Berber (Dell & Elmedlaoui 1985) and Bella Coola (Bagemihl 1991), have been described as having syllabic obstruents (see also Botma 2004: 263).

2.3.3 Vowels

As was mentioned at the beginning of this section, there are three categories of "Finals", one of which is the category of single vowels. In this sub-section, I will present an analytic description of all the single vowels in the surface representation of SX. I use the term ‘single vowel’ instead of monophthong to avoid any association with diphthongs, for there has been a discussion in the literature whether there are diphthongs and triphthongs in Mandarin and other dialects (Zhan 1991; Chan 1997; Wiese 1997; and others). I claim that there are no triphthongs in Mandarin and that there are no triphthongs or diphthongs in SX (see the next subsection). Vowels involved in a combination will also be discussed in the next subsection.

In this subsection I will discuss how single vowels are used as rhymes in SX in surface representation, focusing on the surface vowels, vowel nasalization, and their different distributions. In SX, there are ten single vowels in the surface representation, used as rhymes in the syllables. They are presented in the vowel diagram in (57):

(57) Ten single vowels in SX:



In (57), the circled vowels are rounded. The vowel diagram in (57) offers a clear picture of what single vowels[26] may occur in the rhyme of SX syllables and where they are located. SX has more front vowels than back vowels and more unrounded vowels than rounded vowels, which is a natural arrangement because front vowels are naturally unrounded. We can easily acoustically locate these vowels in the above vowel diagram, except [ʅ], which is a very remarkable phone in SX as well as in Mandarin. [ʅ] is usually regarded as an apical vowel which exists in many Chinese dialects. There has been a discussion of the phonetic and phonological status of the apical vowel [ʅ] in Chinese (Karlgren 1915–1926; Chao 1968;

[26] The ten vowels in (57) are all single vowels which occur alone as the Rhyme in the SX syllables, excluding vowels which only occur in combinations such as [ɪ], [ə], [ɑ], and [ɒ].

Kratochvil 1968; Ladefoged & Maddieson 1990; Wiese 1997). Wiese (1997) assumes that [ɿ] in Chinese is not a vowel, but a syllabic fricative. I argue that [ɿ] in SX as well as in Mandarin is an apical vowel and an allophone of /i/. I will present my analysis of the apical vowel [ɿ] in the following subsection.

2.3.3.1 Apical vowel

Apical vowels are phonetically vowels, which are produced with the tip of tongue touching the anterior portion of the palate. Thus, they are also called fricative vowels (Ladefoged & Maddieson 1996). The contact location is in the denti-alveolar. There are several apical vowels in Chinese dialects such as [ɿ ʅ ʮ ʯ].[27] In SX, there is only one apical vowel [ɿ], as shown in (57). Its phonetic and phonological status and its distribution in SX are very similar to that in Mandarin. Wiese (1997: 239) claims, for Mandarin, that [ɿ] is a pseudo-sound that should not have any place in either a phonological or a phonetic description. He regards [ɿ] as a syllabic fricative, as shown in (58):

(58)	Mandarin		Wiese's assumption
	a. [sɿ^{51}][28]	'four'	[sz̩]
	b. [ʐɿ^{51}]	'day'	[zʐ̩]
	SX		
	c. [sɿ^{33}]	'four'	
	d. [zɿ^{22}]	'word(s)'	

Wiese argues that [ɿ] in (58a) and (58b) is "a syllabic consonant identical in place and continuancy to the preceding fricatives". However, Wiese does not give any evidence for denying [ɿ] the status of a vowel. He proposes a filter (1997: 242) to rule out high vowels preceded by [+cor] consonants, as illustrated in (59):

(59) $$* \begin{bmatrix} +\text{cons} \\ -\text{back} \\ +\text{cor} \end{bmatrix} \begin{bmatrix} -\text{cons} \\ -\text{back} \\ +\text{high} \end{bmatrix}$$

---

[27] [ʮ] and [ʯ] are also apical vowels in some Chinese dialects. They are the rounded counterparts of [ɿ] and [ʅ] respectively, differing from the rounded front glide [ɥ].

[28] The syllables in (58a) and (58b) carry Mandarin tones, which are different from those in SX.

Wiese's filter in (59) attempts to state that after [+cor] consonants, front high vowels are not acceptable so that [ʅ] may be a syllabic fricative. However, he realizes that the filter cannot be correct as stated in (59) because [t] and [t^h] are also [+cor] and [ti] and [t^hi] are well-formed syllables in both Mandarin and SX. The main reason for Wiese's assumption is that [in] and [iŋ] are well-formed and wherever [i] is acceptable, [in] and [iŋ] are also acceptable with the same initial consonants in Mandarin, such as [pi^{55}] 'close', [pin^{55}] 'guest' and [$piŋ^{55}$] 'soldier' while *[ʅn] or *[ʅŋ] is never possible in any case. However, whether [i], [in] or [iŋ] can be preceded by the same consonant is simply a matter of phonotactics. For example:

		Mandarin		SX	
(60)	a.	[t^hi^{55}]	'shave'	[t^hi^{33}]	'shave'
		*[t^hin]		*[t^hɪn]	
		[$t^hiŋ^{55}$]	'listen'	[$t^hɪŋ^{33}$][29]	'listen'
	b.	*[si]		*[si]	
		*[sin]		*[sɪn]	
		*[siŋ]		*[sɪŋ]	

The examples in (60) show that /i/ and /iŋ/ can occur after /t^h/, but /in/ cannot, while /i/, /in/ or /iŋ/ cannot occur after /s/[30] in either Mandarin or SX, because of their phonotactics. Wiese claims that the nucleus preceded by /s/ in (58) must be a syllabic fricative [z̩] because [sz̩] is acceptable, as shown in his argument in (58), while *[sz̩n] or *[sz̩ŋ] is not found. This is ill-formed only because of the Sonority Sequencing Principle (SSP).[31] Wiese argues that if the nucleus in (58) were a vowel, [ʅn] or [ʅŋ] should also be allowed in Mandarin. However, the examples in (60) prove that [V], [Vn] and [Vŋ] could have a different distribution. There are more examples such as in (61):

[29] [ɪ] is an allophone of /i/ in SX, which will be discussed in chapter 3.
[30] Here '//' is used instead of '[ ]' for the purpose of indicating that the same underlying phonemes are involved in both Mandarin and SX.
[31] SSP: Sonority increases towards the syllable peak and decreases towards the syllable margins (see Clements 1990; Roca 1994; Morelli 1999).

(61) Mandarin SX

a. [mo^{51}]	'mill'	[mo^{22}]	'mill'
b. *[mon]		*[mon]	
c. *[moŋ]		[moŋ]	'dream'

The examples in (61) show that [o] is acceptable after [m], but [on] or [oŋ] is not possible after the same consonant in Mandarin while in SX [on] is not possible after [m]. However, it would be senseless to argue that [o] in [mo] is not a vowel but a syllabic consonant because there is no *[mon] or *[moŋ]. In SX, [ɤ] can occur after different onset consonants, but [ɤ] can never be followed by any consonant, disallowing any combination of *[ɤC] in surface representation, such as *[ɤŋ] and *[ɤʔ], while [Vŋ] and [Vʔ] are well-formed combinations in SX. For example:

(62)

[$dɤ^{31}$]	'head'	*[dɤŋ]	*[dɤʔ]
[$fɤ^{35}$]	'deny'	*[fɤŋ]	*[fɤʔ]
[$lɤ^{22}$]	'leak'	*[lɤŋ]	*[lɤʔ]
[$tsɤ^{35}$]	'walk'	*[tsɤŋ]	*[tsɤʔ]
[$k^hɤ^{33}$]	'button'	*[k^hɤŋ]	*[k^hɤʔ]

However, there is no reason to doubt that [ɤ] is a vowel. Neither Wiese's filter in (59) nor his argument for *[sz̩n] and *[sz̩ŋ] can support his denying [ɿ] the status of a vowel. If, as Wiese assumes, [ɿ] is a syllabic fricative identical in place and continuancy, it must be an allophone. Then what is the underlying segment− a phonemic /z/ or /z̩/? If a syllabic fricative itself is a distinctive phoneme, what is its phonological property as an underlying phoneme? The facts suggest the contrary, viz. [ɿ] is a vowel and phonologically is in complementary distribution with /i/ in both SX and Mandarin. Phonetically, [ɿ] has formant structure, according to Howie (1976).

Some Chinese scholars (see Li, Yu, Chen & Wang 2004: 257–258) present a comparative analysis of vowel formants between Standard Chinese (SC) and Shanghai-Accented Standard Chinese (ASH),[32] among which the formants of [i] and [ɿ] pronounced by male and female are shown in the following table:

[32] SC refers to Mandarin and SAH refers to the standard Chinese spoken by Wu native speakers, like SX natives.

(63) Average values of F1, F2, F3 of [i] and [ɿ] by male and female speakers:

Vowel	F1 (Bark)		F2 (Bark)		F3 (Bark)	
Male:	BJ	SH	BJ	SH	BJ	SH[33]
[i]	3.11352	2.96491	13.6248	14.04572	16.10078	16.20306
[ɿ]	4.05798	3.90083	9.92734	10.55437	15.19762	15.3489
Female:	BJ	SH	BJ	SH	BJ	SH
[i]	3.35486	3.28039	15.21406	15.32132	17.0888	17.42927
[ɿ]	4.50622	4.1378	11.0423	12.19668	15.76756	16.05551

Table (63) shows the formants of [i] and [ɿ] for male and female speakers of Mandarin and Wu, indicating that [ɿ] in both Wu and Mandarin is a vowel in aspect of acoustics. [ɿ] is an allophonic vowel which is made with the tip of tongue touching the upper articulator and is thus called apical vowel. [ɿ] only occurs after dental fricatives and affricates in SX, as shown in (64):

(64) [tsɿ^{35}] ‘paper’
[dzɿ^{22}] ‘late’
[ts^{h}ɿ^{33}] ‘wing’
[zɿ^{22}] ‘word
[sɿ^{33}] ‘try’

Following Bright (1978), Ladefoged and Maddieson (1996: 163) regard “the dental sibilant as being apical”.[34] Presumably, the reason that [ɿ] only occurs after [ts ts^{h} dz s z] in SX is that the apical vowel [ɿ] shares with the apical sibilants the [+apical] feature, [35] as proposed by Williamson (1977). Thus, the apical vowel is made with the tongue in essentially the same position as in the corresponding fricatives or affricates. In short, [ɿ] is a vowel, whether it is called fricative vowel or apical vowel, but an allophonic vowel of the phonemic /i/ in SX as well as

[33] BJ refers to Beijing speakers who speak SC; SH refers to Shanghai (Wu) speakers who speak ASH.
[34] According to Trask (1996), sibilants refer to fricatives and affricates and the feature [+sibilant] is similar to [+strident] by nature. In SX, the dental sibilants include [ts ts^{h} dz s z].
[35] In Williamson’s (1977) feature system, [±apical] is used as a feature to replace [±distributed] in SPE. Neither [apical] nor [distributed] is used for the feature specifications of the 29 consonants in SX, as shown in (36).

in other Chinese dialects. This can be expressed by a phonological rule, as presented in (65):

(65) $/\text{i}/ \rightarrow [\text{ɿ}] \;/\; \begin{bmatrix} +\text{cons} \\ +\text{apical} \end{bmatrix}$ —

Rule (65) shows that the high front vowel /i/ becomes an apical vowel [ɿ] when following an apical consonant. The distribution of [ɿ] and [i] will be discussed in chapter 3.

2.3.3.2 Other vowels

Out of the ten single vowels, as shown in (57), I claim only six are underlying phonemic vowels, viz. /i/, /u/, /e/, /ɤ/, /o/ and /a/, which contrast with each other, as shown in (66):

(66) [di^{22}] 'earth'
[du^{22}] 'ferry'
[de^{22}] 'pocket'
[dɤ^{22}] 'bean'
[do^{22}] 'big'
[da^{22}] 'wash in pan'

The examples in (66) show that the six vowels, /i/, /u/, /e/, /ɤ/, /o/ and /a/ all can occur alone as the rhyme in an open syllable contrasting with each other. This suggests that they are all phonemic vowels. The underlying vowel inventory will be discussed in detail in chapter 3. Another vowel [y] can also occur as the rhyme of a syllable in surface representation, although its distribution is almost as limited as [ɿ]. It can only occur after the lateral [l], the five alveolo-palatal affricates and fricatives and the two phonetic onsets [ɦ] and [ʔ], leaving many systematic impossibilities. Some examples are presented in (67):

(67) [ly^{31}] 'donkey'
[tɕy^{33}] 'expensive'
[tɕ^{h}y^{35}] 'take'
[dʑy^{22}] 'live'
[ɕy^{52}] 'book'
[ʑy^{22}] 'tree'
[ʔy^{33}] 'silted'
[ɦy^{13}] 'rain'

I assume that the front high rounded vowel [y] is not a phonemic vowel because of its limited distribution. This gives rise to the question as to what the underlying vowel is in case of surface [y]. I will argue that the underlying form of [y] is not a single phonemic vowel, but underlyingly a GV combination /iu/ or /wi/, which merges into [y] in surface representation because of an OCP constraint. This issue will be discussed in the GV sub-section (§2.3.5.1), where more details of vowel distribution and the phonological motivation for /iu/ or /wi/ to merge into [y] will be discussed (see chapter 3).

2.3.4 Vowel nasalization

Some vowels have to be, or can be, nasalized when occurring as finals in SX. Among the vowels represented in (57), there are three vowels that can also occur nasalized in the rhyme, viz. [ẽ], [ɛ̃] and [ɵ̃]. Consider the following examples:

(68) [k^{h}ẽ^{33}] 'look at/watch'
[pɵ̃^{33}] 'half'
[pɛ̃^{52}] 'class'

Of the three examples in (68), only [ẽ] has an oral counterpart [e], a phonemic vowel that can appear alone as the rhyme. The other two vowels, [ɛ̃] and [ɵ̃], have no oral counterparts appearing alone as the rhyme in surface representation. I assume that the three nasalized vowels in (68) are not underlying vowels. Vowels usually become nasalized in a phonetic environment that is conducive to nasalization. For example, in English, [æ] is nasalized to [æ̃] in [θæ̃ŋk] 'thank' because of the following nasal. In most cases, vowel nasalization is predictable. Why do the three vowels in (68) have to be nasalized in a syllable of SX? Compare the cognates in (69) of Mandarin and SX, both of which developed from Middle Chinese:

(69)

Mandarin	SX	
[nan^{35}]	[nɛ̃^{31}]	‘difficult’
[hwan^{35}]	[vwɛ̃^{31}]	‘return’
[t^{h}wan^{55}]	[t^{h}ɵ̃^{33}]	‘swallow’
[k^{h}un^{51}]	[k^{h}wɵ̃^{33}]	‘sleepy’
[pen^{35}]	[bẽ^{22}]	‘pail’
[kan^{35}]	[kẽ^{35}]	‘catch up’

The examples in (69) show that in Mandarin all these syllables end with an alveolar nasal, whereas in SX the rhymes consist of nasalized vowels, missing the final alveolar nasal in surface representation (which confirms CODA-COND (16)). This phenomenon suggests that the underlying forms of the syllables above in SX could also have a final alveolar nasal. Strong evidence for this analysis comes from examples of nasal gemination between a syllable ending in a nasal or a nasal vowel and a syllable beginning with a vowel, as shown in (70):

(70)

Monosyllable		Disyllable			
a. [ts^{h}oŋ^{52}]	‘dash’	[ts^{h}oŋ.a][36]	→	[ts^{h}oŋ^{52}.ŋa]	‘Dash!’
b. [nɛ̃^{31}]	‘difficult’	[nɛ̃.a]	→	[nɛ̃^{31}.na]	‘Difficult!’
c. [t^{h}jɛ̃^{52}]	‘heaven’	[t^{h}jɛ̃.a]	→	[t^{h}jɛ̃^{52}.na]	‘God!’

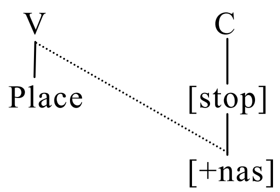
The syllables in (70) are subject to certain phonetic and phonological environment where liaison between the syllables occurs. Usually liaison does not occur between a consonant-final syllable and a vowel-initial syllable if both syllables have full tones. I assume it is because a full-tone syllable must have an onset which is required to assign the register feature, as is stipulated by the onset-condition constraints in (26). Liaison or gemination never occurs between two full-tone syllables in SX (the details of the process of liaison and gemination will be discussed with more examples in chapter 4). Liaison or gemination only occurs when the second syllable is toneless and has no phonological onset. The syllable [ts^{h}oŋ] in (70a) ends with a velar nasal. In the disyllabic form of (70a), the first syllable-final velar nasal is followed by another syllable [a], which has no real lexical meaning but only expresses a certain emphasis, so that it is an unstressed syllable and is toneless. In this context, liaison triggers gemination of the final nasal in the preceding syllable, or phonetically, the

[36] The dot “.” here indicates a syllable boundary.

final nasal [ŋ] becomes ambisyllabic, producing a phonetic onset of the following syllable. In this way, the prosodic word [ts^{h}oŋ.a] is uttered with a geminated velar nasal as [ts^{h}oŋ.ŋa]. In this form, [ŋ] does not autosyllabically become the onset of the second syllable because [ŋ] is still required to be the coda of the first syllable for its integral bimoraic status as a TBU. As for the examples in (70b) and (70c), I assume that the underlying forms of the surface nasalized vowels are VN combinations, which, when followed by a toneless vowel-initial syllable, triggers the phonetic restitution of the lost syllable-final nasal, which is realized as nasal insertion, as shown in the disyllabic forms in (70b) and (70c). In connected speech, there can be a nasal consonant in the onset of the second syllables following SX words such as those in (70), which strongly suggests that both [nɛ̃] in (70b) and [t^{h}jɛ̃] in (70c) also have a final nasal in the underlying form.

There is cross-linguistic evidence that the surface syllable-final nasal vowels are underlying sequences of an oral vowel + a nasal consonant. For example, Portuguese nasal vowels have been analysed as phonological vowel + nasal consonant sequences and the syllables ending in a nasal vowel are treated as closed syllables (Parkinson 1983). More details of the process of the phonological change of the vowel nasalization will be discussed in chapter 4. However, the syllable structure of the final should be VN([+cor]) underlyingly, viz. a vowel followed by an alveolar nasal. The final alveolar nasal in SX was debuccalized through historical attrition (nasal debuccalization will be discussed in chapter 4) so that it drops off and the [+nasal] feature spreads to the preceding vowel, producing vowel nasalization. This can be expressed by the [+nasal] feature spreading (see van de Weijer 1994: 200), as illustrated in (71):

(71)



The feature-spreading structure in (71) can also be represented by a vowel nasalization rule, as follows:

(72) V → Ṽ / /__ N/

The vowel nasalization rule in (72) shows that a vowel becomes nasalized when followed by a nasal underlyingly, which suggests that the deep structure of the surface [k^h ẽ ^{33}] 'watch/look at' is /k^hVn^{33}/ in SX. Nasalized vowels are just the surface forms. The phonological motivation for vowel nasalization will be discussed in detail in chapters 3 and 4. In chapter 3, I will present my analysis of the underlying vowels of the three surface nasal vowels, and in chapter 4, I will discuss why vowel nasalization only happens in the case of [+cor] nasals in the coda, and not with a [+dors] nasal in SX.

2.3.5 Complex Finals

There are four kinds of complex finals in SX. They are GV, VC, VV and GVC. SX has a limited number of rhymes, mainly because there are very few postnuclear consonants in the coda position. In this section, I will discuss all these four types of final combinations. I will introduce the Diphthong Constraint and the OCP(H) constraint in SX. There has been discussion in the literature whether there are diphthongs and triphthongs in Mandarin and other dialects (Zhan 1991; Chan 1997; Wiese 1997; and many others). I claim that there are no triphthongs in Mandarin and that there are no triphthongs or diphthongs in SX. I will present my analysis of these issues in this section.

2.3.5.1 GV

GV is a combination of a glide and a vowel. In SX there are 12 such combinations as surface syllable finals. They are: [jo], [ja], [je], [jɤ], [wa], [we], [wo], [jẽ], [jɛ̃], [wẽ], [wɵ̃], and [ɥɵ̃], five of which have a nasalized vowel. Of these 12 GV combinations, five are preceded by the glide [w], six by [j], and one by [ɥ], which is the glided counterpart of the rounded vowel [y]. In SX [ɥ] only occurs in combination with [ɵ̃], which is also a [−back] rounded vowel. I assume that [ɥ] is in complementary distribution with [j] because [j] is more often used and mostly precedes unrounded vowels such as [ɤ], [a], [e], and [ɛ̃], with the exception of back rounded [o]. This phenomenon suggests a rule or an equivalent constraint, as in (73) and (74), respectively:

(73) $/j/ \rightarrow [ɥ] \; / \; \_\_ \begin{bmatrix} +\text{round} \\ -\text{back} \end{bmatrix}$

or

(74) $*[j] \begin{bmatrix} +\text{round} \\ -\text{back} \end{bmatrix}$

The rule in (73) says that /j/ becomes rounded when followed by a [–back] rounded vowel, and the constraint in (74) rules out such combinations as *[jy] and *[jɵ], permitting the existence of [jo] in SX. This roundedness rule is well supported by the existence of [ɥɵ̃] in SX, which, however, gives rise to the question why [j] in [jo] does not get rounded. There is cross-linguistic evidence that [–back] roundedness is marked and [+back] roundedness is unmarked, which suggests that marked features can trigger assimilation more strongly than unmarked features.

The rule in (73) makes it possible that [ɥɵ̃] exists in SX. Undoubtedly, [ɥ] is an allophone of the underlying glide /j/ in [ɥɵ̃]. The details of the distribution of glides will be discussed in chapter 3. All the 12 GV combinations are well-formed in lexical syllables when preceded by an onset. For example:

(75)	[ɕjo^{33}]	'blood'	[ɦwe^{31}]	'return'
	[ɕja^{33}]	'write'	[ɦwo^{22}]	'speech'
	[ȵɥɵ̃^{13}]	'soft'	[gwɛ̃^{22}]	'circle'
	[tɕ^{h}jɤ^{33}]	'autumn'	[hwɵ̃^{33}]	'happy'
	[pjẽ^{52}]	'edge'	[ɦwa^{22}]	'bad'

The examples in (75) show that all these combinations begin with a glide [w], [j] or [ɥ] in surface representation. The syllabic status of the prenuclear glides has been the topic of many Chinese linguistic studies. I claim that the prenuclear glides in SX are not in the onset (which will be discussed in detail in chapter 4). This gives rise to the question whether these well-formed GV combinations are in the nucleus and are therefore diphthongs. The topic of the representation of diphthongs is a controversial one (cf. Laver 1994; Casali 1996; Trask 1996; Wiese 1997). According to Trask (1996), a diphthong is "a single syllable nucleus which begins with one vowel quality and changes more or less smoothly to a second quality, as in [ju] and [aj]. Usually one of the two vocalic elements is more prominent than the other, this other consisting only of a preceding glide (an on-glide, as in [ju]), or a following glide (an off-glide, as in [aj])". I formalize Trask's definition of a diphthong as follows:

i. Diphthongs form the nucleus of a syllable.
ii. Within the diphthong, the vowel quality changes (involving high, low, back and front).
iii. Either part of the diphthong is a glide.

GV combinations in SX look like diphthongs in many other languages, satisfying Trask's properties (ii) and (iii). Traditional Chinese phonology claims that there are not only diphthongs but also triphthongs in Mandarin as well as in other Chinese dialects (Wang 1963, 1985; Chan 1985; Norman 1988; Wiese 1997; and others). This claim is mainly based on the acceptable combinations in Mandarin, such as [ai], [au], [wa], [wo], [ei], [ou], [wei], [wai], [jau], [jou],[37] etc. The major motivation to assume that Mandarin has diphthongs and triphthongs is to deal with the on-glides in the syllable structure because every 'diphthong' is a combination of a monophthong and either an on-glide or an off-glide (GV or VG) and every 'triphthong' is a combination of a monophthong and both an on-glide and an off-glide (GVG) in Chinese.

Wiese (1997) proposes an underspecification analysis of the Mandarin vowel system and claims the existence of triphthongs in Mandarin, regardless of their position in the syllable structure. On Wiese's (1997: 219) assumption, there are at least four well-formed triphthongs in Mandarin, viz. [jau], [wai], [jou], and [wei], all of which have a GVG structure. Glides play a crucial role in forming a diphthong or triphthong. It is essential to identify the phonological status of glides in the syllable structure when determining whether the sequence is a diphthong or triphthong. As widely claimed, a diphthong must be the single syllable nucleus. However, as attested in all the Chinese literature, is that the prenuclear glide in GV or GVG is never included in the Chinese poetic rhyming. Consider the examples of Mandarin in (76):

(76)	A		B	
a.	[tɕja^{55}]	'family'	[ma^{55}]	'mother'
	[kwa^{55}]	'melon'	[la^{55}]	'pull'

[37] All the off-glides in Chinese dialects are written as [i] and [u], rather than [j] and [w], because I assume that VG in Chinese is a diphthong.

b.	[xwai^{35}]	'chest'	[ts^{h}ai^{35}]	'finance'
	[kwai^{55}]	'well-behaved'	[k^{h}ai^{55}]	'open'
c.	[t^{h}jau^{55}]	'choose'	[t^{h}au^{55}]	'hollow out'
	[p^{h}jau^{55}]	'flow'	[p^{h}au^{55}]	'throw'

In the examples in (76), all words in column A have prenuclear glides and those in column B have no prenuclear glides. The A-B pairs of words in groups (a), (b) and (c) rhyme very well with each other.[38] Obviously, the prenuclear glides are not in the rhyming unit in Mandarin, which is also true in SX. Besides, the prenuclear glides are not TBUs since they are non-moraic (Howie 1976), so that they are not in the nucleus constituent.[39] The nucleus constituent must be fully counted towards the rhyme. In a prosodic unit, a diphthong must be an integral constituent of Nucleus. Thus, any combination of GVG in Mandarin can never be a phonological triphthong; nor can GV phonologically be a diphthong. In all the GV combinations of SX, G is never counted in the rhyme and there is no VG combination in SX, as shown in (75). The evidence from the data strongly suggests that there are no phonological triphthongs in Mandarin and that there are neither triphthongs nor rising diphthongs in the SX surface representation. I propose that only (falling) VG is a real diphthong and that (rising) GV is not a diphthong in Mandarin or any other Chinese dialect.

2.3.5.2 VV

In SX, there is only one VV form in complex final combinations, which is [ɑɒ]. It is very frequently found in syllables with an initial consonant, such as the following:

(77)	[sɑɒ^{35}]	'little/few'
	[pɑɒ^{52}]	'wrap'
	[ɕjɑɒ^{35}]	'small'
	[dʑjɑɒ^{31}]	'bridge'
	[ŋɑɒ^{31}]	'tolerate'

[38] This is the Chinese poetic rhyming pattern.

[39] The position of prenuclear glides in the syllable structure will be discussed in more detail in chapter 4.

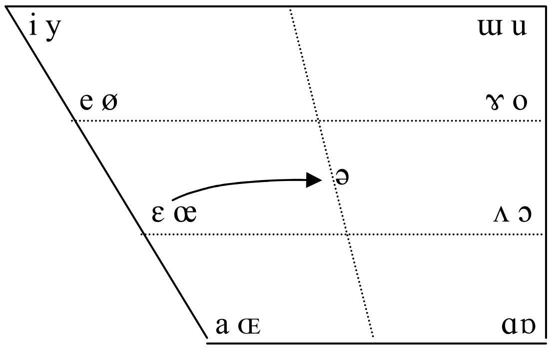
VV combinations as the nucleus of a syllable can be found in many other languages such as [ɛə] in English (Heffner 1949) and [ɔɐ] and [eɐ] in German (Rohler 1999, from International Phonetic Association, Corporate Author International Phonetic 1999). In terms of diphthongs, one element of the combination is always a glide in GV or VG, according to Trask's definition, as discussed in the previous sub-section. However, VV combinations such as [ɛə], [ɔɐ] and [eɐ] are also regarded as diphthongs, because there is a vowel quality change or tongue movement from the first vowel to the second vowel and both vowels form a single nucleus of the syllable in English and German, respectively. More specifically, they are falling diphthongs since the second, schwa-like, element has the weaker intensity. The consideration of [ɛə], [ɔɐ] and [eɐ] as diphthongs challenges Trask's diphthong definition (ii). Moreover, whether [ɑɒ] in SX is a diphthong remains a question.

Before answering this question, I first introduce a well-attested principle: the Obligatory Contour Principle (Leben 1973; Goldsmith 1976; McCarthy 1986; Yip 1988; Wiese 1997):

(78) Obligatory Contour Principle (OCP)
Adjacent melodic autosegments cannot be identical.

The validity of the OCP in (78) has been confirmed in numerous studies, not only for features of tone but also for segmental features (see (80) below)). It is assumed that every feature occupies a tier of its own, ultimately associated with a root node, R, which encodes, by bundling temporally co-occurring segmental features, the notion of a segment and is also situated on a tier of its own (Wiese 1997). If the features of the two adjacent segments of a diphthong are the same, there will not be any tongue movement, which should begin "with one vowel quality and changes more or less smoothly to a second quality (Trask 1996)." [ɛə] in English is a diphthong in which the tongue begins from front mid and changes to central mid, as is illustrated in the vowel chart in (79):

(79)



Singh & Singh (1976) propose an important criterion for diphthong formation in that "diphthongs involve an appreciable amount of tongue body movement within the perimeter of one syllable". However, [ɑɒ] in SX is composed of [ɑ] and [ɒ], which are identical in height and position and only differ in [±round] so that they are regarded as a pair of vowels (cf. (79)). There are generally eight pairs of such vowels, as shown in (79), none of which is used or regarded as a diphthong in the world's languages according to Ladefoged and Maddieson (1996). Casali (1996:40) claims that V_1 and V_2 in a diphthong must differ in at least two features of dimension (height and position) and regards sequences like /ae/ and /ao/ as ill-formed because the two vowels are "not sufficiently distinct from each other". I argue against Casali's claim that if there is a place movement between V_1 and V_2 in a vowel combination and if the combination is the integral nucleus of a syllable, it is a well-formed diphthong. In my view, /ae/ and /ao/ are diphthongs if they occur as a single nucleus in a language since there is a tongue movement in their articulation. However, accepting /ae/ and /ao/ as diphthongs will challenge Trask's diphthong criterion (iii). With the [ɑɒ] combination, no tongue movement is involved in its articulation. It is well attested in the literature of the world's languages that a well-formed diphthong is always faithful to the OCP in terms of features of height and position and that violation of the OCP always fails to produce a well-formed diphthong (Casali 1996; Ladefoged & Maddieson 1996; Trask 1996; Wiese 1997). For example:

(80) a. [ɛə] R R
+front –front

b. [ai] R R
–high +high

c. *[iy] R R
+front +front
(*or* +high +high)

d. *[ɑɒ] R R
+back +back
(*or* +low +low)

The examples in (80) show that the OCP must be observed in the formation of a diphthong. We conclude that [ɑɒ] is not a diphthong in SX, but a true VV combination. What matters whether [ɑɒ] is a diphthong or not is a matter of a vowel system of SX. The exclusion of [ɑɒ] as a diphthong stipulates the *DIPHTHONG constraint in SX, which helps to decide that the merged [y] is from [wi] or [ju] rather than [uj] or [iw]. The latter is a well-formed diphthong and does not violate OCP (H) (see (83) below), so that they will not merge into [y], as it is in Mandarin.

2.3.5.3 VC

In this subsection, I will give a brief introduction to all the surface VC forms which constitute the rhyme of a syllable in SX. There are altogether ten such complex finals in VC structure in SX. They are [əŋ], [ɪŋ], [aŋ], [ɒŋ], [oŋ], [oʔ], [aʔ], [əʔ], [ɪʔ], [ɛʔ]. In these ten surface VC combinations, the final C is always either one of the two consonants: the velar nasal /ŋ/ or the glottal stop /ʔ/. This satisfies CODA-COND in (16) which says that the coda in the SX syllables can only be [ʔ] or [ŋ]. I list some examples of CV combinations in (81):

(81)	[nəŋ^{31}]	'able'	[ɦəʔ^{3}]	'join'
	[tɕɪŋ^{52}]	'gold'	[zɪʔ^{3}]	'enter'
	[ŋaŋ^{22}]	'hard'	[baʔ^{3}]	'white'
	[zɒŋ^{31}]	'taste'	[pɛʔ^{5}]	'eight'
	[loŋ^{31}]	'dragan'	[moʔ^{3}]	'ink'

The examples in (81) show that the ten VC finals can be divided into two groups: one with five Vŋ combinations, and the other five with Vʔ combinations. The vowels of the two columns of VC combinations are almost the same except for one case in each set, which is [ɒ] in [ɒŋ] and [ɛ] in [ɛʔ]. The distribution of vowels in combinations will be discussed in chapter 3 and the details of the phonotactic constraints of segment sequences in SX will be discussed in chapter 4.

2.3.5.4 GVC

The maximal final combination structure in SX is GVC. This means that a vowel is preceded by a glide and followed by a consonant. There are altogether nine such complex finals: [jaŋ], [waŋ], [wɒŋ], [joŋ], [woŋ], [woʔ], [joʔ], [jaʔ], [wɛʔ]. As with the VC structure, the GVC structure can only have [ŋ] and [ʔ] in the final C position. I list some examples in (82):

(82)	[ɕjaŋ^{52}]	'scent'	[k^{h}woʔ^{5}]	'wide'
	[vwaŋ^{31}]	'horizontal'	[tɕ^{h}joʔ^{5}]	'lack'
	[kwɒŋ^{52}]	'light'	[tɕjaʔ^{5}]	'foot'
	[dʑjoŋ^{31}]	'poor'	[vwaʔ^{3}]	'slide'
	[ɦwoŋ^{31}]	'red'		

The examples in (82) show that there are five GVŋ and four GVʔ finals in SX. In §2.3.4.1 and §2.3.4.3, I have discussed the structures GV and VC in complex finals of SX. GVC could be either a combination of GV plus C or G plus VC. On the basis of the data in (82), I assume that the GVC structure in SX is G plus VC rather than GV plus C, because we have more VC combinations with the same phonemes than combinations of GV with the same phonemes as those in GVC. This assumption is also well supported by the fact that GV in GVC is not an independent single constituent in SX, while VC is the integral rhyming unit in GVC since G is excluded from the nucleus constituent in SX. So, I may formalize: GVC is G + VC. The status of the syllabic structure of G will be discussed in detail in chapter 4.

The data in (82) also show that *[juŋ] and *[wuŋ] do not exist in SX just as there is no *[wu], *[ju], *[wi], or *[ji] in GV combinations in SX. I assume such GV combinations are not acceptable in the SX surface representation because they violate OCP for height features. This suggests a specific OCP constraint on height in SX, as stated in (83):

(83) OCP(H)
Two adjacent segments cannot be identical in height in GV combinations.

In (78), I introduced OCP as one of the criteria of diphthong formation, which mainly involves two vowel parameters, height and position. OCP(H) only involves height. OCP(H) is a surface constraint which only applies to GV combination in surface representation, ruling out *[ji], *[jy], *[ju], *[ɥi], *[ɥu], *[wi], *[wy] or *[wu] on the surface, while underlyingly /ju/ and /wi/ could be possible since we do have /jo/, /je/, /jɤ/, /ja/, /we/, /wo/ and /wa/ in SX, formalized in a combination of GV underlyingly. However, I assume that OCP in both height and position (OCP(H&P)) when applied to GV combination is an underlying constraint which rules out such combinations as */jy/, */ji/ and */wu/. OCP(H) is a remarkable phonological constraint in SX, since [ju], [ɥu] and [wi] are well-formed combinations or sequences in surface representation in some other Chinese dialects such as Mandarin. I list some examples to show how the OCP(H) constraint works in SX for words of the same lexical meaning in the same combination in Mandarin:

(84)

Mandarin	SX	
[tɕju^{35}]	[tɕjɤ^{33}]	'nine'
[ji^{35}]	[ɦi^{31}]	'move'
[ɥu^{214}]	[ɦy^{13}]	'rain'
[k^{h}wi^{55}]	[k^{h}we^{52}]	'loss'

The fact that SX has such syllables as in (84), different from Mandarin, is not a coincidence. Because of the OCP(H) constraint in SX, if there should be any GV combination identical in height underlyingly, some adaptation should be made in surface representation to satisfy OCP(H), as shown in the examples in (84).

There is another remarkable final combination in SX, [jɑɒ], which has remained undiscussed so far. In fact, [ɑɒ] in SX behaves like a long vowel phonologically, though it is a combination of [ɑ] and [ɒ] phonetically, as discussed above, so that [jɑɒ] behaves the same as a GV combination in terms of syllabic constituency. This will be discussed in chapter 3.

2.4 Phonemic [ʔ] in SX

Although the initial glottal stop [ʔ] in SX and other Wu dialects, as well as in other languages in the world, has been a problematic segment in terms of its phonological status in the syllable-initial position of the languages that have it, there seems not to be much discussion over the syllable final [ʔ]. SX has the glottal stop [ʔ] both in the initial and in the final, possible even in the same syllable, as shown in (85):

(85)	[ʔɛʔ^{5}]	'duck'
	[ʔɪʔ^{5}]	'one'

As was discussed in §2.2.2, the initial glottal stop [ʔ] in SX plays the phonetic role of indicating the beginning of an underlying onsetless high-register syllable, rather than any phonological role in the syllable structure. However, the status and properties of the syllable-final [ʔ] differ phonologically from that of the initial [ʔ]: it is phonologically present as the syllable final while it only has phonetic status as the syllable initial. This can be supported by the fact that there are near minimal pairs with and without the final [ʔ] in SX. For example:

(86)	a. [toʔ^{5}]	'supervise'	c. [baʔ^{3}]	'white'
	b. [to^{52}]	'many'	d. [ba^{31}]	'card'

The examples in (86) show that (a)-(b) and (c)-(d) are two near minimal pairs (differing in tones). It is also true that syllables with different tones have different lexical meanings in SX. Tones and segments are the integral syllable constituents which decide the lexical meaning. The syllables of (86a) and (86c) both have the tone feature [h] and have the final glottal stop [ʔ] which is dominated by one mora (the moraic structure of syllables will be discussed in chapter 4). The syllables of (86b) and (86d) both have the tone features [hl] with the tone pitch falling down. It is obvious that the final glottal stop [ʔ] plays a phonological role in deciding the syllable tone feature thus deciding the lexical meaning of the syllable. We therefore assume that it is underlyingly present.

According to *Qieyun*, a book of ancient Chinese phonology, Middle Chinese had [p], [t], [k], [m], [n] and [ŋ] in the coda. Syllables ending in

[p], [t] or [k] fell into one group and had *entering* tone,[40] similar to that of Modern SX. In this way, the syllables with the same tone but different final stops had different lexical meanings, because the different stops in the coda were distinctive phonemes. Some modern Chinese dialects still have [p], [t] or [k] in the coda such as Xiang, Hakka, Min and Cantonese (Zhan 1991). These three final stops – [p], [t] and [k] in Middle Chinese – merged into the glottal stop [ʔ] in the coda in modern Wu dialects through historical attrition. However, it is believed that the disappearance of the final stops had a significant influence on the tones in Wu dialects (Zhan 1991; Cao 1998, 2002, 2004). In one way or another, the final glottal stop [ʔ] still has a similar phonological function to that of the final [p], [t] or [k] in Middle Chinese. It is widely agreed that Chinese underwent a process by which the voicing distinction on initial consonants was transformed into a tonal distinction, doubling the number of tones (Hombert 1978; Yip 2002, also see above). I assume that tones developed as a result of the loss of some phonemes originally. Much cross-linguistic evidence shows that a syllable-final consonant influences the tones on the preceding vowels (Haudricourt 1954; Pulleyblank 1962; Wang 1963; Hombert 1978; Baxter 1992; Chen 2000). Hombert (1978: 92) also finds that the effect of a glottal stop on the pitch of the preceding vowel is widely attested, e.g. in Vietnamese, Burmese, and Middle Chinese.

On the basis of the diachronic and synchronic studies, I assume that the syllable-final glottal stop [ʔ] (which results from the debuccalization[41] of the syllable final stops [p], [t] and [k] in SX as well as other Wu dialects) has phonological status in the syllable structure, making [toʔ] and [to] or [baʔ] and [ba] a pair of different lexical items. The details of the debuccalization of the syllable-final stops will be discussed in chapter 4.

[40] The term *entering* is one of the four tones, viz. *even*, *rising*, *going* and *entering*, which are translated from the traditional Chinese terms, *ping*, *shang*, *qu* and *ru*, respectively. The *entering* tone only occurs on the syllables ending with a stop in Chinese. The Chinese tonology will be discussed in chapter 5.

[41] Debuccalization, also called deoralization, is a phonological process in which a consonant segment loses its oral articulation (see Humbert 1995; Trask 1996). The debuccalization of the SX syllable-final stops and nasals will be discussed in chapter 4.

2.5 Summary

Through phonetic observation and phonological analysis of the consonants and vowels of SX in surface representation, focusing on their distribution, phonetic and phonological realizations, constraints and rules, I presented the total of 29 surface consonants in SX, as shown in (35), and presented the feature specifications for the 29 consonants in (36). I have also presented a clear picture of all the 48 Finals, as shown in (41), in the forms of syllabic C, V, VV, GV, VC, and GVC, which remarkably excludes a combination of VG in SX. That is, there are no postnuclear glides in SX syllable structure while VG is, as a matter of fact, a very common segment sequence in many of the world's languages, including Mandarin and other Chinese dialects. This chapter also presents 14 surface vowels in SX, as shown in the vowel chart of SX in (87):

(87) The Vowel Chart of SX

	Front		Central		Back		
Close	i	y				u	High
	ɿ						
	ɪ						
Mid	e			ɵ	ɤ	o	Mid
			ə				
	ɛ						
Open							Low
	a				ɑ	ɒ	
	–round	+round	–round	+round	–round	+round	

The Vowel Chart (87) shows that there are altogether 14 vowels in SX which appear in surface representation in the rhyme of syllables either by themselves or in combinations. Besides the 14 vowels, there are also three glides: [j], [ɥ] and [w], in surface representation. However, among these 14 surface vowels, I assume that there are only six phonemic vowels underlyingly. These are /i/, /e/, /a/, /ɤ/, /o/ and /u/. The discussion of the 14 vowels, their phonetic and phonological behaviour, and why only these six vowels are the underlying phonemes in the vowel inventory of SX, will be the topic of the next chapter.

3 The Underlying Vowel Inventory of Shaoxing

3.1 Introduction

All languages have an inventory of (possibly abstract) sound categories with which words are represented, referred to as 'phonemes' or 'segments'. These segments will be phonetically manifested differently in different phonetic contexts, due to both universal and language-specific factors (Goldsmith 1995: 2). It has often been observed that the typical Chinese language has a large number of allophones in complementary distribution, which can therefore be derived from a smaller number of phonemes (Yip 1996). This chapter will discuss the underlying vowel inventory in SX and the distribution of these vowels, to account for the nature of the basic units of speech sounds and the relationships between these units and their contextual variants. After presenting an analysis of the distribution of the 14 surface vowels of SX, I argue that the underlying vowel inventory of SX includes only six phonemic vowels: /i u e ɤ o a/, and thus constitutes a preferred vowel inventory among the world's languages that have a six-vowel system.

3.2 The Arrangement of Surface Vowels

Crothers (1978) presents a study of vowel inventories in the world's languages and formalizes general patterns in vowel systems, such as the following (de Boer 2001: 90):

a. The number of height distinctions in a system is typically equal to or greater than the number of backness distinctions.
b. Languages with two or more central[1] vowels always have a high vowel.
c. The number of vowels in a column of central vowels cannot exceed the number of vowels in the front or back column.
d. The number of height distinctions in front vowels is equal to or greater than the number in back vowels.

[1] In de Boer (2001), the original word is 'interior', which is replaced by 'central' to fit in with the common description in general linguistics.

Maddieson (1984a) presents a systematic, statistical investigation of the 317 languages in UPSID and makes some interesting generalizations, e.g. that front vowels are usually unrounded (94.0%), and back vowels usually rounded (93.5%). This coincides with Jones' (1968) primary and secondary cardinal vowels; high front vowels are more frequent than high back vowels. According to Maddieson's (1984a) segmental analysis (based on 317 languages in UPSID), in a generally symmetric vowel system, there are obvious asymmetries, such as the one that vowels in the mid range are more common than high vowels; low vowels are substantially less common, amounting to only 20.5%; central vowels are considerably less common, amounting only to 22.2%; unrounded vowels are considerably more frequent than rounded vowels, namely 61.5% vs. 38.5%.

As was discussed in the previous chapter, there are 14 surface vowels in SX, including [ʅ i ɪ y e ɛ ɵ ə ɤ u o a ɑ ɒ], which, according to the major three vowel parameters of position, height, and rounding (Ladefoged & Maddieson 1996: ch. 9), can be classified as follows:

(1)	Front		Central		Back		Total
	−round	+round	−round	+round	−round	+round	
High	[i], [ɪ] [ʅ]	[y]				[u]	5
Mid	[e] [ɛ]		[ə]	[ɵ]	[ɤ]	[o]	6
Low	[a]				[ɑ]	[ɒ]	3
Total	6	1	1	1	2	3	14
	7		2		5		

The table in (1) presents the 14 surface vowels in SX, and shows that there are five high vowels, the same number as that of all back vowels; among the five high vowels, four are front while only one is back; there are two central vowels, i.e. fewer than front or back vowels; the number of high vowels is about 1.7 times more than that of low vowels; one out of seven of the front vowels is rounded and two out of five of the back vowels are unrounded. Although the arrangement of the 14 surface vowels of SX in table (1) gives a rather asymmetrical picture, it still follows the general pattern of vowel systems of the world's languages as observed by Crothers (1978) and Maddieson (1984a), as mentioned above. However, these 14 vowels in (1) all belong to the surface representation of SX. Vallée (1994, cited from de Boer 2001),[2] who investigated the UPSID, found that the maximum number of different vowel qualities used in any language in the sample is 15. SX seems to have almost the maximum number of vowels in a vowel system, though there is still a lot of room in both the articulatory and the acoustic space. In the next section, we examine the vowel system from a phonological perspective.

3.3 The Vowel Phonemes of Shaoxing

De Boer (2001) agreed, on the basis of his vowel simulation with standard parameter settings and through optimisation, that the most common vowel system (88% of the 49 vowel systems he analysed) is a symmetrical five-vowel system with /i u e o a/. The result of his vowel simulation is consistent with Maddieson's (1984a) statistics of the 317 languages in UPSID, from which he also concluded that the most common number of vowel phonemes in a language is five, and the most common number of distinctive vowel qualities in a language is also five, viz. /i u e o a/. In this chapter, I present my analysis of the vowel distribution in SX and I claim that out of the 14 surface vowels only six are underlying, phonemic vowels, viz. /i u e ɤ o a/, which is very similar to the widely attested and most common vowel system of the world's languages.

There are three principles which have to be taken into consideration when determining the underlying segment inventory of a language: (a) which allophone has the widest distribution; (b) which allophone most appropriately represents the phonetic range of variation of all allophones; (c) which allophone is the one from which other allophones can be most simply and naturally derived (Maddieson 1984a: 163). In this section I will present an analysis of the distribution and the phonological behaviour of these 14 surface vowels, aided by an extensive OT analysis, so as to identify which vowels are the underlying phonemes and explain how the 14 surface vowels can be reduced to only six phonemic vowels underlyingly. Yip (1996: 757) points out that 'to derive rich surface inventories from more parsimonious underlying inventories, it was necessary to postulate abstract underlying forms'. It is common cross-linguistically that the underlying representations (UR) can be very

[2] There are languages that have more vowel phonemes, but these will use other processes, such as length, nasalization, and pharyngealization, not quality, in order to distinguish vowels.

abstract from the surface representations (SR). To establish certain relations between the abstract UR and the rich SR, rule-based theories are always inefficient, because the rule: A → B /__C does not tell why A will not become D, E or F when they exist in SR, which frequently requires more separate rules. In this case, a constraint-based approach is more economical, in which any abstract form can be the input and a set of outputs is generated for each input, and inspected by the ranked constraint set (see Yip 1996). However, the rule is also included in this chapter to describe the change that takes place. Sometimes, I use both the rule to describe how A becomes B and OT to explain why A becomes B not D, E or F.

This section presents an OT analysis of the relations between the abstract UR of the six vowels and the rich SR of the 14 allophones in SX. According to the three criteria mentioned above, i.e. their phonetic and phonological behaviour, and their distribution, the 14 surface vowels will be divided into four classes in the analysis: high front vowels, high back vowel, mid vowels, and low vowels.

3.3.1 High front vowels

Table (1) shows that in SX, there are four high front vowels in surface representation. They are [i], [y], [ɪ], and [ɿ], among which [ɿ] is an apical vowel, as was discussed in chapter 2. All these high front vowels can occur in the rhyme, either by itself of in a combination. Consider the following examples:

(2)	a.	[sɿ^{33}]	‘try’	[dzɿ^{31}]	‘late’
	d.	[tɕ^{h}i^{33}]	‘go’	[pi^{35}]	‘compare’
	c.	[ɦy^{13}]	‘rain’	[ʑy^{22}]	‘tree’
	b.	[mɪŋ^{22}]	‘life’	[zɪʔ^{3}]	‘enter’

The examples in (2) show that of the four high front vowels, three occur in open syllables and one in closed syllables in SX. It is remarkable that there are four high front vowels in one language. English has only two front high vowels: [i] and [ɪ] (Heffner 1949). However, these four high front vowels are in complementary distribution. I will argue that only /i/ is an underlying vowel in SX, and I will explain how the other three allophones are derived.

3.3.1.1 Distribution of high front vowels

First of all, front vowels are usually unrounded in most languages (Ladefoged & Maddieson 1990), so that rounded front vowels are always less frequent and less common in distribution in any language. The examples in (2) may suggest how the four high front vowels in SX are distributed. For example, [i] and [ɿ] cannot occur after the same consonant and neither can be followed by a consonant. In fact, the phonetic and phonological behaviour of [i] and [ɿ] and their distribution were discussed in chapter 2, where I postulated a phonological rule (see (65), ch.2) as in (3):

(3)
$$/\text{i}/ \rightarrow [\text{ɿ}] \ / \ \begin{bmatrix} +\text{cons} \\ +\text{apical} \end{bmatrix} \_\_$$

The rule in (3) shows that /i/ is realized as [ɿ] when preceded by a [+apical] consonant in SX, indicating that [ɿ] is an allophone of /i/. This is also supported by the examples in (2). The consonants in (2a) are [s] and [dz], which are both dental sibilants, specified as [+apical], as discussed in chapter 2. [apical] is not an underlying feature, but a phonetic feature in SX. Usually, coronal sibilants (which include fricatives and affricates) can be classified into apical and laminal (see Ladefoged & Maddieson 1996: 164), as shown in (4):

(4) Types of sibilants:

a.	dental:	apical	e.g. [tsɿ], [ts^{h}ɿ], [dzɿ], [sɿ], [zɿ]
b.	post-alveolar:	laminal	e.g. [tʃi], [dʒi], [ʃi], [ʒi]
c.	alveolo-palatal:	laminal	e.g. [tɕi], [tɕ^{h}i], [dʑi], [ɕi], [ʑi]
d.	retroflex:	apical	e.g. [tʂʅ], [tʂ^{h}ʅ], [ʂʅ]

Types of sibilants (4) show that dental and retroflex sibilants are apical (see also Bright 1978) and apical sibilants precede apical vowels which complementarily distribute with non-apical high front vowel /i/. As was discussed in chapter 2, apical vowels are produced with the tongue in essentially the same position as in the corresponding sibilants. Thus, apical vowels must be different from the different articulators of the preceding apical sibilants, as shown in (4a) and (4d). SX has only dental apical fricatives and affricates, so that it has only one apical vowel [ɿ]. Mandarin has both dental and retroflex apical fricatives and affricates, so it also has the apical vowel [ʅ]. [s] and [z] in English are laminal alveolar (see Ladefoged & Maddieson 1996: 164), so that they allow such syllables as [si] and [zi].

Because apical sibilants are usually specified as [+strident], the following apical vowels also sound strident acoustically, which is realized by spreading a manner feature of the preceding consonant to the following vowel. Ladefoged & Maddieson (1996) refer to an apical vowel as strident. Strident vowels have a constriction between the part of the tongue below the epiglottis and the tips of the arytenoid cartilages in the upper part of the larynx. This constriction results in these vowels having a specific phonation type. Traill (1985) suggests that the strident vowels may be regarded phonologically as pharyngealized breathy voiced vowels. However, such a phonation type has a certain commonality cross-linguistically, e.g. in the Caucasian languages and the Khoisan languages (Traill 1985). The phonation type of strident vowels in SX resembles vowel devoicing in many languages. For example, in French, [i] is devoiced after a voiceless obstruent in the onset; in Japanese there is a contrast between the voiceless allophones of /i/ and /u/ between voiceless obstruents, as in [ki̥ʃi] 'shore' and [ku̥ʃi] 'comb'. In short, the apical vowel [ɿ], which has a strident phonation type, is an allophone of /i/ when preceded by an apical sibilant in SX, as shown in (3).

The examples in (2) also show that [ɪ] occurs in the rhyme of a syllable only when in combination with [ɪŋ] or [ɪʔ], whereas [i] cannot occur in either combination. Therefore, [ɪ] is also in complementary distribution with [i] and is likely to be an allophone of /i/. As was discussed in chapter 2, Yang and Yang (2000) assume that SX has no nucleus vowel [ɪ] in surface representation and that [ɪŋ] and [ɪʔ] do not occur either, but [iəŋ] and [iəʔ] occur instead, which are regarded as the finals Middle Chinese had (see Chao 1928). As a matter of fact, Modern SX did go through considerable phonological changes, especially in its rhymes, having lost many coda consonants such as [m], [p], [t] and [k], and having more simple rhymes instead of complex rhymes (Chao 1928). Systematically speaking, since modern SX has such final combinations as [ia], [ie], [io], [iɤ], [aŋ], [uŋ], [əŋ], and [ɒŋ], it is also likely to have /iŋ/ in its underlying system of syllable structure, or any of its allophones in such a combination. Thus, there is stronger motivation to allow for [ɪŋ] and [ɪʔ] rather than [iəŋ] or [iəʔ] in the surface SX final combinations. What is more important is that we do hear such rhymes as [ɪŋ] and [ɪʔ] in modern SX, as shown in (2). In conclusion, I claim that there is a surface high front vowel [ɪ] which only occurs in combinations of [ɪŋ] and [ɪʔ] in SX and is also an allophone of phonemic /i/. Its distribution can be formulated in the following rule:

(5) /i/ → [ɪ] / __ C

As was mentioned in chapter 2, vowel length is underspecified in SX, unlike Thai or Japanese (Rosner 1994). Instead, vowels in SX can be phonologically specified with [tense] as follows:

(6)	i	y	ɪ	ɿ	u	e	ɛ	ə	ɵ	ɤ	o	a	ɑ	ɒ
[tense]	+	+	−	+	+	+	−	−	−	+	+	+	−	−

Articulatorily speaking, [+tense] vowels are usually longer than [−tense] vowels. All the [+tense] vowels shown in (6) are bimoraic and can occur in stressed open syllables, which is required by the tonal system of SX. The well-formed syllable [ɪʔ] also satisfies the tonal system of SX, in which syllables ending with the glottal stop [ʔ] have *entering* (high level) tones, [5] or [3], differing in register, which are phonetically short but phonologically still bimoraic if stressed, since the syllable-final [ʔ] is also moraic in SX.[3] As a result, [ɪ] is licensed when followed by a consonant like other [−tense] vowels such as [ɛ] and [ə] in [ɛʔ] and [əʔ], respectively, which suggests that [−tense] vowels have to be followed by a consonant.

The examples in (2) show that [y] can also occur alone as the rhyme and can contrast with six phonemic vowels in certain environments, as shown in (7):

(7) a.[ɦi^{31}] 'move'
b.[ɦu^{31}] 'lake'
c.[ɦe^{22}] 'harm'
d.[ɦɤ^{31}] 'attend'
e.[ɦo^{31}] 'river'
f.[ɦa^{31}] 'shoe'
g.[ɦy^{31}] 'surplus'

---

3 The weight status of the syllable-final stop differs from Chinese dialect to dialect. In Cantonese, the syllable-final stop is weightless when it follows a long vowel, e.g. [taːp] 'pile' (see Yip 1996, 2002).

The examples in (7) give rise to the question whether [y] should be regarded as a phonemic vowel, just like the other six vowels, since the seven syllables above are minimal pairs in surface representation. This is a difficult topic. As the tables of the Finals by Chao and Campbell (see (38) and (39) in chapter 2) show that [y] cannot occur alone as a Final but occurs in combinations such as [yɥ] in both tables. Accordingly, the syllable in (7g) should be [ɦyɥ^{31}], which, I argue, is unacceptable in the SX surface representation because both [y] and [ɥ] are [+high, +front, +round] so that [yɥ] badly violates the OCP. I assume that [y] is only a surface vowel and its underlying form could be either /wi/ or /ju/ since we have two glides [j] and [w] in GV combinations in SX such as [wa], [we], [wo], [ja], [je], [jɤ] and [jo], which were discussed in chapter 2. As was also discussed in chapter 2, SX has GV combinations but not VG. In most of the world's languages, usually GV is a rising diphthong and VG is a falling diphthong, which means SX has rising combinations but no falling combinations. If there is any underlying diphthong in SX, it is most likely to be /iu/ rather than /ui/, because the former is a rising combination but the latter is a falling combination, according to the sonority scale (Durand 1990). However, as a GV combination, there may be /ju/ and /wi/ underlyingly in SX. I assume that [y] is the result of segment merger of /ju/ or /wi/ in surface representation in SX.

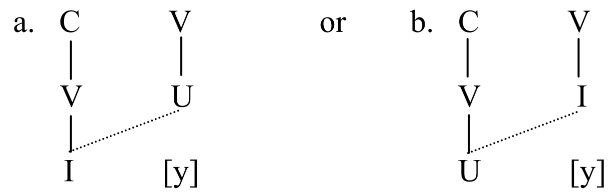
3.3.1.2 Segment merger

Merger is a phonological change in which a previously existing contrast between two or more phonemes is lost. There are two types of merger: a merger applying only in restricted contexts, thus introducing a neutralization, is a conditioned merger, and one which applies in all contexts, thus reducing the number of phonemes in the language, is an unconditioned merger (Trask 1996).

There is strong phonological motivation why underlying /ju/ or /wi/ should merge into [y] in surface representation in SX. First, as was discussed in chapter 2, OCP(H) (see (83), ch.2) rules out *[+high][+high] combinations in the SX surface representation, so that either [ju] or [wi] is not acceptable while [je], [jo], [ja], [jɤ], [wo], [wa], [we], [wɛ̃], and [wə̃] are well-formed in SX. Secondly, /j/ and /w/ are both specified for [−cons] and excluded from the onset position in SX (syllable structure in SX will be discussed in chapter 4), and each segment of the GV combination (/ju/ or /wi/) mainly differs in backness and roundedness. During the process of merger, /j/ or /i/ becomes [y] when rounded; or /w/ or /u/ becomes [y] when fronted, which is a result of merging the features of [−back] and [+round], formulated in such a rule as follows:

(8) $$\left.\begin{array}{l}/\text{j}/\text{ or }/\text{ i}/\\ /\text{w}/\text{ or }/\text{u}/\end{array}\right\} \rightarrow [\text{y}] \left\{\begin{array}{l} /\ [\underline{+\text{round}}]^{4} \\ /\ [-\underline{\text{back}}]\end{array}\right.$$

The rule in (8) can also be expressed by element structure (as in Dependency Phonology, cf. ch.1), as shown in (9):

(9)


The element structure in (9) shows how the two elements of either /j/ and /u/ or /w/ and /i/ merge into [y]. This can be formulated in a simple rule, as shown in (10):

(10) $$\left.\begin{array}{l}/\text{ju}/\\ /\text{wi}/\end{array}\right\} \rightarrow \ [\text{y}] \quad / \quad [\ ]$$

The rule in (10) says that underlying /ju/ or /wi/ merges into [y] in surface representation to avoid the violation of the OCP(H) in GV combination. More examples are given in (11):

(11)

[tɕʰy^{35}]	'fetch'
[dʑy^{22}]	'live'
[pɪʔ^{5}ɕy^{33}]	'must'
[ɦy^{22}mɪŋ^{31}]	'fisher'
[ʔy^{52}ɕjoʔ^{5}]	'blood stasis'

In short, [y] is not an underlying vowel, but a merged vowel in surface representation, resulting from the neutralization of a combination, /ju/ or /wi/ in SX.

---

[4] The underlines in (8) mean where the underlying vowels or glides are.

3.3.1.3 Phonemic /i/

On the basis of the analysis presented above, I claim that of the four high front vowels in SX, only /i/ is a phonemic vowel. Both [ɿ] and [ɪ] are allophones of /i/, which can be formulated as follows:

(12)

$$/\text{i}/ \rightarrow \begin{cases} [\text{ɿ}] \;/\; [+\text{cons}, +\text{apical}] \;\_\_ & \text{(a)} \\ [\text{ɪ}] \;/\; \_\_\,\text{C} & \text{(b)} \\ [\text{i}] \;/\; \text{elsewhere} & \text{(c)} \end{cases}$$

The rule in (12) shows that the three high front vowels are in complementary distribution in (a), (b) and (c) and that only /i/ is a phonemic vowel. The distribution in (12a) was discussed in chapter 2 and above in this subsection. We agreed that [ɿ] is an apical vowel with an articulation of the tongue in essentially the same position as in the corresponding apical sibilants. This strongly suggests a constraint that when the onset consonant is an apical sibilant, the following high front vowel will have the same value of [+apical], as stated below in (13):

(13) AGREECV[apical]
An apical consonant must agree with the following high front vowel in value for the status of apical.

AGREECV[apical] in (13) stipulates that [ɿ] in SX only occurs after apical dental consonants which include [ts ts^h dz s z] according to the types of sibilants proposed by Ladefoged & Maddieson (1996) (see also Williamson 1977; Bright 1978). The constraint AGREECV[apical] rules out such syllables as *[tsi], *[ts^{h}i], *[dzi], *[si] and *[zi] in SX.

The rule in (12b) shows that /i/ becomes [−tense] when followed by a consonant, which suggests a simple constraint: *[i]C.

3.3.2 The high back vowel

It is believed that all languages have /i u a/ (Maddieson 1984a; Ladefoged & Maddieson 1990, 1996). Naturally, SX also has these vowels. Compared with the four surface high front vowels in SX as shown in table (1), there is only one high back vowel [u], even in surface representation. The proportion of 4:1 between high front and high back is quite a striking asymmetry, which is very rare in the languages covered in Maddieson (1984a). In this subsection, I will discuss the distribution of /u/ and its possible allophone(s) and the phonological motivation for postulating a single high back vowel.

Unlike the high front phonemic vowel /i/ which has two allophones in complementary distribution, /u/ does not have any allophone in open syllables, so it has a wider distribution, as shown in the examples in (14):

(14) a. [pu^{35}] 'compensate'
b. [tu^{35}] 'block'
c. [su^{35}] 'count'
d. [ku^{35}] 'old'
e. [hu^{35}] 'fire'
f. *[ɕu]

The examples in (14) show that [u] can occur after many different initial consonants as a nucleus vowel, but not after alveolo-palatal consonants, including [tɕ], [tɕ^{h}], [dʑ], [ɕ], [ʑ], and [ȵ]. This suggests such a constraint in SX that [u] cannot occur after alveolo-palatal consonants, as stated in (15):

(15) *ALV-PAL[u]
[u] cannot occur after alveolo-palatal consonants.

In fact, not only is [u] disallowed after alveolo-palatal consonants, but no vowels except [i], [ɪ] and the glide [j] can occur after alveolo-palatal consonants. These consonants share the same specifications of [+high] and [−back] with the high front vowels, according to the *SPE* feature system (see Chomsky & Halle 1968). These two feature specifications for the alveolo-palatal consonants can be proved by the nasal palatalization rule that says the alveolar nasal [n] becomes the alveolo-palatal nasal [ȵ] when followed by a high front vowel. Palatalization results from spreading the feature of [+high] and [−back], so that [ȵ] is specified as [+high] and [−back]; so are the alveolo-palatal fricatives and affricates. Therefore, the distribution of alveolo-palatal consonants can be formalized in a constraint in (16):

(16) AGREECV[+H, −B]
A [+high, −back] consonant must agree with the following vowel in value for the features of [+high] and [−back].

However, the apical vowel [ɿ] is also specified as [+high, −back], because it is an allophone of /i/. But [ɿ] cannot occur after alveolo-palatal consonants, which suggests a constraint ranking that AGREECV[apical] dominates AGREECV[+H, −B] so that [ɿ] can only occur after [ts], [tsʰ], [dz], [s] and [z] but never occur after an alveolo-palatal, which can only precede [i], [ɪ] and [j] for their agreement in value of [+high] and [−back] between the onset consonants and the nucleus vowels.[5]

As was discussed in the previous subsection, we assume a /ju/ or /wi/ combination in SX underlyingly, which, however, violates the surface constraint OCP(H) so that /ju/ or /wi/ merges into [y]. [y] is also a [+high, −back] vowel and can also follow the alveolo-palatal consonants.

3.3.3 Mid vowels

3.3.3.1 Introduction

In the surface representation of SX, there are six mid vowels: [e], [ɛ], [ə], [ɵ], [ɤ] and [o], which, according to the place parameter, can be divided into three categories: front mid vowels [e] and [ɛ], central mid vowels [ə] and [ɵ], and back mid vowels [ɤ] and [o], two for each place, in a very symmetrical system. Mid vowels in SX share more in common with many other languages than the high front vowels, among which are the remarkable apical vowel [ɿ] and a merged [y] in surface representation, as was discussed previously. Among the six mid vowels, only the rounded central vowel [ɵ] and the unrounded back vowel [ɤ] are uncommon in the world's languages. According to Maddieson (1984a), among the 317 languages in UPSID, there are only five languages that have phonemic /ɵ/ and four languages that have phonemic /ɤ/. In this subsection, I will present my analysis of the distribution of the six surface mid vowels in SX and I assume that among these six surface mid vowels only /e/, /ɤ/ and /o/ are phonemic. My analysis in this section is mainly based on OT theories, in which any underlying representation will give the right output, if the phonotactic constraints outrank FAITHFULNESS (Yip 1996). The form of the phonotactic constraints is driven by the observed surface forms in SX.

---

[5] [high] and [back] are not used for the consonant feature specifications in §2.2.5 in chapter 2. According to *SPE* (Chomsky & Halle 1968), post-alveolar, alveolo-palatal, retroflex, palatal and velar consonants are [+high]; velar and glottal consonants are [+back]. Thus, in SX, [+high, −back] consonants are only alveolo-palatal consonants, including [tɕ tɕʰ dʑ ɕ ʑ ȵ] (see (35) in chapter 2).

3.3.3.2 Major-feature constraints

It is usual to analyse Mandarin as having an underlying four vowel system of /i u a ə/, and to derive the mid vowels [e ɛ o ɔ] by spreading frontness or rounding from adjacent segments (Chao 1934). According to Chao (1934), Mandarin has rich surface mid vowels, excluding [e] and [o] as phonemic vowels (the discussion of Mandarin vowel system is outside the scope of my dissertation). In contrast, SX has more phonemic mid vowels, although some underlying representation forms can be very abstract from the surface forms. Consider the following examples:

(17)	a. [e]:	[de^{31}]	'lift'	[tze^{33}]	'vegetable'
	[ɤ]:	[dɤ^{31}]	'head'	[tzɤ^{33}]	'bad smell'
	[o]:	[do^{31}]	'take'	[tzo^{33}]	'wrong'
	b. [ɛ]:	[tɛʔ^{5}]	'build up'	[pɛ̃^{52}]	'class'
	[ə]:	[təʔ^{5}]	'get'	[nəŋ^{31}]	'able'
	[ɵ]:	[tɵ̃^{33}]	'stew'	[huɵ̃^{33}]	'happy'

The examples in (17a) show that [e], [ɤ] and [o] can stand alone as the rhyme and can occur after the same initial consonant and with the same tones, which suggests that [e], [ɤ] and [o] are all contrastive with each other and thus are presumably phonemic vowels (the contrastive distribution of [e], [ɤ] and [o] with different initials will be presented in table (54) in chapter 4). The examples in (17b) show that [ɛ], [ə] and [ɵ] do not stand alone or in oral contrast as the rhyme in the syllables. Instead, they are the rhyme only either when nasalized or when followed by a consonant. In chapter 2, I discussed vowel nasalization and the VC structure. That discussion showed that there are only three nasalized vowels: [ẽ], [ɛ̃] and [ɵ̃], which are contrastive with each other, as shown by the examples in (18):

(18) a. [dzẽ^{31}] 'sink'
b. [dzɵ̃^{31}] 'pass on'
c. [dzɛ̃^{31}] 'disabled'

It was also discussed in chapter 2 that nasalized vowels only occur in surface representation, as shown in (18), and the underlying syllable structure of the nasalized vowels is assumed to be /VN/ (a vowel followed by a nasal) underlyingly, which was accounted for by the nasalization rule (see (71) and (72), ch.2). However, this leaves as yet unanswered the question what the underlying vowels are for the underlying /VN/ structure of the syllable rhyme.

We have postulated two phonemic high vowels: /i/ and /u/ in the two previous subsections; we have proposed three phonemic mid vowels: /e/, /ɤ/ and /o/, since these three mid vowels are in contrastive distribution (as mentioned above) and they can occur in open syllables after many different onset consonants; we also set up the low vowel /a/ as a phonemic vowel since it is believed that all the world's languages have /i u a/ (Maddieson 1984a; Rosner 1994; Ladefoged & Maddieson 1990, 1996). As a result, we have six phonemic vowels: /i/, /u/, /e/, /ɤ/, /o/, and /a/, as I suggested previously. Since there are five underlying GV combinations, viz. /ja/, /je/, /jɤ/, /jo/ and /ju/ ([y] in surface) in SX, I assume that the underlying vowels in /VN/ combinations of the nasalized vowels might be underlyingly represented as /iN/, /uN/, /eN/, /ɤN/, /oN/ and /aN/.[6]

I will present an OT analysis of the three nasalized vowels in the SX surface representation. If it is true that surface representation is derived from underlying representation by spreading certain features (Chao 1934), I invoke the well-established constraint IDENT-I/O(F) (Pulleyblank 1996; Kager 1999; Yip 2002) in order to formalize the relations between the allophonic vowels and the underlying phonemes in SX. I divide IDENT-I/O(F) into three different specific constraints according to the three major vowel parameters, as follows:

(19) IDENT-BACK
Input-output identity for the feature [back].

(20) IDENT-HIGH
Input-output identity for the feature [high].

(21) IDENT-ROUND
Input-output identity for the feature [round].

Before I can work out a constraint ranking, let us return to the arrangement of 14 surface vowels of SX in (1) and the vowel chart of SX (see (87), ch.2), both of which show that SX has more distinctions along

[6] The final /N/ is nasality. The possible phonemic nasals for /N/ will be discussed in detail in chapter 4.

the height dimension than along the front-back dimension. For example, all else being equal and only varying the height parameter, there are six front vowels from the highest point to the lowest, viz. [i], [ʅ], [ɪ], [e], [ɛ] and [a], while there are only two vowels going from front to back, as in the pairs of [y] and [u], [e] and [ɤ], [ɵ] and [o], or [a] and [ɑ] if only the parameter of position is changed with the other two parameters unchanged. Even in the primary Cardinal Vowel system described by Jones (1975), there are [i], [e], [ɛ] and [a] with four height levels while there are only [i] and [u], [e] and [o], [ɛ] and [ɔ], or [a] and [ɑ] with two positions at the same level. Both the surface vowel inventory of SX and the primary Cardinal Vowels suggest that the height dimension is more "active" and plays a more important role in constructing a vowel system. Ladefoged & Maddieson (1996) also find that all languages have some variations in vowel quality that indicate contrasts in the vowel height dimension, rather than in the front-back dimension and that the languages of the world make a much more limited use of the front-back and rounded-unrounded dimensions. The roundness parameter plays the least active role among the three parameters in constructing a vowel inventory, because great majority of the world's languages have a predictable relationship between the phonetic Backness and Rounding dimensions (Ladefoged & Maddieson 1996). Front vowels are usually unrounded and back vowels are usually rounded, so that the unroundedness of front vowels and roundedness of back vowels can be regarded as predictable. This is also true for the SX surface vowel system as illustrated in (1). Therefore, IDENT-HIGH is more highly ranked than IDENT-BACK and IDENT-ROUND is the least important of the three in formalizing the relations between allophonic vowels and underlying vowels. As a result, the constraint ranking is IDENT-HIGH ≫ IDENT-BACK ≫ IDENT-ROUND.

3.3.3.3 Tense vs ATR

The three constraints above concern the three major features based on the three parameters of constructing vowels in terms of height, position and rounding. Since SX has made a better use of the height dimension in constructing its vowels (which, however, fits into the tendency of the world's languages) the features of [±high] and [±low] are inappropriate to distinguish the four height levels of the 14 surface vowels in SX, which are classified as in (22):

(22)	[i], [ɿ], [y], [ɪ], [u]	[4 high]
	[e], [ə], [ɵ], [ɤ], [o]	[3 high]
	[ɛ]	[2 high]
	[a], [ɑ], [ɒ]	[1 high]

The binary feature framework with [±high] and [±low] can only distinguish three height levels, viz. [+high, −low], [−high, −low], and [−high, +low], excluding the possibility of *[+high, +low]. However, [e] and [ɛ] are a pair of [−high, −low] vowels in SX, as shown in (22), which is common in many other languages, e.g. [o] and [ɔ] in English (Jones 1975). The feature [tense] is usually used to distinguish between [i] and [ɪ], [u] and [ʊ], [e] ands [ɛ], and [o] and [ɔ]. The feature [tense] plays a role in the phonology of RP. For example, the [−tense] vowels cannot occur in final position in a stressed syllable while the [+tense] vowels can (cf. /biː/ *bee* and */bɪ/). The same is true in SX, as shown in (6). The phonological role of [tense] will be discussed later in this chapter.

Halle & Clements (1983) observe that ATR and tense do not seem to contrast in any language.[7] This leads one to assume that ATR and tense might be different names for a single dimension of contrast (see also Yip 1996). There has been some discussion about ATR and tense cross-linguistically (e.g. Stewart 1967; Lindau 1979; among others). Ladefoged & Maddieson (1996) propose that [ATR] should be reserved for the cases wherein tongue root position alone is distinctive. They assume that the distinction in Romance and other languages traditionally referred to in terms of [tense/lax] should not be expressed in terms of [ATR] because the tongue root gesture is not separable from the raising of the tongue body.

Articulatorily speaking, tense vowels are produced with a tongue body or tongue root configuration involving a greater degree of constriction than that found in their lax counterparts; this greater degree of constriction is frequently accompanied by greater length (tense vowels vs. lax vowels) whilst ATR vowels are produced by drawing the root of the tongue forward, expanding the resonating cavity of the pharynx and probably raising the tongue body. There can be some difference between ATR and tense in the articulation of some vowels, especially back vowels. Ladefoged & Maddieson (1996: 304) note: "The high back retracted tongue root vowel is always further back than its counterpart, rather than

[7] For a detailed explanation of [ATR], see Ewen & van der Hulst (2001: 14–21) and Ladefoged & Maddieson (1996: 300–305).

further forward, as in the case for the traditional lax back vowels. Lax vowels of all kinds are normally taken to be more centralized. Retracted tongue root vowels do not always have this characteristic.” Lindau (1979) also points out that there are differences between ATR and tense/lax characterizations of vowels in the acoustic domain.

However, the differences between ATR and tense are minor and differ among languages. Ewen & van der Hulst (2001) present their discussion of ATR and tense and find that the schwa [ə] is [−tense] in a language with [±tense] division, like English, and it is also [+ATR] in a language with ATR vowel harmony, like Akan. In SX, the apical vowel [ʅ] is produced with the tip of the tongue raising, touching the anterior portion of the palate and the body of the tongue being pulled back passed the hard plate to position of the posterodorsum, instead of advancing the tongue root. Thus, I assume that the apical vowel [ʅ] is [−ATR], although it is phonologically [+tense]. However, it is not necessary that both ATR and tense should be used for feature specifications in one language. With the additional minor feature [ATR] to the other four major features, the 13 surface vowels (except for the [+apical] vowel [ʅ]) in SX can be distinguished by the following feature specifications, as shown in (23):

(23)

	y	ɪ	i	u	a	e	ɤ	o	ɛ	ə[8]	ɵ	ɑ	ɒ
[high]	+	+	+	+	−	−	−	−	−	−	−	−	−
[low]	−	−	−	−	+	−	−	−	−	−	−	+	+
[back]	−	−	−	+		−	+	+	−	+	−	+	+
[round]	+	−	−	+	−	−	−	+	−	−	+	−	+
[ATR]	+	−	+	+	−	+	+	+	−	+	+	−	−

The feature specifications in (23) show that in distinguishing the 13 surface vowels in SX there are five features involved, among which four are major features ([high], [low], [back] and [round]), based on the three major parameters of vowels (height, position and roundedness), enough to distinguish the six underlying phonemic vowels (shaded part); [ATR] is a minor feature (so termed by Ladefoged & Maddieson 1996) applied to distinguishing the surface allophonic vowels in SX. The feature specifications in (23) also show that the vowels such as [a], [ə] and [ɵ] differ in the features of [tense] and [ATR].

[8] The central mid schwa [ə] is specified as [+back] in most languages.

According to the phonetic and phonological properties, [a] is unspecified for [back] (cf. Ewen & van der Hulst 2001). There is cross-linguistic evidence that specification of [back] for [a] differs from language to language (Ladefoged & Maddieson 1996, among others). For example, [a] is a front vowel in the Cardinal Vowel system (Jones 1975); [a] is [+back] in *SPE* (Chomsky & Halle 1968: 332); [a] is a low central vowel in a Bavarian dialect[9] (Traunmüller 1982, cited from Ladefoged & Maddieson 1996); [a] in SX is also unspecified for [back].

In (23), [e] and [ə] are distinguished by [back], and [e] and [ɛ] are distinguished by [ATR]. I do not mark the [nas] feature for nasalized vowels, because it is not distinctive for any possible underlying vowel but derives from an underlying nasal coda. We do not yet know what are the underlying vowels for the surface nasalized vowels. It is not necessarily the case that a nasalized vowel is an allophone of its oral counterpart if there is one. Maddieson (1984a) points out that vowels with nasalization sometimes have different qualities from their closest oral counterpart. There is an argument (Wright 1980) that the introduction of a nasal formant at low frequencies – around 200 Hz prompts the speaker to raise the first formant, i.e. produce a more open vowel quality, so as to make perceptual room for the nasal formant. There is certainly morphophonological evidence for such a mechanism in French, cf. synchronic *fine* [i] ~ *fin* [ɛ̃], *une* [y] ~ [œ̃], cf. Lat. *lento* ~ Fr. *lente* [ɑ̃]. However, an oral counterpart of the nasalized vowel is always an ideal candidate for its underlying form for the faithfulness. I will present my analysis of the phonological motivation for the vowel nasalization in SX and I assume vowel nasalization in SX may involve fronting rather than lowering.

For the addition of the minor feature [ATR], I propose one more IDENT-I/O(F) constraint as follows:

(24) IDENT-ATR
Input-output identity for the feature [ATR].

IDENT-ATR is a feature constraint that plays a lesser important role in the phonological system, e.g. with respect to syllable structure, as discussed in the previous section, so that it should be ranked lower than the constraints for major features. Thus, we propose a constraint hierarchy like the following: IDENT-HIGH ≫ IDENT-BACK ≫ IDENT-ROUND ≫ IDENT-

[9] A dialect of Old High German spoken in medieval Bavaria.

ATR, by which the three nasalized vowels are derived from some underlying vowels.

3.3.3.4 OT analysis

Bearing in mind the feature specifications in (23), we can work out the most suitable underlying phonemes of the nasalized vowels with the constraint ranking discussed above by means of an OT analysis. As I assume that vowel nasalization involves fronting, the underlying vowel for [ɵ̃] is proposed to be /o/. Consider the tableau in (25):

(25)

Input /o/	IDENT-HIGH	IDENT-BACK	IDENT-ROUND
a. ☞ [ɵ̃]		*	
b. [ẽ]		*	*!
c. [ɛ̃]		*	*!

According to the feature specifications in (23), the tableau in (25) shows that candidates (b) and (c) violate IDENT-ROUND, so that candidate (a), [ɵ̃], is the optimal output as the surface nasalized vowel of the underlying phonemic /o/. In the tableau above, IDENT-ATR is usually not listed unless it is relevant to the analysis, for it is ranked low. Now let us propose a similar OT analysis on the assumption that the underlying vowel of the nasalized vowel [ɛ̃] is /a/, as shown in (26):

(26)

Input /a/	IDENT-HIGH	IDENT-BACK	IDENT-ROUND	IDENT-ATR
a. [ɵ̃]	*		*!	*
b. [ẽ]	*			*!
c. ☞ [ɛ̃]	*			

Since /a/ is unspecified for [back], IDENT-BACK is irrelevant to the analysis in (26). The tableau in (26) shows that the three candidates (a), (b) and (c) all violate IDENT-HIGH; candidate (a) violates IDENT-ROUND and is first to be ruled out; [ẽ] violates IDENT-ATR and is also ruled out; candidate (c) is the winner for the surface nasalized vowel of the underlying phonemic vowel /a/.

The nasalized [ẽ] has the oral counterpart /e/. Since all the constraints concerning the relations between the underlying representation and surface representation of the nasalized vowels are faithfulness constraints, the oral counterpart /e/ certainly best satisfies FAITHFULNESS and is the optimal candidate. This can be expressed through the following OT analysis:

(27)

Input /e/	IDENT-HIGH	IDENT-BACK	IDENT-ROUND	IDENT-ATR
a. [ɵ̃]			*!	
b. ☞ [ẽ]				
c. [ɛ̃]				*!

The tableau in (27) shows that candidates (a) is ruled out because it violates IDENT-ROUND; [ɛ̃] is also ruled out because it violates IDENT-ATR; candidate (b) is the winner because it has all the identical features its oral counterpart has. Thus, I conclude that in SX the underlying form of the nasalized [ẽ] is its oral counterpart [e]. Through the analysis of the three OT tableaus, we can conclude that the three nasalized vowels, [ẽ], [ɵ̃] and [ɛ̃], are the surface allophones of the underlying phonemic /e/, /o/ and /a/, respectively. This suggests that a fronting process takes place in some way in vowel nasaliztion, as shown in (28):

(28) /e/ → [ẽ]
/o/ → [ɵ̃]
/a/[10] → [ɛ̃]

The illustration in (28) shows that the two phonemic vowels (/o/ and /a/) get fronted when nasalized in surface representation in SX. I assume that such a fronting process in the SX vowel nasalization is place assimilation of the underlying final nasal [n], which I will be discussing in next subsection.

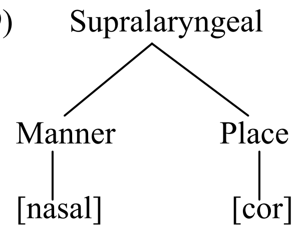
3.3.3.5 Phonological motivation

At this point, some questions may arise: Why should the underlying phonemes /o/ and /a/ become [ɵ̃] and [ɛ̃] in surface representation when nasalized, rather than [õ] and [ã], respectively, just like [ẽ]? What is the (phonological or phonetic) motivation for this change? Why can [ẽ], [ɵ̃] and [ɛ̃] not be phonemic vowels? Cross-linguistically, assimilation or dissimilation is frequently involved in diachronic or synchronic phonological changes. The processes of assimilation or dissimilation are a

[10] /a/ in SX is unspecified for [back]. However, it is obviously fronted when nasalized.

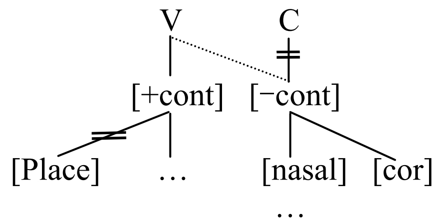
matter of feature spreading, progressive or regressive (Hall 2001). Vowel nasalization in SX is a diachronic process, which came about by debuccalization of final nasals through historical attrition (which will be discussed in detail in chapter 4). It is well known that SX had final [m], [n] and [ŋ] in Middle Chinese times. But in Modern SX only [ŋ] remains and [m] and [n] disappeared, both of which are specified for [−back] (or [+ant]) (see *SPE* 1968: 177). However, in nasal debuccalization, the Place component is lost and what remains is nasality which receives the default Place specification of [cor] (or the element [I] in a Dependency approach). When nasalization occurs, the feature of nasality ([N]) spreads leftward to the preceding vowel, involving [+nasal] and [cor]. The former is a feature of manner and the latter is a feature of place. According to feature geometry (McCarthy 1988), [nasal] and [cor] are in two feature domains under the Supralaryngeal Node, as shown in (29):

(29)



The feature geometry in (29) shows that [nasal] is under the manner domain and [cor] under the place domain and that [nasal] and [cor] are in sister relationship and under the same Supralaryngeal Node, which enables the spreading of both the features of manner and that of place. Van de Weijer (1994, 1996) actually claims that the feature of manner dominates that of place and both features can spread, which is well-attested cross-linguistically (also see Clements 1985; McCarthy 1988). For example, the underlying English prefix ‘in-’ should be ‘ir-’, ‘il-’ or ‘im-’ according to the manner and place features of the following cross-morpheme consonant. The feature spreading in vowel nasalization in SX can be captured by the feature geometry, as shown in (30):

(30)



The feature geometry in (30) shows that when the feature [nasal] spreads to the preceding vowel, the Place feature which is dominated by Manner feature spreads together with it and the original Place feature gets disassociated from the vowel, making the nasalized vowel fronted, which is the phonological and phonetic processes taking place in vowel nasalization in SX. Since [+cont] is the default value of a vowel, it does not get disassociated from the vowel, so that [−cont] has no effect on the vowel. I assume that in SX vowel nasalization, the Place component of the final nasal is lost by debuccalization so that the default Place feature spreads to the preceding vowel while in the English vowel nasalization (e.g. [θæ̃ŋk] ‘thank’) the following nasal is not lost and the Place feature does not spread. In SX vowel nasalization, the spreading of [I] element changes the vowel quality from /a/ and /o/ to [ɛ̃] and [ɵ̃], respectively, when /e/ did not change to a different vowel because /e/ is already a front vowel, both /e/ and [ẽ] having [I] element.

The cross-linguistic evidence strongly suggests that nasal deletion in vowel nasalization may occur diachronically or synchronically for different phonological reasons. In SX vowel nasalization, the syllable-final nasal is debuccalized and the contrastive property of a nasal is now carried by the vowel so that the nasalized vowels are long enough to bear full tones of the lexical syllables, as Halle (1995: 214) explains that since debuccalization does not affect the timing slot of the phoneme, deletion is accompanied by lengthening of the preceding vowel.

The nasalized vowels in SX, [ẽ], [ɵ̃] and [ɛ̃], need not be phonemic vowels underlyingly, though there are some languages in which nasalized vowels are in full contrastive distribution with their oral counterparts and thus are phonemic vowels such as Dan, Zande, Sara and Burmese (Maddieson 1984a). In the world’s languages, the most frequent nasalized vowels are [ĩ ã ũ] also the counterparts of the most frequent oral vowels [i a u] (Ladefoged & Maddieson 1996). In SX, the nasalized [ẽ] has an oral counterpart /e/ which is the underlying form of the nasalized vowel because the surface [ẽ] is derived from the underlying /eN/.

The assumption that the underlying form of the nasalized vowels in SX is /VN/ is also well supported by the fact that there are no ṼC combinations such as *[ẽʔ], *[ɛ̃ʔ] or *[ɵ̃ŋ] in that Ṽ has a final nasal in the coda underlyingly. In short, /e/, /a/ and /o/ are the underlying phonemic vowels of the nasalized [ẽ], [ɛ̃] and [ɵ̃] in the SX surface representation, respectively.

3.3.3.6 Schwa in Shaoxing

Of the six mid vowels in the SX surface representation, one is schwa [ə], which is a common vowel in many other languages (Maddieson 1984a) and the most commonly used vowel in English (Wikipedia 2001).[11] However, [ə] in SX is not as frequent as other vowels. Schwa is usually specified as having many minus specifications in its feature matrix,[12] as shown below:

(31)

	ə
high	−
low	−
front	−
back	+
rounded	−

Perhaps, due to its remarkably negative feature specification, [ə] can be easily assimilated in certain phonetic or phonological environment. In SX, [ə] only occurs in VC structure, either [əŋ] or [əʔ]. It can never constitute a rhyme when standing on its own after the onset so that it can never contrast with the proposed six phonemic vowels. Thus, [ə] is not a phonemic vowel in SX. It is not easy to decide of which phonemic vowel [ə] is an allophone. Consider the distribution of all the VC syllables in SX, as shown in (32):

[11] Wikipedia (www.wikipedia.org) is a multilingual encyclopedia designed to be read and edited by anyone.
[12] Schwa [ə] is phonetically a central mid vowel, although it is phonologically specified as [+back], as shown in (23) and (31).

(32)

Vŋ		Vʔ	
[dɪŋ^{31}]	'stop'	[tɪʔ^{5}]	'fall down'
[dəŋ^{13}]	'wait'	[təʔ^{5}]	'get'
[daŋ^{22}]	'stroll'	[taʔ^{5}]	'build up'
[doŋ^{22}]	'cave'	[toʔ^{5}]	'inspect'
[dɒŋ^{31}]	'sugar'	[tɛʔ^{5}]	'correct'

From the distribution shown in the data in (32), I assume that the most likely underlying vowel(s) of [ə] would be /e/ or /ɤ/, because they do not occur in VC combinations. However, let us first make a reverse OT analysis (from the surface vowel to identify its possible underlying form), to see which candidate is the optimal output as its underlying phoneme, according to the constraint ranking discussed earlier, as shown in (33):

(33)

Input [ə]	IDENT-HIGH	IDENT-BACK	IDENT-ROUND
a. /i/	*!	*	
b. /u/	*!		*
c. /e/		*!	
d. ☞ /ɤ/			
e. /o/			*!
f. /a/	*!		

The tableau in (33) shows that candidates (a), (b) and (f) all violate the first constraint and are ruled out; the candidate /e/ is [−back], so that it is ruled out for violating IDENT-BACK; the candidate /ɤ/ is the optimal output as the underlying phoneme of schwa [ə] in SX. Now let us make a similar OT analysis of the five surface vowels in the 'Vʔ' column in (32) with /ɤ/ as the input, to see if the result is the same, as shown in (34):

(34)

Input /ɤ/	IDENT-HIGH	IDENT-BACK	IDENT-ROUND
a. [ɪ]	*!	*	
b. ☞ [ə]			
c. [a]	*!		
d. [o]			*!
e. [ɒ]	*!		*

The tableau in (34) presents the same result as that in (33) so that it is self-evident that the underlying vowel in the surface [əʔ] combination is /ɤ/. This can also be heuristically supported by data from other Wu dialects. Here are some examples from Qingyuan[13] (Cao 2001) as shown in (35):

(35)	Qingyuan	SX	
	[kʰuɤʔ⁵]	[kʰəʔ⁵]	'thirsty'
	[ʔdɤʔ⁵]	[təʔ⁵]	'obtain'
	[tɤʔ³⁴]	[dəʔ³]	'special'
	[kʰɤʔ⁵]	[kʰəʔ⁵]	'carve'
	[sɤʔ⁵]	[səʔ⁵]	'block'

The OT analysis in (33) and (34) can also be formulated in a rule as follows:

(36) /ɤ/ → [ə] / __ ʔ

There is another strong piece of evidence that /ɤ/ and schwa [ə] have close relations and show similar phonetic and phonological behaviour in SX. For example, schwa is used as an insertion vowel in many languages while in SX /ɤ/ is always used as an insertion vowel, e.g. in loanwords. As was discussed in chapter 2, there is no onset complex in SX so that any CC or CCC cluster in a source language always has /ɤ/ inserted between the consonant cluster when borrowed into SX. For example:

(37)	English	Loanwords in SX	
	a. [kloun]	[kʰɤloŋ]	'clone'
	b. [gri:n]	[kɤlɪŋ]	'Green (name)'

As for the other vowels in 'Vʔ' column, as shown in (32), I assume that the surface vowel of the underlying /e/ is [ɛ] because they share most similarities in features and are always regarded as a pair of mid vowels which only differ in ATR cross-linguistically. I present an OT analysis of the relations between the five surface vowels in 'Vʔ' column in (32) and the underlying phonemic /e/, as shown in (38):

[13] Qingyuan is also one of the Wu dialects which has eight tones, including [334], [52], [33], [221], [11], [31], [5] and [34] (Cao 2002).

(38)

Input /e/	IDENT-HIGH	IDENT-BACK	IDENT-ROUND
a. [ɪ]	*!		
b. [ə]		*!	
c. [a]	*!		
d. [o]		*!	*
e. ☞ [ɛ]			

The tableau in (38) shows that the candidates [ɪ] and [a] violate the constraint for height and are therefore ruled out; candidate [ə] and [o] violate IDENT-BACK because both candidates are [+back]; candidate [ɛ] does not violate any of the three major-feature constraints. Thus, it is the optimal output as the surface allophone of the phonemic /e/. Of the five surface vowels in 'Vʔ' combinations, two have the identical forms with the phonemic vowels. The derivation of all the five surface combinations can be formalized in the following rules:

(39)
$$\left.\begin{array}{lll} /\text{i}/ & \rightarrow & [\text{ɪ}] \\ /\text{a}/ & \rightarrow & [\text{a}] \\ /\text{o}/ & \rightarrow & [\text{o}] \\ /\text{e}/ & \rightarrow & [\text{ɛ}] \\ /\text{ɤ}/ & \rightarrow & [\text{ə}] \end{array}\right\} \_\_\,\text{ʔ}$$

The rules in (39) show that [ɪ], [a], [o], [ɛ] and [ə] are the surface variants of underlying /i/, /a/, /o/, /e/ and /ɤ/, respectively, in 'Vʔ' combinations. The examples in (32) also show that vowels between 'Vŋ' column and 'Vʔ' column are all the same except [ɒ] in [ɒŋ] and [ɛ] and [ɛʔ]. The possible underlying vowel for the surface [ɒ] in [ɒŋ] combination will be discussed in the next subsection of low vowels. Of the five rules in (39), three underlying [+tense] vowels become [−tense] when followed by a consonant. This can also be formulated as follows:

(40)
$$\left.\begin{array}{l} [-\text{low}] \\ \textit{or} \\ [-\text{round}] \end{array}\right\} \rightarrow \ [-\text{tense}] \ \ / \ \ \_\_\ \text{C}\ \$$$

Rule (40) says that any [−low] or [−round] vowel will become [−tense] when followed by a syllable-final consonant. The rule in (40) is supported by the data of SX and also coincides with the syllable structure and tonal structure in SX because, phonetically, [+tense] vowels are articulated longer than [−tense] vowels and the rhymes of V and VC, as the weight unit, are phonetically equal in length and phonologically bimoraic when stressed. The examples in (32) also show that there are such vowels as [ɪ], [ɛ], [ə], [ɒ], [a] and [o] in surface VC combinations. Among them, [ɪ], [ɛ], [ə] and [ɒ] are [−tense]. However, articulatorily and acoustically speaking, the [+tense] rounded vowel [o] and low vowel [a] also sound much shorter when followed by a syllable-final consonant than when in an open syllable. Such an acoustic difference can be clearly manifested in the following syllables when spoken:

(41)	[toŋ^{52}]	'east'	[to^{52}]	'many'
	[dzoŋ^{31}]	'worm'	[dzo^{31}]	'tea'
	[ŋaŋ^{22}]	'hard'	[ŋa^{22}]	'stay'
	[saŋ^{35}]	'save'	[sa^{35}]	'sprinkle'

From all the phonetic and phonological evidence of SX discussed above, especially with regard to the surface realization of vowels in terms of syllable structure, I propose a mora-deletion rule in (42):

(42) $$\underset{\mu\mu}{\overset{\mathrm{V}}{|}} \;\rightarrow\; \underset{\mu}{\overset{\mathrm{V}}{|}} \;/\; \underline{\quad}\ \mathrm{C}\ \$$$

The rule in (42) says that a bimoraic vowel becomes a monomoraic vowel when it is followed by a syllable-final consonant, so that the mora-deletion rule is only realized in syllable structure rather than in feature because [long] is not a distinctive feature in SX. Rule (42) is made possible because the syllable-final consonant is also moraic in SX, which will be discussed in chapter 4. However, the rule in (42) captures all the phonetic facts of vowels in SX.

In short, through the analysis above, I conclude that among the six surface mid vowels, only /e/, /ɤ/ and /o/ exist underlyingly in SX, and [ɛ], [ə] and [ɵ] are allophonic vowels of these three phonemic vowels, respectively (except when [ɛ̃] is in nasalized form, its underlying vowel is /a/) in surface representation.

3.3.4 Low vowels

There are three low vowels in SX, viz. [a], [ɑ] and [ɒ]. Among the three surface low vowels, [a] can stand alone as the rhyme after an onset consonant and contrast with other phonemic vowels, so it is undoubtedly a phonemic vowel in SX. However, [ɑ] and [ɒ] remain questionable because there is disagreement about the existence of [ɑ] in the SX surface vowel inventory, as was discussed in chapter 2. According to the two versions of the Final inventory (see “Yang & Yang’s Finals in SX” (40) and “Zhang’s Finals in SX” (41) in chapter 2), the following syllables can be transcribed differently, as shown in (43):

(43)	Yang	Zhang	
	a. [pɒ^{52}]	d. [pɑɒ^{52}]	‘wrap’
	b. [dʑjɒ^{31}]	e. [dʑjɑɒ^{31}]	‘bridge’
	c. [k^{h}ɒ^{52}]	f. [k^{h}ɑɒ^{52}]	‘knock’

The reason I claim that [ɒ] only occurs in the combination [ɑɒ] in open syllables in the SX surface representation, as shown in (43d, e, f) above and as discussed in chapter 2, is that the tonal structure of SX requires the syllable rhyme to be long enough for the purpose of realizing full tones when stressed. Thus, phonetically, every nuclear vowel in an open syllable must be [+tense] so that it is phonetically heavy enough to be bimoraic. This is supported by the fact that all vowels are [+tense] in open syllables and all [−tense] vowels are followed by a consonant, as discussed previously.

As was mentioned above, [ɒ] is specified as [−tense]. [tense] is a very important feature in determining the vowel system of SX. There is cross-linguistic evidence that [−tense] vowels have different phonological behaviour. For example, in English, [−tense] vowels cannot occur in syllable-final position, while [+tense] vowels can (cf. /biː/ ‘bee’ vs. */bɪ/, for example) (Ewen & van der Hulst 2001). I observe that, like English, a simple [−tense] vowel cannot be the syllable final in SX, which will crucially exclude [ɒ] as a phonemic vowel. However, there are more than enough reasons to make this assumption. As was discussed in chapter 2 and previously in this chapter, the length of a vowel is underspecified so that there is no contrast between [+long] and [−long] vowels. As a monosyllabic language, almost every syllable is stressed in SX, except some syllables which are grammatical particles or affixes, and only a stressed syllable is a full-tone TBU. Articulatorily speaking, a [−tense] vowel is pronounced shorter than a [+tense] vowel. Thus to make the rhyme of a syllable phonetically long enough to be a full-tone TBU, the SX syllable structure phonologically requires its syllable final (all that is left after the onset consonant) to be either a [+tense] vowel or a combination of VC or VV. Let us see how vowels of [+tense] and [−tense] are distributed in the syllable structure. Consider the following data in SX:

(44)	[+tense]		[−tense]	
	[mi^{13}]	‘rice’	[nəŋ^{31}]	‘able’
	[tsʅ^{35}]	‘paper’	[zɪʔ^{3}]	‘enter’
	[so^{52}]	‘sand’	[pɛʔ^{5}]	‘eight’
	[ku^{35}]	‘ancient’	[ɕɪŋ^{52}]	‘new’
	[he^{35}]	‘sea’	[pɵ̃^{33}]	‘half’
	[fɤ^{35}]	‘deny’	[zɒŋ^{31}]	‘taste’
	[ŋa^{13}]	‘we/us’	[ȵiɛ̃^{22}]	‘check’

The data in (44) show that simple vowels are all [+tense] vowels when they make up the whole syllable final. When a [−tense] vowel is in the rhyme, it is never alone, but either it is followed by a coda consonant, or it is nasalized (recall that nasalized vowels are phonetically longer than their oral counterparts (Rosner 1994)); phonologically, nasalized vowels are underlying vowel + nasal sequences in SX, as was discussed in chapter 2 and previously in this chapter. The data in (44) strongly suggest that a simple [−tense] vowel cannot be the final of a syllable. Thus, I assume that there is a segment filter in SX syllable structure to make sure every syllable is properly structured in terms of segments, as shown in (45):

(45) $$C \; {}^{*}\begin{bmatrix} \text{V} \\ -\text{tense} \end{bmatrix} \; \$$$

The segment filter in (45) stipulates that a simple [−tense] vowel is not acceptable in an open syllable in SX. This segment filter will naturally filter out [ɒ] as the whole syllable final, so that the syllables in (43a, b, c) are ill-formed. Only syllables like (43d), (43e) and (43f) are acceptable, as was also presented in the Final inventory of Chao’s ((38), ch.2) and Campbell’s ((39), ch.2). The phonetic and phonological motivation why [ɒ] never occurs alone but in combinations of [ɑɒ] or [ɒŋ] in the SX surface representation is to satisfy the segment filter so as to be heavy enough for bimoraic status. Acoustically, the rhymes of the syllables in (43d, e, f) are factually as long as [ɑɒ], not as short as [ɒ]. However, both [ɑ] and [ɒ] are allophonic vowels in surface representation. The underlying phonemes of these two allophonic vowels can also be worked out through an OT analysis, with the same constraint ranking as shown in (46):

(46)

Input [ɑ]	IDENT-HIGH	IDENT-BACK	IDENT-ROUND
a. /i/	*!*	*	
b. /u/	*!*		*
c. /e/	*!	*	
d. /ɤ/	*!		
e. /o/	*!		*
f. ☞ /a/			

The tableau in (46) shows that candidates (a), (b), (c), (d) and (e) all violate IDENT-HIGH once or twice[14] and are ruled out together while candidate (f) does not violate any of the three constraints and is surely the winner. Thus /a/ is the optimal underlying phoneme for the allophonic vowel [ɑ], which is also phonetically satisfying because of the articulatory similarities between [a] and [ɑ]. Then we come to the analysis of [ɒ] in the same approach. But since [ɒ] only occurs in the combination [ɑɒ], we should also apply the OCP to the analysis to eliminate any possible sequence of the two exact same segments. This is not acceptable underlyingly in SX phonology. The OCP is inviolable in SX, so that it dominates the other faithfulness constraints. In the following analysis I propose [ɑɒ] as the input and /a/ as the first V of the combination for the output because /a/ is already decided as an underlying vowel for the allophonic [ɑ] through the analysis in (46). The violation of the output candidates only refers to those by the second V of the combination, as shown in (47):

[14] As is shown in (23), [ɑ] is specifies as [−high, +low] and [i] and [u] are specified as [+high, −low], so that candidates (a) and (b) violate IDENT-HIGH twice for the specifications of [high] and [low].

(47)

Input [ɑɒ]	OCP	IDENT-HIGH	IDENT-BACK	IDENT-ROUND
a. /ai/		**!	*	*
b. /au/		**!		
c. /ae/		*	*!	*
d. /aɤ/		*		*!
e. ☞ /ao/		*		
f. /aa/	*!		*	*

The tableau in (47) shows that candidate (f) violates the OCP and is the worst candidate, so it is ruled out; candidates (a) and (b) violate IDENT-HIGH one once more than (c), (d) and (e) so that they are also ruled out; candidate (c) is also ruled out by violating IDENT-BACK; candidate (d) finally is ruled out because it violates IDENT-ROUND; candidate (e) is the winner. Thus, /ao/ is the optimal underlying form of the surface combination [ɑɒ]. This result is also satisfying for the acoustic similarity between the input and output. However, according to my analysis in chapter 2, /ao/ is the only diphthong in SX underlyingly, which is not acceptable in surface representation because of the surface constraint *DIPH,[15] having the surface form [ɑɒ] consequently. The reason why there is no diphthong in the SX surface representation is not clear so far. However, it was widely accepted that diachronically the nucleus of the SX syllables have become shorter than that in Middle Chinese times, having lost the Middle Chinese diphthongs such as [əu], [ou], [ai] and [au], some of which are still retained in some other Wu dialects (Chao 1928; Cao 2002). One hypothesis would be that during the shortening of the nucleus, all diphthongs were missing and became monophthongs, but [ɒ] was ruled out by the segment filter in (45). Thus, SX has [ɑɒ] not only to satisfy the segment filter but also follow the tendency of losing diphthongs. The change from /ao/ into [ɑɒ] can be formalized as follows:

(48) $$/\mathrm{ao}/ \rightarrow \begin{bmatrix}+\mathrm{back}\\ +\mathrm{low}\end{bmatrix} \;\Big/\; [+\underline{\mathrm{low}}]\,[+\underline{\mathrm{back}}]$$

The rule in (48) shows that when VV is a [+low]+[+back] combination, both VV will become [+low, +back] in surface representation. As

[15] I claim that there is no diphthong (*DIPH) in the SX surface representation and that [ɑɒ] is not a diphthong, as was discussed in chapter 2.

presented in (23), both [ɑ] and [ɒ] are specified as [+low] and [+back]. It is an interesting linguistic phenomenon in SX that in GV there is constraint OCP(H) (*[+high][+high]) so that /ju/ merges into [y] while in VV [+low][+low] is preferred so that /ao/ changes into [ɑɒ], which suggests that *DIPH dominates OCP(H) in SX. In fact, the two phonological changes both involve merger. In /ju/, the two segments merge into a [−back, +round] segment [y] and in /ao/ the two segments merge into a [+low, +back] combination rather than a single segment because the feature [+round] does not merge. Otherwise, the merged segment would be [ɒ] which is ruled out by the segment filter in (45). The only difference is that the high vowels merge into a front vowel and low vowels merge into (a) back vowel(s), which just fits in with the general vowel inventory that high vowels are more likely to be in front and low vowels are more likely to be in back (Maddieson 1984a).

In the data of SX, there are also such syllables as [zɒŋ^{31}] 'taste' and [fɒŋ^{52}] 'square', in which [ɒ] does not occur as a combination of [ɑɒ]. This phenomenon gives rise to the question if the underlying form of [ɒŋ] is /oŋ/. The problem is that in fact there are well-formed syllables such as [kɒŋ^{52}] 'steel' and [koŋ^{52}] 'male' in the SX surface representation. The reason that [ɒ] only occurs in combinations of [ɑɒ] and [ɒŋ] is that both [ɑɒ] and [ɒŋ] satisfy the segment filter in (45). If /oŋ/ is also the underlying form of [ɒŋ], how we could have both [ɒŋ] and [oŋ] as surface representation for the same underlying /oŋ/? I assume that [ɒŋ] is a derived form from [ɑɒŋ] so that the underlying form of [ɒŋ] is /aoŋ/, rather than /oŋ/, since /joŋ/, /woŋ/, /waŋ/, and /jaŋ/ are all well-formed underlyingly in SX. As was discussed above, [ɒ] cannot occur alone as the syllable final for its [−tense] or phonetically short duration. A [ɑɒ] combination is just like a long vowel in terms of time duration, so that [ɑɒ] is long enough to satisfy the segment filter but too long to be followed by the final nasal [ŋ]. As a result, [ɑ] is dropped when the syllable ends in [ŋ] in surface representation because the syllable-final nasal is also moraic, while such syllables as [dʑjoŋ^{31}] 'poor' and [ɦwoŋ^{31}] 'red' are also well-formed because the prenuclear glides are weightless. Accordingly, the underlying form of [ɒŋ] is /aoŋ/, not /oŋ/. In short, among the three low vowels, only /a/ is a phonemic vowel; [ɑ] and [ɒ] are allophonic vowels of /a/ and /o/, respectively.

3.3.5 The distribution of glides

3.3.5.1 What is a glide?

Perhaps the most problematic segment type for all theories of phonology is the class of glides (Hyman 2003:77). There have been controversial definitions of what a glide is. According to Trask (1996), a glide is a very brief phonetic vowel which functions in some languages as a phonological consonant; the English glides /j/ and /w/ (as in *yes* and *win*) are brief versions of [i] and [u]. Conventionally, glides are also known as semivowels. Jakobson, Fant and Halle (1963) think that there is only an allophonic difference between semivowels and vowels, which, however, has been challenged in later studies (Rosenthall 1997). Ladefoged and Maddieson (1996) call glides vowel-like consonants. Phonetically speaking, glides are sounds produced with a relatively unimpeded flow of air through the mouth. The constriction is not narrow enough to produce local turbulence, though cavity friction may be heard (Maddieson 1984a). In the *SPE* feature system, glides are [−cons, −voc] segments, which is not really insightful in that in some languages [−cons, −voc] segments also include [r], [h], [ʔ], etc. (Trask 1996). There are two kinds of glides, on-glides and off-glides. The former is a glide occurring at the beginning of a diphthong, such as [j] in [ja] and the latter is one occurring at the end of a diphthong such as [j] in [aj]. In Mandarin, all on-glides are [−voc] but off-glides can be [+voc] like [i] in [xwai^{35}] ‘chest’ and [u] in [$\text{t}^{\text{h}}\text{jau}^{55}$] ‘choose’, in which [ai] and [au] are treated as falling diphthongs. I would say, a glide is [−cons] and [−peak] in a syllabic aspect. However, a real glide should be both phonetically and phonologically a glide. In a CV approach, a glide is a C-dominated V, different from the corresponding high vowels which are V-dominated V. In van de Weijer’s (1994, 1996) element-based segmental structure, the three glides [j], [w] and [ɥ] can be formalized as in (49):

(49) a.		b.		c.	
	C		C		C
	|		|		|
	V		V		V
	|		|		| \\
	I		**U**		**I U**
	[j]		[w]		[ɥ]

The segmental structures in (49) capture what a glide is by nature, free of ambiguity whether they are [+voc] or [−voc], or whether they should be symbolized as [i] or [j], [u] or [w] and [y] or [ɥ].

3.3.5.2 Glides in Shaoxing

The majority of the world's languages (65.2%, Maddieson 1984a) have both /j/ and /w/ as glides, which are closely related to the high vowels /i/ and /u/, respectively, or in complementary distribution with /i/ and /u/, respectively, in many languages (Casali 1996). In SX, there are three glides in surface representation, viz. [j], [w] and [ɥ], which are all on-glides. SX has no off-glide: VG structure is not acceptable in SX. Thus glides in SX are all [−voc]. The three glides can be specified with the following features and thus be distinguished from the identical vowels as in (50):

(50) The glide feature specification[16]:

	Glides			Vowels		
	j	w	ɥ	i	u	y
[voc]	−	−	−	+	+	+
[back]	−	+	−	−	+	−
[round]	−	+	+	−	+	+

The feature specifications in (50) show that the only difference between glides and vowels in SX is [±voc], which is also true with the on-glides in other Chinese dialects. The syllable structure in SX will be discussed in detail in the next chapter. In this subsection I will present my analysis of the distribution of the three glides in SX. Glides in SX occur in two structures, GV and GVC.[17] Consider the following examples:

[16] For convenience, the glides are specified here with the features according to *SPE* (Chomsky & Halle 1968). However, I don't mean that all [−cons] and [−voc] segments are glides, which is of course not true. To distinguish between off-glides and their identical vowels, I would propose the feature [±peak].

[17] Either GV or GVC, excluding the onset C, is a sub-syllabic constituent in SX, which will be discussed in chapter 4.

(51) GV

[ɕja^{35}]	'write'	[kwa^{35}]	'strange'
[ȵjɛ̃^{31}]	'inspect'	[kwɛ̃^{52}]	'close'
[ɦje^{13}]	'too'[18]	[kwe^{33}]	'piece'
[ʔjo^{35}]	'graceful'[19]	[hwo^{52}]	'flower'
[tɕɥɵ̃^{33}]	'donate'	[hwɵ̃^{52}]	'happy'

GVC

[dʑʰjoŋ^{31}]	'poor'	[ɦwoŋ^{31}]	'red'
[ɕjaŋ^{35}]	'think'	[ɦwaŋ^{31}]	'horizontal'
[tɕʰjoʔ^{5}]	'lack'	[kʰwoʔ^{5}]	'wide'
[tɕjaʔ^{5}]	'foot'	[kwɛʔ^{5}]	'scratch'

The examples in (51) show that the glides [j] and [w] in SX seem to be in contrastive distribution with each other in both GV and GVC structures (in spite of the fact that the preceding consonants are in complementary distribution when preceding [j] and [w]. This will be discussed in chapter 4). However, there is one case in which [j] and [w] contrast with each other after the same onset consonants, which involve only alveolar stops and the lateral [l]. For example:

(52)

[ljɤ^{31}]	'flow'	[lwɵ̃^{22}]	'mess up'
[tja^{52}]	'dad'	[twɵ̃^{33}]	'stew'
[tʰjẽ^{52}]	'sky'	[tʰwɵ̃^{33}]	'swallow'
[djɑɒ^{22}]	'exchange'	[dwɵ̃^{31}]	'unite'

The examples in (52) show that [j] and [w] do contrast after the same consonants, though in most cases the onset consonants are in complementary distribution with others when preceding [j] and [w] (see the details of the distribution of consonants and vowels in §4.6.2, ch.4). However, it is true that there are no exact minimal pairs of [j] and [w], differing in either the preceding onset consonant or the following nucleus vowel, as shown in (52). I assume that the different distribution of [j] and [w] is decided by the phonotactics by their different phonological properties, e.g. [±back] and [±round]. Both [j] and [w] are phonologically different glides. The segment [ɥ] is an allophonic glide of [j] when

[18] [je] only occurs in literary style for some syllables borrowed from Mandarin, so it is not the real native SX pronunciation.
[19] [jo] also mostly occurs in literary style for some syllables borrowed from Mandarin.

followed by a [−back] rounded vowel, which can be formalized in a rule as follows:

(53) $/j/ \rightarrow [\text{ɥ}] \;/\; \_\_ \begin{bmatrix} -\text{back} \\ +\text{round} \end{bmatrix}$

The rule in (53) says /j/ becomes rounded when followed by a [−back] rounded segment, which actually involves labial assimilation triggered by the spreading of the [+round] feature of a [−back] vowel, rendering [jo] acceptable because [o] is [+back]. This phenomenon gives rise to the question why a [+back] vowel does not trigger labial assimilation in this case. As was discussed in chapter 2, [–back] roundedness is marked and [+back] roundedness is unmarked. Cross-linguistic evidence shows that marked features can trigger assimilation more strongly than unmarked ones. Underlyingly, there are only two glides in SX, viz [j] and [w].

Glides in SX have very remarkable characteristics in terms of syllabic position and phonological behaviour. Generally speaking, if there is a sequence CGV[20] in a language, usually G is either in the Nucleus so that GV is a diphthong, like [ja] in [ljato] 'boat' in Luganda (Clements 1986), or G is in the Onset so that CG is an onset cluster like [tj] in [etjo] 'to pull' in Okpe[21] which has no diphthong (Casali 1996). However, as was discussed previously, we agree that there is no diphthong or onset cluster in SX. Thus, G in CGV sequence is neither in the Nucleus nor the Onset. This topic will be discussed in detail in chapter 4.

Within one syllable, most VV-like sequences are diphthongs, either rising or falling. But some languages have no diphthongs, like SX and Okpe (Casali 1996), in which V_1 in V_1V_2 sequences is always a glide. Casali (1996: 18, 46) proposes a constraint GLIDEHOOD, saying (i) a glide must be [+high]; (ii) a glide must be [+front] or [+round]. Accordingly, in any VV-like sequences in SX, the first V-like segment must be a glide if it satisfies GLIDEHOOD because there is no diphthong in SX, so there is no VG sequence. The fact that V_1 in V_1V_2 sequences becomes a glide is a cross-linguistic commonality.

What is different between CGV in SX and CGV in many other world's languages is that in other languages when V_1 becomes G it is either in CG sequences as an onset cluster, like in Okpe, or in a GV

[20] C refers consonant; G refers to glide; V refers to vowel.

[21] Okpe is a Benue-Congo language spoken in Nigeria, which is discussed in Hoffman (1973), Pulleyblank (1986), and Omamor (1988) (see Casali 1996).

sequence as a diphthong as in Luganda, while in SX when V_1 becomes G it is neither an onset cluster nor a diphthong, because of the *COMPLEX-ONSET and *DIPH constraints in SX, which brings up a controversial issue of the syllable structure not only in SX but also in all other Chinese languages. The syllabic status of the prenuclear glides is a highly remarkable characteristic in SX. However, glides in SX are predictable because V_1 in V_1V_2 sequences is unexceptionally a glide if V_1 is [+high]. Glides are position-sensitive in SX. There is also cross-linguistic evidence that an underlying distinction between glides and high vowels is not based on a difference in feature content. Rather, it is shown that the difference is structural (Levin 1985).

3.4 The Six-vowel System

From the analyses I have presented in the previous sections, it clearly emerges that the underlying vowel inventory of SX consists of six phonemic vowels, which can be displayed as in (54):

(54)	i		u
	e	ɤ	o
		a	

The underlying vowel inventory of SX in (54) shows that SX has a symmetrical six-vowel system with two high vowels, three mid vowels and one low vowel. This symmetrical six-vowel system fits in with the general pattern of vowel systems of the world's languages. Vowels in the mid range are a little more common than high vowels and low vowels are substantially less common (Maddieson 1984a), while the number of height distinctions in a system is typically equal to or greater than the number of backness distinctions (Crothers 1978). Among the six phonemic vowels, five are the most preferred vowels in the world's vowel system, with /ɤ/ as a remarkable extra phonemic vowel in SX.

Crothers (1978) proposes that 55% of the languages with six-vowel system have /i u e ə o a/. This vowel inventory has only one vowel /ə/, which is different from that of SX. It is true that /ə/ is much more common than /ɤ/ in the world's languages (Maddieson 1984a). However, there is a strong phonological motivation to assume that SX has /ɤ/ instead of /ə/ as one of the six phonemic vowels. In this section I will briefly analyze the phonetic and phonological similarities and differences between /ɤ/ and /ə/ and discuss the phonological mechanism of SX which supports the six-vowel system, as I have presented above. I assume that the present six-vowel system has close relations with the syllable structure of SX, which bears on two main issues: stress and tenseness. These two issues give the answer to why schwa /ə/ is not in the six-vowel system of SX.

3.4.1 Stress

According to Maddieson's (1984a) UPSID analysis, about 21.1% of the languages have schwa /ə/. Phonologically speaking, the schwa is the vowel sound in many unaccented syllables in words of more than one syllable in many languages. It is almost always unstressed. For example, in English this vowel is related to rhythmic factors, which makes a contrast between stressed syllables and unstressed syllables. However, as was discussed in chapter 2, almost every syllable is a lexical word in SX and every syllable can be stressed so as to realize the full tone(s) of the syllable, except when a grammatical particle or an affix is involved. Thus, stress in SX is a realization of full tones which have to be carried by two moras of the rhyme (weight unit) so that the nuclear vowel (if it is all that the rhyme has) in the syllable of SX should be phonetically and phonologically bimoraic (the tonal structure in SX will be discussed in detail in chapter 5). Schwa /ə/ is too lightly pronounced to play such a phonological role in a syllable of SX. Phonetically speaking, /ɤ/ is the most similar to /ə/, which is strongly supported by some SX loanwords as shown in (55):

(55)	English	SX	
	[ˈmaikr**ə**foun]	[maʔ^{22}k**ɤ**^{33}fuŋ^{52}]	'microphone'
	[ˈtʃɔk**ə**lɪt]	[tɕ^{h}jɑɒ^{35}k**ɤ**^{33}lɪʔ^{31}]	'chocolate'

The examples in (55) show that the schwa /ə/ in unstressed syllables in English becomes /ɤ/ in SX loanwords because every syllable in [maʔ^{22}k**ɤ**^{33}fuŋ^{52}] and [tɕ^{h}jɑɒ^{35}k**ɤ**^{33}lɪʔ^{31}] has to be stressed to realize full tones and [ə] and [ɤ] share many phonetic and phonological similarities, as was shown in the OT tableau in (34) and in the examples in (37). However, [ə] only occurs in CV combinations when [ɤ] can occur in an open syllable (or syllable final) in SX. As a result, /ɤ/ is assumed instead of /ə/ in the phonemic vowel system of SX to have phonetic and phonological reality of being stressed in an open syllable.

3.4.2 Tenseness

Specification of [tense] also plays an important role in the six-vowel system of SX and its syllable structure. As was discussed previously, there is a segment filter (45) in the SX syllable structure, which eliminates a [−tense] vowel alone as the syllable final of SX. Schwa [ə] is a [−tense] short vowel so that it is not accepted as a final phonemic vowel in SX because the segment filter requires that all phonemic vowels in open syllables must be [+tense] for the sake of weight. De Boer (2001) proposes that [±tense] is not a phonemic feature but an allophonic feature. This is also true for SX because all [−tense] vowels in SX are in fact allophones of phonemic vowels, and [±tense] is a redundant feature among the phonemic vowels in SX. However, not all languages divide the set of vowels into a tense and a lax subset. Many have two subsets according to tongue-root position, i.e. one [+ATR] set and one [−ATR] set, e.g. /i u e o ə/ are grouped for [+ATR] and /ɪ ʊ ɛ ɔ ɑ/ are grouped for [−ATR]. [ATR] plays an important role in some phonological systems, e.g. vowel harmony in Akan (see Ewen & van der Hulst 2001). Usually, [+ATR] vowels are tensed vowels, so that [ATR] also plays a role in deriving some surface representations from abstract underlying representations, as analysed above. However, in SX, of the 14 surface vowels, only /i u e ɤ o a/ constitute its underlying vowel system because these vowels are in natural class with specification of [tense].

In short, both for the segment filter of the syllable structure and the weight of syllables, specification of [tense] is one of the most important phonetic and phonological factors which make /i u e ɤ o a/ the underlying vowel system of SX. This not only satisfies the tendency of vowel systems of the world's languages to be symmetrical, but also satisfies the phonological demands of SX.

3.5 Summary

In this chapter I have attempted four things. First, I have worked out the constraint ranking for the analysis of the relations between the underlying phonemes and the surface vowels, proposing some constraints and rules which express certain phonological principle in general. Second, I have proposed a segment filter on SX syllable structure, which explains the phonological motivation for the alternation of some underlying phonemes into different variants in surface representation. I claim that a phonemic vowel in SX must be [+tense] to occur in open syllables. Third, I have presented a clear picture of the distribution of all the 14 surface vowels and three surface medial glides of SX, proposing some constraints and rules for either complementary or contrastive distribution. Last but not least, I have worked out an underlying vowel inventory of SX, a symmetrical six-vowel system, including /i u e ɤ o a/, which fits in the general pattern of vowel systems of the world's languages.

4 The Syllable Structure of Shaoxing

4.1 Introduction

The status of the syllable as a linguistic unit, although not uncontroversial, is widely accepted in present-day phonology and phonetics. Syllables may consist of a vowel or diphthong, with onsets and codas of one or more consonants. Syllables may also contain syllabic consonants. Human listeners seem to need syllables as a way of segmenting the speech stream, while speakers use syllables to impose a rhythm of strong and weak beats to language, just like in music. All languages are assumed to have syllables, although the syllable status is sometimes questioned by researchers working on languages with extreme collocations of consonants or vowels, such as Bella Coola (Bagemihl 1991) and Gokana (Hyman 1990, 2003).[1] Levin (1985) claims that phonetic utterances in all natural human languages are made up of syllables. However, the primary evidence for the syllable seems rather phonological than phonetic, in that "the syllable is the phonological unit which organizes segmental melodies in terms of sonority; syllabic segments are equivalent to sonority peaks within these organizational units" (Blevins 1995: 207). The syllable allows the formulation of generalizations both at the segmental level and at higher prosodic levels, which are awkward to express without referring to this constituent. For example, the recurrence of the context {C, #}[2] in phonological rules (Roca 1994) indicates that generalizations are difficult to express without reference to the higher prosodic unit of the syllable. Some examples of evidence for the syllable boundary are given in (1):

(1) Evidence for syllable boundary:
   (a) In German, Dutch, Russian and Turkish (Clements and Keyser 1983: 59; Trommelen 1984; Itô & Mester 2003), obstruents are de-

[1] Hyman (1983, 1990, 2003) argues that the syllable is not universal, based on examples of extreme vowel collocation in Gokana, e.g. [mɛ́ɛ́ ɛ̀ kɔmm̀ kɛ̃ɛ̃̀ɛ̃̀ɛ̃̀ɛ̃ɛ̃́] 'who said I woke him up?'. Even so, this does not necessarily mean that there is no evidence for the syllable in Gokana.

[2] The context {C, #} means that some phonological rules take effect only when the relevant segment occurs before a consonant or at the end of the word. In both cases the target occurs at the end of the syllable. In SX, a monosyllabic language, every word orthographically presented in a Chinese character is a syllable.

voiced when in the final position of the syllable or followed by a syllable-final consonant:

$$[-\text{son}] \rightarrow [-\text{voice}] \ / \ \_\_ \begin{Bmatrix} \text{C} \\ \# \end{Bmatrix}$$

(b) In English (Roca 1994: 134), laterals are velarized (i.e. [+high, +back]) when in the final position of the syllable or followed by a syllable-final consonant:[3]

$$[+\text{lateral}] \rightarrow \begin{bmatrix} +\text{high} \\ +\text{back} \end{bmatrix} \ / \ \_\_ \begin{Bmatrix} \text{C} \\ \# \end{Bmatrix}$$

(c) In English (Ewen & van der Hulst 2001: 18), non-tense vowels cannot occur in final position in a stressed syllable:

$$* \begin{bmatrix} \text{V} \\ +\text{stress} \\ -\text{tense} \end{bmatrix} \#$$

(d) In Turkish (Clements and Keyser 1983: 59), vowels are inserted between the consonants which have different Place features when the cluster is in syllable final position:

$$\emptyset \rightarrow \text{V} \ / \begin{bmatrix} \text{C} \\ \alpha\text{Place} \end{bmatrix} \text{ — } \begin{bmatrix} \text{C} \\ \beta\text{Place} \end{bmatrix} \#$$

If the two consonants share the same Place features, the second consonant is deleted when in syllable final position:

$$\begin{bmatrix} \text{C} \\ \alpha\text{Feature} \end{bmatrix} \rightarrow \emptyset \ / \begin{bmatrix} \text{C} \\ \beta\text{Feature} \end{bmatrix} \text{ — } \#$$

(e) In Shaoxing (discussed in chapter 3), non-tense vowels do not occur at the end of syllable:

$$\text{C} * \begin{bmatrix} \text{V} \\ -\text{tense} \end{bmatrix} \#$$

---

[3] The lateral [l] in English is [−high, −back] elsewhere according to *SPE* (Chomsky & Halle 1968: 176–177).

The examples in (1) show all these phonological changes in different languages refer to the syllable domain; the simple formulation is made by referring to the syllable domain.

There are more examples of the context {C, #} in rules in many other languages. The examples in (1) will suffice to show that the recurrence of these environments in what otherwise appear to be unrelated phenomena cannot merely be attributed to chance. The syllable has played a key role in generative phonology since *SPE* (see for further discussion Cho & King (2003), among others). The syllable is a significant unit in determining how lower-level segmental units are grouped into constituents.

The syllable of every language has a hierarchical structure (Selkirk 1982; Zhang 2000) which may differ from language to language. Although the Sonority Sequencing Principle (SSP) is a universal principle (Hooper 1976; Kiparsky 1979; Selkirk 1984; Harris 1994; among others), (apparent) surface violations against the SSP occur in some languages, such as Georgian (Butskhrikidze 2002), Polish (Rowicka 1999) and Bella Coola (Bagemihl 1991; Cho & King 2003). The Chinese languages, which are usually monosyllabic languages, have a simpler syllable structure in terms of segment sequences than Indo-European languages. However, aspects of the internal constituent syllable structure of Chinese have been problematic for a long time; for instance, with respect to the status of the on-glide. There are a number of different hypotheses with respect to the prenuclear glide: (i) that it is the second member of an onset cluster (Yin 1989; Bao 1990; among others); (ii) that it forms a secondary articulation on the initial consonant (Duanmu 1999, 2000a; Wang 1999); (iii) that it is part of the rhyme constituent (Wang & Chang 2001; among others); (iv) that it is head of yet another constituent, the Final[4], but falls outside of the rhyme (Cheng 1966; Lin 1989, 1990). In this chapter I will discuss previous studies on the internal syllable structure of Chinese, focusing on the status of pre-nuclear glides. My own analysis of the prenuclear glide in SX will also be offered in this chapter (section 4.3.6), and will also be used to cast some light on the syllable structure of Mandarin.

[4] Throughout this chapter, 'Final' (beginning with a capital 'F') refers to all that is left after the initial consonant in the syllable, i.e. to the syllabic constituent recognized in traditional Chinese phonology.

4.2 The Syllable Types of Shaoxing

The most basic syllable (or ‘core’ syllable) is one that consists of a single consonant followed by a single vowel (McCarthy & Prince 1986; Roca 1994; among many others), i.e. CV. Languages whose syllable inventory is limited to CV are reportedly Senufo[5] (Clements & Keyser 1983: 29) and Hua[6] (Blevins 1995: 217). More commonly, however, more syllable types exist in a language, either because the language allows more than one consonant either before or after the vowel, or because it permits onsetless syllables (i.e. without the initial C), or both. Thus, most of the world’s languages also have syllables of the type V, CVC, VC as well as CV. As Clements and Keyser (1983) point out, these augmentations implicitly contain the prediction that the presence of the more complex types in any one language implies the presence of their simpler counterparts. For example, if a language has the VC type, it is most likely also to have CV, V and CVC types.

As was discussed in previous chapters, the maximal syllable type of SX is CGVX (X can be C or V in SX), which implies that SX also has the CV, CVC and CGV syllable types, as shown in (2):

(2)	CV	[ɕi^{33}]	‘opera’
	CVC	[ɕɪŋ^{52}]	‘new’
	CGV	[ɕja^{35}]	‘write’
	CGVC	[ɕjaŋ^{35}]	‘think’
	CGVV	[ɕjɑɒ^{35}]	‘small’

The formalization of syllable types in (2) shows the five possible syllable types with an onset C, though there are some constraints on the sequences between C and V or C and G. The details of all the possible combinations will be provided in the tables in (54), (64), (70) and (73) in the last section of this chapter.

There have been a number of approaches to the syllable initial of Chinese (including SX): they question whether a syllable can begin with a vowel or a glide, or if it has to begin with a zero initial when there is no other consonant in the syllable initial (Chao 1928; Wang 1985; Duanmu 1999, 2000a).

---

[5] Spoken in Northern Côte d'Ivoire and Mali.

[6] A Khoisan language of Botswana that is most closely related to !Xóõ in Africa.

Many authors (e.g. Chao 1968; Wang 1985; Duanmu 1999, 2000a) assume that Chinese (including all dialects) has an obligatory zero initial in the onset when there is no other initial consonant. On this assumption, there is no V or VC syllable type in Chinese. This assumption is mainly based on the fact that there is no liaison between a syllable ending in a consonant and a syllable beginning with a vowel. For example, [mjan^{35}au^{214}] (in Mandarin) 'wadded jacket' cannot be pronounced as *[mjan^{35}nau^{214}] (the question of germination will be discussed later in this subsection) because, they claim that there is a zero initial in front of the vowel, which blocks liaison (Chao 1968; Duanmu 1999). It is true that liaison does not occur in such cases. I claim that there is a phonetic onset [ʔ] in SX. In fact there are two phonetic onset consonants: [ʔ] and [ɦ], standing for the two glottis features [+stiff] and [+slack], respectively, which are required by the consonant-tone correlation, as was mentioned in chapter 2 and will be discussed in detail in chapter 5. It is well documented that voiceless ([+stiff]) initial obstruents induce high tones and voiced ([+slack]) initial obstruents induce low tones cross-linguistically. In SX, when the syllable has no other phonological onset consonant, it has the voiceless glottal stop [ʔ] for high-register syllables and the voiced glottal fricative [ɦ] for low-register syllables. In broad transcription, the initial glottal stop [ʔ] does not appear in SX because it is not a phonological onset (Wang,[7] personal communication 2004) while the initial glottal fricative [ɦ] is never missing because, I assume, first, [ɦ] is also a phonological onset, contrastive with [h]; secondly, [ɦ] is acoustically robust in articulation even as a phonetic onset. However, liaison does not occur between two full-tone syllables in SX even when the second syllable has no phonological onset consonant, e.g. [tsəŋ^{52}aʔ^{5}] 'steam (a) duck', which can never be pronounced as *[tsəŋ^{52}ŋaʔ^{5}]. Chinese is a monosyllabic language in which every syllable is a lexical unit, and the alignment between phonological units and morphological units is never violated. In both SX and Mandarin such liaison can never occur; otherwise the lexical meaning of the compound or phrase would be totally changed. Consider the following examples:

[7] I am grateful to Prof. Wang, Futang, an expert in dialectology in Beijing University, China, for his suggestion.

(3) SX

a. [ʑjaŋ^{33}(ʔ)o^{35}] ‘like mute’
b. [ʑjaŋ^{33} # ŋo^{13}] ‘like me’
c. [ʑja^{33} # ŋo^{13}] ‘thank me’

Mandarin

d. [tjan # **an** # məŋ] ‘Heaven **Peace** Gate’
e. [tjan # **nan** # məŋ] ‘Heaven **South** Gate’

The examples in (3) show that the syllable boundary (which is also a word boundary) plays an important semantic role and any liaison between syllables will cause semantic changes. This phenomenon is stipulated by a universal alignment constraint, after McCarthy & Prince (1993), as stated in (4):

(4) ALIGN-L
The left edge of a lexical unit coincides with the left edge of a full-tone syllable (“no word-initial onset insertion”).

ALIGN-L in (4) stipulates that no onset consonant can be inserted before an underlying vowel-initial syllable which is a lexical unit and has a full tone. Thus, liaison does not occur between two full-tone syllables. This holds for SX as well as Mandarin. But liaison does occur between a syllable ending in a consonant and a syllable beginning with a vowel when the second syllable is toneless. However, ALIGN-L does not take effect on a toneless syllable. Thus, when the vowel-initial syllable has no tone, a phonetic onset is not necessary, which causes liaison to occur. For example:

(5) [ts^{h}oŋ^{52}a] → [ts^{h}oŋ^{52}ŋa] ‘dash ahead’
[ɲɪŋ^{31}a] → [ɲɪŋ^{31}ŋa] ‘(how many) people!’

The examples in SX in (5) both form a disyllabic unit, in which the second (toneless) syllable does not have real lexical meaning and only serves to emphasize the action expressed by the first syllable. Thus, the toneless vowel-initial syllable does not require a phonetic onset for the consonant-tone correlation, allowing liaison to take place. Yip (1980: 191) refers to such cases of liaison as ‘resyllabification’ and argues that resyllabification takes place when the first syllable ends in a consonant whilst the second starts in a vowel. She cites an example in Mandarin from Chao (1968: 803), which demonstrates the absence of a phonetic onset of toneless syllables of the prosodic word in the same way, as shown in (6):

(6)	[rən^{35}a]	→	[rən^{35}(n)a]	'What a man!'
	[njaŋ^{35}a]	→	[njaŋ^{35}(ŋ)a][8]	'Mum!'

Liaison takes place as a result of nasal gemination in the examples in (5) and (6) because the second syllables of each item in (5) and (6) do not have tones as lexically meaningless syllables so that ALIGN-L does not take effect in this case. Such a toneless syllable is akin to a clitic (see also §2.3.2, ch.2).

In an autosegmental view, when liaison occurs, the final consonant of the preceding syllable becomes the onset of the following syllable, as in French, e.g. *nap* + *lite* → *na.plite* 'tablecloth'. In SX and Mandarin, a stressed syllable must be bimoraic, and the coda consonant is moraic and a TBU, as in (5) and (6), so that when liaison occurs, the final consonant of the preceding syllable geminates and becomes the coda of the first syllable as well as the onset of the second syllable. Therefore, the weight of the syllable is also correlated with the lexical boundary between the syllables: An unstressed syllable in SX cannot be a full-tone TBU. Thus, such a syllable is an onsetless syllable phonologically and phonetically, surviving ALIGN-L. As a result, liaison in the examples in (5) and (6) occurs. There is no comparable account on the assumption that vowel-initial words obligatorily have a zero initial, because this would be a general constraint on SX (or Mandarin) words, and no distinction between full-tone (lexically meaningful) and toneless (lexically meaningless) words would be expected. Thus, ALIGN-L in (4) can also be understood as a constraint against liaison across lexical boundary between two full-tone syllables in SX and Mandarin, as shown in (3a) and (3d).

The possible occurrence of liaison supports the idea that there are vowel-initial or glide-initial syllables in SX as well as Mandarin underlyingly. Thus, besides toneless syllables which have no onset, there are also phonologically V (or single syllabic C), GV, VC and GVC/GVV syllable types with full tones, as shown in (7):

[8] In Mandarin, the alveolar nasal does not become palatalized when followed by [i] or [j].

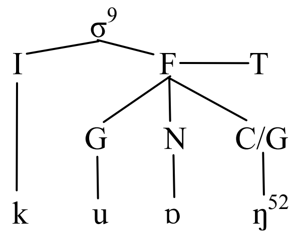
(7) a. C /ŋ̍^{31}/ 'fish'
b. V /a^{35}/ 'short'
c. GV /jɤ^{52}/ 'low (voice)'
d. VC /ɪʔ^{5}/ 'one'
e. GVC /jaŋ^{52}/ 'seedling'
f. GVV /jɑɒ^{33}/ 'want'

However, a (large) minority of speakers use a glottal stop for all words with a zero initial (Chao 1968). The syllables in (7b-f) are also phonetically transcribed as [ʔa^{35}], [ʔjɤ^{52}], [ʔɪʔ^{5}], [ʔjaŋ^{52}], and [ʔjɑɒ^{33}], respectively, following the glottal stop insertion (GSI) rule in (9) in chapter 2. The initial glottal stop [ʔ] in this case is phonetically present, but does not have a phonological position in the underlying syllable structure, since it is fully predictable. This was also discussed in chapter 2. In summary, there are altogether 11 syllable types in the syllable inventory of SX, as shown in (2) and (7).

4.3 The Internal Syllable Structure

While all Chinese languages and dialects have different inventories and arrangements of phonological units, these languages nevertheless show a remarkable similarity with respect to syllable organization. SX, as one of the Chinese dialects, has much in common in its internal syllable structure with other Chinese dialects. It was claimed in traditional Chinese phonology that all Chinese dialects have a syllable structure maximally consisting of an Initial and a Final with a prenuclear glide, a nucleus vowel, a final which can be an off-glide or a consonant, and a tone (Chao 1968; Wang 1985; Norman 1988; among others), as shown in (8) with the example of [kuɒŋ^{52}] 'light' in SX:

(8)

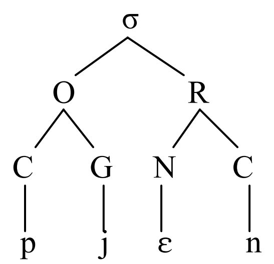


However, as was mentioned in section 1 of this chapter, there has been some controversy with respect to the status of the prenuclear glide since the establishment of the classical OR syllable theory (Pike & Pike 1947; Kuryłowicz 1948; Fudge 1969, 1987; Selkirk 1982, 1984; among others). My analysis, too, will focus on the status of the glide. In this section, I will present an analysis of the different proposals for Chinese syllable structure from the perspective of OR syllable theory, focusing on the prenuclear glide. I claim that OR theory cannot appropriately express the organization of the SX syllable structure. My analysis is divided into six subsections, which each deal with a different approach to syllable structure.

4.3.1 Onset clusters

Some scholars (e.g. Tung 1983; Yin 1989; Bao 1990) argue in favor of an OR analysis of Chinese syllable structure and assume that the prenuclear glide in Chinese syllable structure is the second member of the onset cluster, claiming, therefore, that there is a possible onset complex in Chinese. Taking [pjɛn^{55}] 'edge' (Mandarin) as an example, Bao (1990) represents the syllable structure as in (9) (in the following analysis and throughout this chapter, 'C' is used for consonant in both syllable-initial and syllable-final position):

(9)



[9] In (8), 'T' refers to tone. In traditional Chinese phonology, 'Final' is the tone domain (the location of Tone in the syllable structure will be discussed in chapter 5) and on-glides and off-glides are transcribed as [i] and [u], instead of [j] and [w], respectively.

Bao's argument for this representation is the fact that in all dialects of Chinese the prenuclear glide is not counted in the poetic rhyming system. For example, [twan^{51}] rhymes with [pan^{51}] and [ɕjau^{214}] rhymes with [hau^{214}] in Mandarin. Similar examples also hold for SX, e.g. [tjɑɒ^{35}] 'bird' rhymes with [hɑɒ^{35}] 'good' in SX. According to the OR theory, all the segments in a syllable should be parsed into Onset or Rhyme[10]. Thus if the prenuclear glide is not in the R, it must, by definition, belong to O.

Another piece of evidence for Bao's assumption comes from language games. There is a secret language of Man-t^{h}a (Yip 2003) in the Min language family, in which the sequence [ən] replaces the nucleus and anything that follows it in the first half; in the second half [t^h] replaces the initial consonant. The exact realization of the inserted vowel is quite complex. The lowness of original low nuclei is preserved, but frontness and backness are controlled by the surrounding segments, which can be exemplified with the syllables of [xwey] 'meeting' and [lyaŋ] 'two' as in (10) (the letters 'a', 'b', 'c' above the arrows refer to the different rules that apply; cf. below):

(10) Man-t^{h}a:

	a		b		c	
xwey	→	xwey xwey	→	xwən xwey	→	xwən t^{h}wey
lyaŋ	→	lyaŋ lyaŋ	→	lyɛn lyaŋ	→	lyɛn t^{h}yaŋ

The examples in (10) show that (a) the syllable is reduplicated; (b) the rhyme of the first half is replaced; (c) the onset of the second half is replaced. Importantly, in this language game, medial glides surface in both copies, which they also do in May-ku, another secret language in the Min language family, as shown in (11), in which the syllable is reduplicated, and in which in the first half the sequence [ay] replaces the nucleus and anything that follows it. In the second half [k] replaces the onset. Take [xwey] 'meeting' and [lyaŋ] 'two', for example:

(11) May-ku*:*

	a		b		c	
xwey	→	xwey xwey	→	xway xwey	→	xway kwey
lyaŋ	→	lyaŋ lyaŋ	→	lyay lyaŋ	→	lyay kyaŋ

[10] We will use the spelling *rhyme* to refer to the literary/poetic rhyming system and *rime* to refer to the linguistic constituent in the syllable structure.

The examples in (11) show that the medial glides surface in both copies. In the first half, the medial glides are treated as if they are in the onset, whereas in the second half, they are treated as if they are in the rhyme. This phenomenon has parallels in many language games.

Bao (1990) presents an OR analysis of the May-ku secret language with the assumption that medial glides are in the onset and argues that [t^h] replaces the first consonant of the onset in the second half. His analysis can be summarized as follows:

(a) Reduplicate the syllable.
(b) Assume glides are in the onset and substitute [ay] for the rhyme of the first syllable.
(c) Substitute [k] for the first consonant of the onset of the second syllable.

Apart from several other problems in Bao's analysis,[11] it turns out that language games treat the medial glides differently from dialect to dialect (Chao 1931; Yip 1982, 2003; Walton 1983). Also in Southern Min, there is another secret language La-mi, which treats the prenuclear glide as part of the rhyme. Yip (2003) cites the examples as in (12), which show that the first syllable substitutes [l] for the onset, and the second syllable substitutes [i] for all vocalic material after the onset (the coda remains):

(12) La-mi:

a. hen	→	len hin	'to turn pages'
b. keʔ	→	le kiʔ	'to separate'
c. kya	→	lya ki	'slope'
d. kway	→	lway ki	'strange'

The examples of speech errors from Mandarin (Yip 2003: 792) also show that the medial glides surface in both the onset and the rhyme, as shown in (13):

[11] In the secret language of May-ku, Bao treats [xw] as an onset cluster and thinks that [k] only replaces the first consonant [x] when [w] remains. Such a treatment is against the principle that substitution in a language game always occurs with constituent for constituent, which is usually assumed.

(13)	Speech errors			Rule:
a.	fa xwey[12]	→	xwa xwey	[xw] replaces [f]
	pey kwan	→	pey pan	[p] replaces [kw]
b.	ʂan tʂʰwey	→	ʂan tʂʰwan	[an] replaces [ey], not [wey]
	kʰway ta	→	kʰwa ta	[a] replaces [ay], not [way]
c.	ye çwan	→	ywan çwan	[wan] replaces [e], w in Rhyme

The prenuclear glides in the examples in (13a) behave like Onsets or they fail to behave like Rhymes (13b), but the prenuclear glide in (13c) behaves like Rhyme. However, the affiliation of the glide cannot be predicted from the specifics of neighbouring segments.

The examples above show that secret languages and speech errors may provide valuable psycholinguistic information with respect to syllable structure, and that the proper interpretation of medials as structurally aligning with the Initial or with the Final may vary from Chinese language to language, or from dialect to dialect, or even from person to person. Bao (2000: 307) later claims that “the phonology of a language game need not be identical with the phonology of the language on which the game is based”. However, there is a language game in SX which strongly suggests that the prenuclear glide in SX is not in the Initial but in the Final. It is a simple game popularly played among children, in which a syllable is reduplicated and the rhyme of the first half is replaced by [ɑɒ], as shown in (14):

(14) [ɑɒ] game in SX:

a.pjẽ^{31}	→	pɑɒ^{31} pjẽ^{31}	‘edge’
b.djɑɒ^{22}	→	dɑɒ^{22} djɑɒ^{22}	‘exchange’
c.kʰwoʔ^{5}	→	kʰɑɒ^{5} kʰwoʔ^{5}	‘wide’

The examples in (14) show that in the SX language game the prenuclear glides are replaced together with the rhyme, while the Final of the first syllable is replaced by [ɑɒ], which strongly suggests that the prenuclear glides are not in the onset in SX. Another piece of evidence is the syllable fusion of clitics in SX, in which the syllable of the negation marker [vəʔ^{3}] ‘not’ in SX becomes a clitic, and is phonetically and phonologically fused

[12] In the examples in (13), the off-glides are transcribed as [y] in Yip’s (2003: 792) original version, but as [j] in my transcription, while I transcribe them as [i] if the off-glide is vocalic in this dissertation. In the following examples of language games, the on-glides are also transcribed in [y], following the original source.

with the host syllable. For example, the negative forms [vəʔ^{3}vɛ̃^{31}] 'don't be naughty' and [vəʔ^{3}ʔjɑɒ^{33}] 'do not want to' become [fɛ̃^{52}] and [fjɑɒ^{33}], respectively, after a series of phonological changes (see §5.6 in chapter 5 and Zhang (2005: 69-79) for more information). The phonological changes in these two forms of SX clitics are illustrated in (15):

(15)		a		b		c
[vəʔ^{3}vɛ̃^{31}]	→	[v(əʔ^{3})+(v)ɛ̃ ^{31}]	→	[vɛ̃^{52}]	→	[fɛ̃^{52}]
[vəʔ^{3}ʔjɑɒ^{33}]	→	[v(əʔ^{3})+(ʔ)jɑɒ^{33}]	→	[vjɑɒ^{33}]	→	[fjɑɒ^{33}]

The changes in (15) show that (a) the rhyme of the clitic syllables and the onset of the host syllables are both deleted; (b) the onset of the clitic syllables becomes the onset of the host syllables and the low-pitch tone becomes a high-pitch tone when fused into one syllable (which is made possible only by a change of the register feature. The change or spreading of tonal features will be discussed in detail in chapter 5); (c) the voiced initial obstruent becomes voiceless because of the high-pitched tone (according to the onset-condition constraints in (26) in ch. 2).[13] The original host verb [ʔjɑɒ^{33}] 'want' in (15) contains a prenuclear glide [j]. When the onset is deleted during the process of the syllable fusion, [j] remains in the rhyme. The syllable fusion of clitics in SX is a window on syllabic constituency in SX, and strongly suggests that the prenuclear glide is not in the Initial but in the Final.

To summarize, the poetic rhyming system of SX[14] provides us with evidence that the prenuclear glide is not in the rhyming unit, while the SX language game proves that the prenuclear glide is not in the onset. In spite of the complexity of the status of the prenuclear glide in SX, there is enough evidence to suggest that Bao's (1990) onset-cluster approach is not compatible with SX syllable structure. One other piece of evidence comes from loanwords in SX. As was discussed in chapter 2, there is a constraint *COMPLEX(ONS) which is active in SX loanword phonology. The same is true in all other Chinese dialects, too, including Mandarin. In Zhang (2003: 10), I proposed a basic constraint ranking for loanwords, viz. *COMPLEX(ONS) ≫ MAX-IO ≫ DEP-IO. For example, the optimal

---

[13] Voiced-L:Voiced initial obstruents must have low-register tones on the following vowels; Voiceless-H: Voiceless initial obstruents must have high-register tones on the following vowels

[14] In modern SX, the poetic rhyming system does not include the prenuclear glide, e.g. [dʑjɑɒ^{31}] 'bridge' and [dɑɒ^{31}] 'flea' rhyme perfectly.

loanword of English [brændi] 'brandy' is [baʔlɛ̃di] in SX and [pailanti] in Mandarin, which can be accounted for by an OT analysis as in (16):

(16)

Input /brændi/	*COMPLEX(ONS)	MAX-IO	DEP-IO
a. blɛ̃di	*!		
b. baʔdi		*!	
c. lɛ̃di		*!	
d. ☞ baʔlɛ̃di			**

Another example, which provides evidence with regard to the status of glides, is the loanword *Twain* from English as in *Mark Twain*, in which [tw-] is a complex onset in English. Both in SX and in Mandarin, [tw] is acceptable in principle, e.g. Mandarin [t^hwan^{35}] 'unite' and SX [t^hwə̃^{52}] 'swallow'. However, the status of [w] in SX and Mandarin is different from the status of [w] in English. In the former case, [w], as the pre-nuclear glide, is not in the onset while in the latter case, [w] is in an onset cluster. Thus, the optimal realization of the loanword *Twain* in Mandarin is [t^hu.wen] rather than [t^hwan] or [t^hwen] because of the *COMPLEX(ONS) constraint in Chinese. To capture this behaviour, I propose the following constraint, which preserves onset status:

(17) IDENT-IO(ONS)
A segment which is in onset position in the input corresponds to a segment in onset position in the output.

The loanword realization [tuwen] for English *Twain* can be derived as in (18):

(18)

Input /twein/	*COMPLEX (ONS)	IDENT-IO(ONS)	MAX-IO	DEP-IO
a. tw·en[15]	*!		*	
b. t·wen		*!	*	
c. t·wi.en		*!		
d. t·u.en			**!	*
e. ☞ t·u.w·en			*	*

[15] In order to differentiate between CG as an onset cluster and C·G as a transsyllabic cluster, I put a raised dot between C and G to indicate they are not tautosyllabic.

In the tableau in (18), candidate (a) is ruled out because it violates *COMPLEX(ONS); candidates (b) and (c) are well-formed syllables in Mandarin but [w] in [twen] and [twi.en] are prenuclear glides, violating preservation of syllabic constituency; candidate (d) has both insertion of [u] and deletion of [w] and [i], so (d) is also ruled out; candidate (e) inserts [u] and keeps [w] still in the onset position in the second syllable, satisfying IDENT-IO(ONS), and only violating MAX-IO once. Thus, (e) is the winner. The loanword [baʔlɛ̃di] in (16) also satisfies IDENT-IO(ONS) because when [aʔ] is inserted between [b] and [l], [l] still keeps its onset status. All loanwords with complex onsets follow this rule. More examples from Mandarin are [kɤloŋ] for [kləun] 'clone', [jiŋkɤlan] for [iŋglənd] 'England', etc., and other examples for SX were discussed in chapter 2. Thus, loanword incorporation strongly suggests that there is no complex onset complex in Chinese languages, including SX.

Interestingly, the loanword for [njuː] in [njuːjɔːk] 'New York' is [njou] in Mandarin and [ɲjɤ] in SX, both keeping [j] between the onset consonant and the nucleus vowel with no vowel insertion between as in English [njuː]. At first sight this would seem to be a violation of our constraint *COMPLEX(ONS) or IDENT-IO(ONS), but the solution lies in the fact that neither [j] in English [njuː] nor [j] in Mandarin [njou] or SX [ɲjɤ] is in the onset. [j] in English [juː] is best analyzed as the first element of a diphthong[16] while [w] in English [twein] is the second element of the onset cluster. Chomsky and Halle (1968: 192–193) also observe that the glides [w] and [j] behave differently when in prevocalic position.[17] The examples in (16) and (18) provide strong evidence that there is no onset cluster in the Chinese languages that we have discussed.

A last strong piece of evidence against the onset analysis is the phonological system of *Fanqie,*[18] which will be discussed in the next subsection. In summary, we have so far argued that the prenuclear glide in SX does not form the second member of a complex onset.

---

[16] See 'diphthong' in Trask (1996).

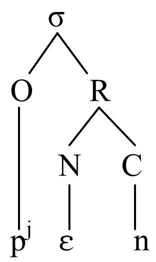
[17] The syllabic status of on-glides is a cross-linguistic issue, which has been widely debated. For example, van der Veer (2006) claims that from a moraic perspective, onglides in Italian can belong to the onset, to the nucleus, sharing a mora with the following vowel, or to a bimoraic nucleus according to their underlying structure.

[18] *Fanqie* is a method for pronouncing Chinese characters in Middle Chinese (6th–10th century). It will be further explained later in this subsection.

4.3.2 Secondary articulation

Duanmu (1990, 2000a) argues that the prenuclear glide is not in the rhyme in Mandarin but disagrees with the assumption made by Tung (1983), Yin (1989) and Bao (1990) that the prenuclear glide is the second member of an onset cluster. Duanmu's (1990, 2000a) argument against an onset cluster analysis in Chinese is based on the principle of the sonority hierarchy (Harris 1983; Selkirk 1984; among others). He argues that if *pj*-, *lw*-, *nj*- or *mj*- were CG clusters in Chinese, there should also be more common clusters such as *pl*-, *kl*-, *pr*-, *kr*-, because the latter have a larger sonority difference between the two segments of the cluster sequence than the former and thus should be more preferred than the former as consonant clusters. For this reason, Duanmu (1990, 1999, 2000a) claims that the prenuclear glide in Chinese forms a secondary articulation on the onset consonant, which can be illustrated as in (19):

(19)



The main arguments on which Duanmu (1990, 1999, 2000a) bases this claim can be summarized as follows:

(a) Howie's (1976) phonetic experiment shows that the prenuclear glide is not a TBU.[19]
(b) The poetic rhyming unit does not include the prenuclear glide.
(c) The Mandarin syllable maximally has CVX; i.e. only three segment slots.
(d) Syllables in Chinese have an obligatory onset. When there is no onset consonant before the nucleus vowel there is an obligatory zero onset; however, prenuclear glides are never preceded by a zero onset.
(e) Phonetically, in syllables such as [swei] 'age' in Mandarin, [s] sounds like [s^w], different from [s] in [swei] 'sway' in English.

[19] Phonologically speaking, the TBU domain coincides with the weight domain, in other words, with the segments that can be considered moraic. In Chinese, the rhyme is the weight domain; this will be discussed later in this chapter.

Although Duanmu presents a list of arguments for his analysis, the data in SX do not support the assumption that the prenuclear glide forms a secondary articulation on the onset consonant. I will go through each point of his argumentation above to see if they fit the data in SX.

Firstly, a number of phonetic experiments have been done to prove that the prenuclear glides are non-moraic and outside the tone domain (Howie 1976; Lin 1995). These experiments may provide strong evidence that the prenuclear glides are not in the rhyme in Chinese, but they do not prove that the glide is a secondary articulation. The secondary articulation analysis might lead one to expect that the length of C is approximately equal to CG. However, this is not the case: Lu (2005: 20) performed some acoustic experiments[20] which show that "the significant results of all CGVX syllables longer than CVX syllables suggest that the extra duration indeed comes from the contribution of the glide". There is also cross-linguistic evidence that CG always results in compensatory lengthening for the syllable even if CG becomes a contour segment or complex segment[21] in surface representation (see Sagey 1986a, 1986b).

Secondly, it is true that the prenuclear glide is not counted in the rhyming unit of the modern poetic rhyming system in Chinese languages, including SX. However, in traditional Chinese phonology, the Chinese syllable is divided into Initial and Final, which differs from the division into Onset and Rhyme in modern syllable theory, and the prenuclear glide is included in the Final, but excluded from the Rhyme. A strong piece of

[20] Lu's (2005: 16-18) acoustic experiments with Mandarin spoken in Taiwan were done in the University of Illinois at Urbana-Champaign. The experiments test the durations of CGVX and CVX syllables in different conditions, including word-initial position, word-final position, in the carrier sentence or in a natural sentence. The results show that the word-initial CGVX syllables in a carrier sentence will have the longest duration, with word-final CVX syllables in a natural sentence to be the shortest (within CGVX: 278 vs 238ms, t=9.54, p<.00001, within CVX: 272 vs 228ms, t=10.86, p<.00001) and syllable in the initial position is longer than in the final position (within CGVX: 271 vs 245ms, t=5.79, p<.00001, with CVX: 263 vs 237 ms; t=6.15, p<.00001). The independent t-test shows that the CGVX syllables are significantly longer than CVX syllables (CGVX 257 ms vs CVX 250 ms; t=2.501, p=.012). A paired T-test was also performed and the significant level of the results was further enhanced (t=3.847, p<.0001). The CVX syllables in general are 3% shorter than the CGVX syllables.

[21] In Sagey's (1985) analysis of Kinyarwanda (one of the Bantu languages), both contour segment and complex segment occupy one segment slot. She claims that if CG becomes a contour segment or complex segment in Kinyarwanda, the x-slot left by G will be taken by the adjacent vowel, resulting in compensatory lengthening.

evidence against the assumption of secondary articulation is the phonological system of *fanqie*.

Fanqie was a guidebook to pronouncing the Chinese characters in Middle Chinese, applied in the ancient Chinese phonology books such as *Qieyun* (Shui 601), *Tangyun* (Sun 751) and *Guangyun* (Chen 1008).[22] Basically, the *fanqie* method uses two characters to represent the pronunciation of a syllable for a new character. The Initial is represented by one character with the same onset, and the Final and the tone are combined and represented with the other character which has a matching rhyme and tone. Thus 'Initial' + 'Final & tone' gives the pronunciation of the whole character/syllable. For example, the syllable [dong] 'east' is represented by [dɤ] + [huŋ], resulting in [d + uŋ]. We can thus see where the ancient scholars put the syllable division. When it comes to syllables with a prenuclear glide, the glide is usually with the second syllable, which suggests that the prenuclear glide is in the rhyme. Consider the following examples[23] of *fanqie* from Wang (2003):

(20)	[du]+[ljau]	→	[djau]	'string'
	[du]+[njɛn]	→	[djɛn]	'field'
	[xu]+[kjɛm]	→	[xjɛm]	'suspect'
	[ɣu]+[kwɑn]	→	[ɣwɑn]	'a surname'
	[du]+[kwɑn]	→	[dwɑn]	'unite'

The examples of *fanqie* in (20) show that the prenuclear glide is with the rhyme, although we occasionally find examples in which the first syllable has a medial glide or both syllables have medial glides. For example, [tjau] is constructed with [tjan]+[ljau]; [ɕje] is constructed with [ɕjoŋ]+[tɕje], etc. (Wang 2003). However, in principle, the *fanqie* system treated the prenuclear glide as part of the rhyme in Chinese. The *fanqie* system, of course, reflects the phonetic perception of Middle Chinese syllables. *Fanqie* may shed light on the phonological system for Middle Chinese and may not represent the syllable structure of modern Chinese. However, modern SX does retain a large number of phonetic and phonological characteristics of Middle Chinese (Karlgren 1915–1926; Chao 1928; Zhan 1991), as was

[22] *Qieyun* (601), *Tangyun* (751) and *Guangyun* (1008) are the three authoritative ancient Chinese phonology books, among which only *Guangyun* has been kept in an intact copy in China. For more information, see Wang (1963, 1985).

[23] The transcription of the examples in (20) is based on the proposals in *Guangyun (cf.* Wang 2003: 54–57). Tones are omitted, since they are irrelevant for our point, and because the exact pronunciation of the tones in Middle Chinese is unclear.

mentioned in previous chapters. This at least suggests that SX may still follow the syllable division as suggested by the *Fanqie* system.

Thirdly, Duanmu (1999) presents another reason for his secondary articulation assumption: the Mandarin syllable has maximally three segment slots: CVX, which has no position for the prenuclear glide in the syllable structure. Duanmu bases his assumption of three segments per syllable on the phonetic phenomenon that all stressed syllables of Mandarin are roughly equal in length. Thus, in his view, if G has its own segment slot, CGVX should be longer in duration than CVX. However, Lu's (2005) acoustic experiments showed that the duration of CGVX syllables is longer than that of CVX syllables, even though the difference in length between CG and C is small (CGVX 257 ms vs CVX 250 ms; t=2.502, p=.012) in this case, taking into account the fact that onset and medial glide are non-moraic and do not count as weight units (Howie 1976; Lin 1995).

Fourthly, Duanmu (1990, 1999, 2000a) follows the traditional analysis (Wang 1963, 1985; Chao 1968; among others) that the Chinese syllables have an obligatory onset. He assumes that when V is a high vowel, the zero onset is often realized as the corresponding glide, and when V is a mid or low vowel, the zero onset is usually one of the segments [ɦ ɣ ʔ ŋ] in Mandarin, with some speaker variation (Chao 1968; Duanmu 1999). However, the zero onset only occurs when the onset is empty, because every Chinese syllable has, and also must have, one onset slot in Duanmu's view. But when the onset is filled by a glide, no zero consonant occurs. Below are some examples from Duanmu (1999: 479, 481):

(21)	a. aw	ɦaw	ɣaw	ʔaw	ŋaw	'concave'
	cf.					
	b. waa	(*ɦwaa	*ɣwaa	*ʔwaa	*ŋwaa)	'frog'
	c. jaa	(*ɦjaa	*ɣjaa	*ʔjaa	*ŋjaa)	'crow'

The examples in (21) show that the phonetic onset (if there is one) does not occur before [w] or [j] whilst it does occur before a vowel-initial syllable. Duanmu first made the claim that the prenuclear G is in the onset. Then he assumes that CG forms one single sound, for G is only a secondary articulation. On such an assumption, he explains that the reason why the zero onset cannot be used with a prenuclear glide is that the zero onset is not an independent phoneme and cannot have a secondary articulation. In Duanmu's view, the zero onset should also occur before a prenuclear glide if this is not a secondary articulation. On Duanmu's secondary articulation assumption, a CGV(C) syllable is possible only because G is a secondary articulation depending on the preceding C and *ʔG, *ŋG, *ɣG, or *ɦG is not possible, as shown in (21), because G, as a secondary articulation, cannot be realized in a zero onset. Duanmu may not have realized the fact that [w] and [j] in (21b) and (21c) are the onset by themselves so that it is clear enough that the zero onset never occurs in the case of (21b) and (21c) because the onset is not empty and as it is, zero onset never occurs before C (e.g. *ʔC). We should not confuse the prenuclear glides and the onset glides when talking about the Chinese syllable structure. The former always refers to G between C and V in CGV, and the latter is G in GV unless G cannot be in the onset of a language in question. However, the glides in Mandarin can be in the onset (Wang 1963, 2003), which is strongly supported by the existence of such syllables as [wu] and [ji] in Mandarin like many other languages. But the glides in SX cannot be in the onset and there is no such syllable as *[wu] or *[ji] in SX, because [(C)wu] and [(C)ji] badly violate OCP(H) (see (83), ch.2) when both segments are in the Final constituent. It is also true that the glides [w] and [j] can be preceded by the phonetic onset [ʔ] in SX when there is no other consonant in the onset, as exemplified in (22):

(22)	[ʔwe^{52}]	'feed'	[ʔjɤ^{52}]	'low (voice)'
	[ʔwa^{52}]	'slope'	[ʔjɑɒ^{33}]	'want'

The glottal stop in the examples in (22) is only a phonetic onset because a full-tone syllable must have an onset in surface representation, which was discussed in §4.2. The insertion of the glottal stop [ʔ] as a phonetic onset for the purpose of assigning tone feature ([+stiff]/[+slack]) (which will be discussed further in chapter 5), as in (22), suggests that prenuclear glides in SX are not in the onset position. In fact, the impossibility of the zero onset occurring before the glides, as in (21), also suggests that the glides in Mandarin can be in the onset as independent phonemes.

Finally, Duanmu (1990, 1999) contrasted the secondary articulation in syllables such as [swan^{55}] 'sour' and [swei^{51}] 'age' in Mandarin, in which [s] sounds like labialised [s^{w}], with words like [swɒn] 'swan' and [swei] 'sway' in English, in which [s] does sound as labialized as [s] in the Mandarin [swan^{55}] and [swei^{51}], and he argues that the difference in pronunciation corresponds to a difference in syllabic affiliation: secondary articulation in Mandarin, cluster in English. It is true that some lip rounding is expected to assimilate in words like [s^{w}wan^{55}] and [s^{w}wei^{51}] mentioned above, but we prefer to put this down to some form of fast-speech assimilation; at any rate it does not tell us anything about the underlying syllabic status of the prenuclear glide.

Duanmu (1999) and Wang (1999) argue that the treatment of prenuclear glides as secondary articulation on onset consonants can reduce the number of finals or rhymes[24] in Mandarin from 37 to 11. However, this analysis increases the number of phonemic consonants which can precede a rhyme from 18 to 50, producing a very large and unnatural consonant system. If we try to lay out a system of underlying phonemes, the picture will be much worse because if palatalised consonants and labialised consonants are all independent consonants, they must all be phonemic consonants. Consider the following examples: (where the consonants show secondary articulation on Duanmu's assumption):

(23)	[tan^{55}]	'single'	[t^{h}an^{55}]	'beach'
	[t^{j}an^{55}]	'tumble'	[t^{hj}an^{55}]	'sky'
	[t^{w}an^{55}]	'end'	[t^{hw}an^{55}]	'rapid (flow of water)'

The examples in (23) show that [t], [t^{j}], [t^{w}], [t^{h}], [t^{hj}], and [t^{hw}] would all be phonemic consonants in Mandarin because they are all distinctive in the minimal pairs. The consonant system of Mandarin would become extremely involved if all palatalized and labialized stops, affricates, fricatives, nasals, and the lateral are also counted as phonemic consonants, as Duanmu (1990, 1999) and Wang (1999) assume. More important, however, is that there is no phonological evidence that the prenuclear glides form a secondary articulation on the onset consonants. As was discussed in previous chapters, there is sufficient evidence to prove that the prenuclear glides in SX are not in the onset.

Another piece of cross-linguistic evidence is related to the fact that if a consonant with a secondary articulation is a phonemic consonant, i.e. contrastive with a plain counterpart consonant, it may well occur in both the onset and the coda if the language in question also allows the plain consonant in the coda position (Laver 1994). Consider the following examples from different languages:[25]

[24] They include all types of combinations of VC, GV, GVC and GVG as well as all simple vowels in surface representation.

[25] The examples in (24) of the segments with secondary articulation in the coda position are all from Laver (1994). The example for the onset position in (24a) is from Padgett (2001); the one in (24b) is from Grønnum & Basbøll (2001); the one in (24c) is from www.utexas.edu/courses/lin380k/7.pdf;

(24) a. Velarization in Russian:

[dal]	‘distance’	[dal^{ɣ}]	‘gave’
[bi:]	‘be(imp.)’	[b^{ɣ}i:]	‘yellow’

b. Laryngealization in Danish:

[vɛn]	‘friend’	[vɛn̰]	‘turn’
['hɛnʌ]	‘happens’	['hɛn̰ʌ]	‘hands’

c. Pharyngealization in Qatari Arabic:

[χal]	‘to leave’	[χal^{ʕ}]	‘vinegar’
[sad]	‘to prevail’	[s^{ʕ}ad]	‘name of letter’

d. Labialization in Tabassaran:

[naq]	‘yesterday’	[naq^{w}]	‘grave’

The examples in (24) show that the phonemic consonants with secondary articulation in the languages above contrast with the plain consonants in coda position as well as in onset position. However, in Mandarin, the alveolar nasal /n/ occurs either in coda or in onset and is palatalized or labialized when followed by the prenuclear glide [j] or [w], respectively, but [n^{j}] or [n^{w}] never occur in the coda, as shown in (25):

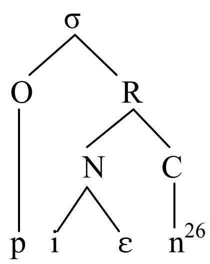
(25) [n^{j}jou^{35}] ‘cow’
[n^{w}wo^{35}] ‘move’
[lan^{35}] ‘blue’
but
*[_n^{j}] #
*[_n^{w}] #

Phonotactically, in SX only the velar nasal /ŋ/ can occur in either the onset or the coda, and never occurs before a prenuclear glide (see ((70) and (73)). In short, we conclude that secondary articulation in SX, both palatalization and labialization, is only a phonetic phenomenon of assimilation, caused by the medial glides in surface representation.

4.3.3 The rhyme constituent

The previous analysis shows that the prenuclear glide in SX is not in the onset – neither in the onset cluster, nor as a secondary articulation –, suggesting that it may well be in the rhyme. Some authors (e.g. Wang & Chang 2001, among others) argue that the prenuclear glide is in the rhyme constituent in the Chinese syllable, as illustrated in the following syllable structure:

(26)



This argument is mainly based on the OR syllable model and the fact that the prenuclear glide does not belong to the onset. Since the prosodic hierarchy is usually assumed to demand exhaustive parsing of entities into the immediately higher constituent, this would mean that segments should all be parsed into Onset or Rhyme (Yip 2003). Thus, if a segment does not belong to the onset, it must be in the rhyme. Based on the traditional Chinese phonological system which treated the prenuclear glide as the rhyme part of the syllable (as *fanqie* did, which was discussed in the previous subsection), Wang & Chang (2001) performed an experiment in which Chinese subjects without any phonological knowledge were asked to fuse two syllables into one, to see which one is preferred, as shown in (27):

(27) tao … pjan → tan/tjan;
pjan … tao → pao/pjao;

In another experiment, Wang and Chang (2001) asked the subjects to decide the preferred way to break a syllable into two, providing two alternatives to choose from, as shown in (28):

(28) pjao → pjan + tjao / pan + tjao

The results of the two experiments showed that the subjects preferred to classify the glides with the rhyme, which follows the traditional

[26] In (26), the glide is transcribed as [i] rather than [j], for it is in the nucleus and [iε] is a diphthong if GV is under the nucleus node.

phonological system that classifies the prenuclear glides as part of the rhyme. This also shows that the prenuclear glides are not in the onset. The question that still needs to be answered is that if the prenuclear glide is not in the onset, is it necessarily in the rime? This question is fundamental to the well supported OR syllable model, especially since there is strong evidence that the prenuclear glide in SX is neither in the onset, nor in the rime.

In one sense, I follow Duanmu's (1990, 1999) and Wang's (1999) argument that the prenuclear glides are not counted in the Chinese poetic rhyming system. This is uncontroversial in the literature of Chinese, including SX. For example, [tɕjan^{55}] 'wicked' and [ʂan^{55}] 'mountain' rhyme, and [tɕ^{h}jau^{55}] 'knock' and [tau^{55}] 'knife' rhyme in Mandarin. As a matter of fact, these syllables rhyme with each other not because the prenuclear glides are excluded from the rhyming system but because in none of these syllables was there a prenuclear glide in Middle Chinese. According to the *fanqie* system in *Guangyun* (1008), the syllables for *wicked* and *mountain* were [kan] and [san], respectively; and those for *knock* and *knife* were [kau] and [tɑu], respectively (Wang 2003), just as in modern SX, the syllables for *wicked* and *mountain* are [kɛ̃^{52}] and [sɛ̃^{52}][27], respectively, and the syllables for *knock* and *knife* are [kɑɒ^{52}] and [tɑɒ^{52}], respectively, both pairs forming perfect rhymes.

I assume that speakers may still regard these syllables in Mandarin as rhymes just as in Middle Chinese, although the rhymes of the syllables have greatly changed over many centuries. However, this assumption lacks evidence and needs further support. The *fanqie* system reflected the phonology of Middle Chinese (about 1000 years ago) which cannot really provide evidence about the syllable structure of modern Chinese. In modern SX, the prenuclear glides are not counted in the poetic rhyming system, although some suggest that the poetic rhyming system should not be the same thing as the rhyme of a syllable structure (Yip 2003) and that poetic matching frequently involves things that are not constituents in any theory of syllable structure (Holtman 1996). However, the concept of the rhyme of a syllable structure originally comes from the rhyming table of the Middle Chinese phonology (Chao 1941; Karlgren 1954; among others) which divided a (stressed) syllable into two parts: '*shengmu*', identical to the onset, and '*yunmu*', the rhyming part in the Chinese poetic writing system between the 7th and 10th centuries A.D.

[27] See chapter 3 for detailed discussion that the nasalized vowel [ɛ̃] is the phonetic realization of underlying /an/. Neither of the syllables involves the prenuclear glide [j].

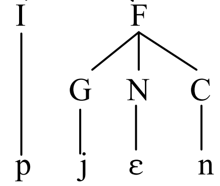
The evidence that the prenuclear glides are not in the rhyme in SX is not only related to the poetic rhyming system, but also to the phonetic and phonological properties of the prenuclear glides in the syllable organization. Phonetically speaking, Howie's (1976) and Lin's (1995) phonetic experiments show that tone in Chinese is carried by the rhyme, in that the F_0 contour does not start during the period of voicing of a (voiced) onset consonant or during the prenuclear glide, but at the beginning of the nuclear vowel. Thus, the prenuclear glide in Chinese is not part of the TBU. If we wish to identify the TBU as a single, identifiable phonological domain, i.e. as the rhyme, parallel to cases of many other languages, we must conclude that the prenuclear glide is neither in the onset nor in the rhyme in SX. We will explore the implications of this view in the next section.

4.3.4 Head of the Final

The analysis presented above suggests that the prenuclear glides are neither in the onset nor in the rhyme in SX. This casts doubt on the classical OR syllable model. Almost all of the argumentation surrounding the status of the prenuclear glides in Chinese is cast within the OR syllable model, so that only two views are possible: one that argues that the prenuclear glide is in the onset (because it is not in the rhyme), and the other that argues that the prenuclear glide is in the rhyme (because it is not in the onset). There is support, as we have seen, for and against both views.

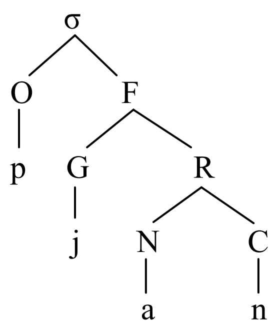
In traditional Chinese phonology, Chinese syllables were also divided into two parts: Initial and Final (not Onset and Rhyme). Initial corresponds to the onset in general linguistic terms; Final is all that is left after the Initial, including the prenuclear glide, the nucleus and the coda or offglide. This can be illustrated in the following tree diagram (Xue 1986, among others):

(29)



The syllable structure in (29) reflects the traditional Chinese phonology view on syllable structure in Middle Chinese. However, this traditional syllable structure becomes problematic in the face of the modern Chinese dialects in that (a) the Final is not the rhyming unit in poetic writing; (b) the glide is not part of the TBU, so that N and C as a rhyming unit and as TBU should be a single constituent in the structure, excluding G. This issue has generated much discussion both at home and abroad. Many authors (e.g. Cheng 1966; Lin 1989, 1990; among others) have made great efforts to combine the traditional Chinese syllable structure with the OR model, developing a new Chinese syllable template, in the hope of rescuing the well established OR theory and adhering to the organization of Chinese syllable constituents. Among other things, this led to the syllable structure proposed for Chinese by Cheng (1966) and Lin (1989, 1990), among others, as is represented in (30):

(30)



Cheng (1966) claims that in the Chinese syllable structure, as shown in (30), G is the head of the Final and outside of R, which is a constituent composed of N and C. The notion that G is head of F remains problematic because heads are always obligatory, while G is optional in Chinese. However, this syllable structure explains how segments are parsed into syllabic constituents under one syllable node in Chinese phonology. In this structure, R is the weight unit of the syllable and also the rhyming unit in literary writing; F is also a constituent structure higher than R, which is a remarkable sub-constituent only postulated, as far as I am aware, for Chinese. The existence of the Final is supported in traditional Chinese phonology, especially on the basis of the *fanqie* system and different language games. It is true that the prenuclear glide is treated differently in different language games or from dialect to dialect, as was discussed above. I assume that this is because the F structure is available in the syllable. In some language games, the onset is replaced so that the prenuclear glide is left with the rhyme, while in other games the rhyme is replaced so that the glide is left with the onset. On this view, the prenuclear glide is neither in the rhyme nor in the onset. However, structurally speaking, if one were to make a break in the syllable in terms of paradigmatic substitution, it is obvious that the most likely break would be between the Initial and the rest of the syllable, the Final (Walton 1983). The evidence for the F constituent in SX syllables, then, can be summarized as follows:

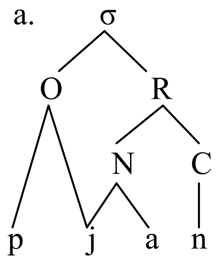
(a) The prenuclear glides are not in the onset, as suggested by loanword incorporation and language games in SX.
(b) The prenuclear glides are not in the rhyme, which is suggested by the poetic rhyming system in SX and also proven by phonetic experiments (Howie 1976; Lin 1995) which show that medial glides are not part of the TBU in Chinese.
(c) The prenuclear glides and the rhyme are treated as one constituent in the SX language game.
(d) The prenuclear glides are usually treated as the rhyme part in the *fanqie* system of Middle Chinese, many phonetic and phonological characteristics of which are still retained in SX.

However, the syllable structure in (30) does not seem to be compatible with modern syllable theory. The problematic area with the syllable structure in (30) is, in my opinion, also a conceptual problem. If G is the head of F, O must be the head of σ (Syllable). Structurally speaking, a head of a constituent is an obligatory element. For example, N is the head of R, since N is obligatory. But the fact is that both O and G are optional in SX as well as in Mandarin. Some notions of the proposal in (30) are problematic. Walton (1983) also points out that notions such as 'head of the final' would not be characterized in any explicit way. Besides, the structure in (30) would be language-specific. However, since the syllable is universal, there must be a generalized framework under which the internal structure of syllables could vary, within strictly defined limits, from language to language. I will discuss approaches to this issue in the next section.

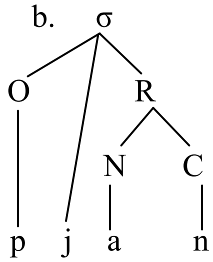
4.3.5 Other options within OR theory

Although the syllable structure in (30) can express the organization of the Chinese syllable in some ways, there are some problems recognizing this as a general syllable template, as was mentioned above. The complexity of the status of the prenuclear glides is mainly caused by the ambiguous role they play in the syllable structure, e.g. in language games of different dialects, or speech errors. Taking the OR syllable theory as the framework to work with and based on the data of language games in different Chinese dialects, Yip (2003) proposes two other options without further discussion, both illustrated below: in (31a), prenuclear glides could be in both Onset and Rime, just as 'ambisyllabic' segments have been argued to be part of two syllables simultaneously; or in (31b), prenuclear glides might be in neither Onset nor Rime, associated directly to the syllable node. These two options are formalized in the following tree diagrams, respectively:

(31) a.



b.



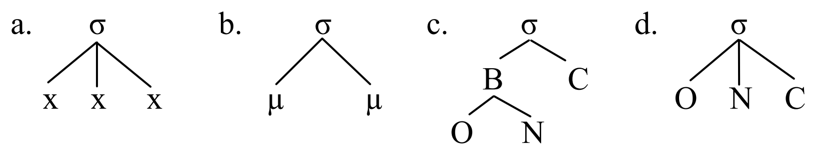
The structure in (31a) shows a representation in which the prenuclear glide is in both Onset and Rime. The structure in (31b) shows a representation in which the prenuclear glide is in neither Onset nor Rime, associated directly to the syllable node. However, both proposals are problematic in one way or another. First, there is hardly any phonological motivation for the prenuclear glide to be in both Onset and Rhyme, similar to the evidence for ambisyllabicity of some consonantal segments. Rather, the evidence against the glide being in the onset as well as the evidence against the glide being in the rhyme combines to weigh against this proposal. Thus, the ambisyllabicity-like medial glide position as in (31a) is not possible for SX syllable structure. The structure in (31b) also has a conceptual problem. As Yip (2003) points out, the prosodic hierarchy is usually thought to demand exhaustive parsing of entities into the immediately higher constituent. There is no particular reason why the medial glide should not observe this principle. Also it is well attested cross-linguistically that a constituent is maximally binary branching (Fudge 1969，1987； Blevins 1995) whether it is of a segment structure or of a syntactic structure, or as Radford (1997: 98) puts it, "all non-terminal nodes are binary". Thus, the structure in (31b) does not fit the binary principle of a prosodic hierarchy; nor does it fit the factual data of the language.

4.3.6 A syntactic approach

In the previous subsections, I have discussed six different approaches to Chinese syllable structure, focusing on the status of the prenuclear glide. Although there are also some problems with the structure in (30), this structure is better suited than others to express the internal syllable organization of SX. I assume that the problem lies with the OR syllable model, which may not be a universal model of syllable structure. The OR model, which goes back to Chinese scholars of the Song dynasty who developed rhyme tables in *Guangyuan* (1008) (Chao 1941; Karlgren 1954, among others), was later developed by Pike & Pike (1947), Kuryłowicz (1948) and many others. However, a number of linguists have realized the limitations of the OR model in structuring syllables in different languages. As Roca (1994: 143) puts it, we cannot assume that the OR model is unanimously accepted, since there are important differences between authors with regard both to the number of constituents they allow for and to the explicit labeling of nodes. There have been many different approaches to the organization of syllables other than OR binary branching model. Blevins (1995: 212) introduces some of these models:

(a) Flat syllable structure (i.e. without sub-constituents other than the segments themselves) (Anderson 1969; Kahn 1976; Clements & Keyser 1983).
(b) Moraic approaches: $\sigma \rightarrow C_0\, \mu\, (\mu)\, C_0$ (Hyman 1985; McCarthy & Prince 1986; Hayes 1989).
(c) Binary branching with Body: $\sigma \rightarrow$ Body Coda; Body $\rightarrow$ Onset Nucleus (McCarthy 1979; Vennemann 1984).
(d) Ternary branching: $\sigma \rightarrow$ Onset Nucleus Coda (Hockett 1955; Haugen 1956; Davis 1985).

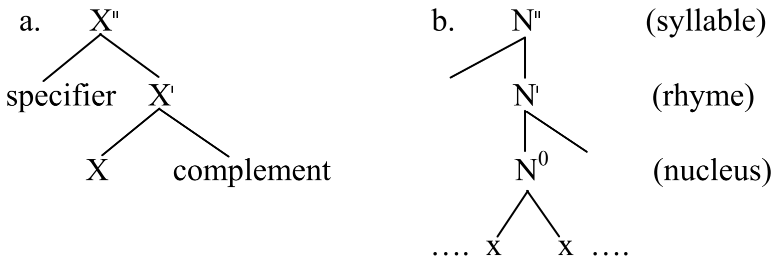
These four models can be represented in (a), (b), (c) and (d), respectively, as follows:

(32) 

Apart from these attempts, there are even more approaches to the syllable. Hall (1992) claims that the syllable in German dominates the subsyllabic constituents of onset, nucleus, and coda, without a rhyme. Geudens & Sandra (2001) present an experiment on acquisition of syllable structure and argue that Dutch-speaking pre-reading children between six and seven show no evidence for Onset and Rhyme units in their explicit phonological awareness. Based on the status of the prenuclear glide of Chinese, Yip (2003) argues that evidence for the existence of Onset and Rhyme constituents is scanty and inconsistent, and that the simpler models are sufficient to capture the facts.

Levin (1985, and later Ritter 1995; Calabrese 1999) argues that Syllable structure is universally of an X-bar type (33a) so that each syllable that is a maximal projection contains one and only one endocentric head, and the rhyme and syllable levels as projections of the head. This can be formalized as in (33b):

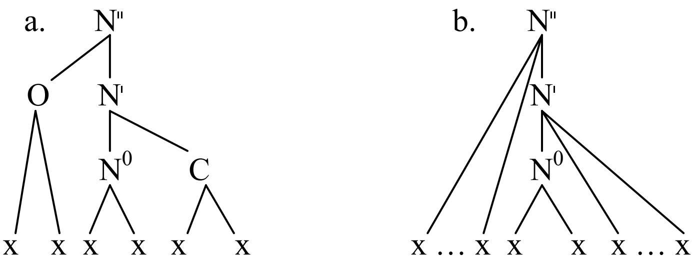
(33)



The structure in (33a) is syntactic X-bar structure. Syntactically, specifier means that the grammatical function is fulfilled by certain types of constituent which (in English) precede the head of their containing phrase. For example, in a sentence such as *What did John do*? *what* is the specifier of the CP headed by the inverted auxiliary *did*; the head is the key word which determines the properties of the phrase, which is the centre of the phrase; the complement is an expression which combines with a head word to project the head into a large structure of essentially the same kind. Moreover, complements bear a close morphological, syntactic and semantic relation to its head.

The structure in (33b) is an X-bar syllable structure which builds on the idea that the syllable is a headed constituent in which the coda and onset can be treated as the complement and specifier, respectively. In syntactic terms, both complement and specifier are in principle optional (Hornstein 2004: 191) and the head is obligatory, which captures the facts of the syllabic organization in SX. Like the OR model, this model (referred to as the X-bar model) must also stipulate that only material adjoined to N^0 is potentially weight-determining.[28] An attractive aspect of this model is its parallel to syntactic structure: it is suggested that sentence-level syntactic structure may have a counterpart in syllable structure, which is also hierarchically structured. From the X-bar model in (33b), it can be assumed that the onset is the specifier of N'', and the coda is the complement of N'. The X-bar model is a nucleus-centered model, because only the nucleus is obligatory in a syllable. Syntactically speaking, N^0 is the head of N' and also the head of N'', i.e. N^0 is the head of the rhyme and also the head of the syllable in terms of syllabic structure. The basic X-bar model can be worked out in various ways, differing, for instance, in the complexities of syllable constituents. Van der Hulst & Ritter (1999) argue in favour of a structure in which an onset and a coda constituent are postulated, both of which are allowed to branch (34a). Another view is to allow multiple adjunctions directly to the N' and N'' levels, as illustrated in (34b):

(34)

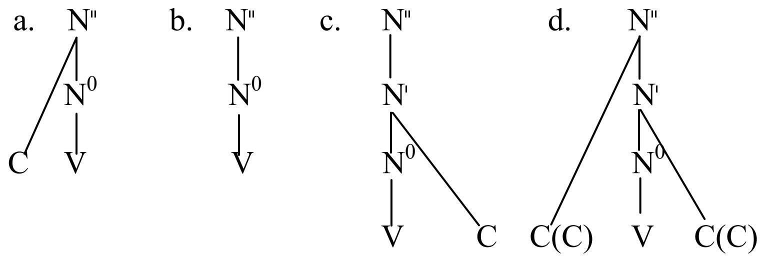


The difference between the two views is related to the complexity of the syllable margins. I will not discuss the difference between these two views. Since SX has no complex onsets or complex codas, the syllabic margins are rather simple, with only one onset consonant and one coda consonant maximally. For the sake of simplicity, I prefer the model in (34b) or simply the model in (33b), which captures the facts of the syllable structure of SX in a straightforward way.

The X-bar syllable structure in (33b) can be generalized as a universal template for syllable structure in that all the syllables of the world's languages must have a nucleus (or syllabic peak), which is the obligatory part, while onset and/or coda are optional, differing from language to language. As van der Hulst & Ritter (1999) explain, the X-bar structure can be implemented in many different ways, allowing various types of syl-

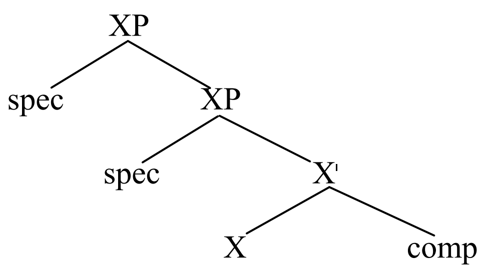
[28] Levin (1985) proposes that whether elements of the Project-N' are weightful is language-specific.

lable structure, including CV, V, VC, CVC and CCVCC, all of which can be represented in the X-bar structure, as shown in (35):

(35)


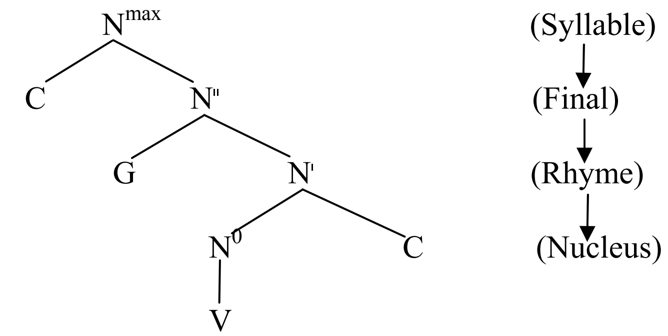
In the X-bar schema, X'' is X^{max} and X^0 is X^{min}. However, both in syntactic structure and in syllabic structure, only X^0 and X^{max} are universally required, as shown in (35). The existence (or non-existence) of an intermediate projection N' is a language-specific matter. Syllabically speaking, if a language has only a simple CV structure (like Hua (Blevins 1995: 217)), it has the $[N''[N]]$ structure; if a language maximally has a CVC structure (like Cairene (Blevins 1995: 217)), it has the $[N''[N'[N]]]$ structure. Blevins (1995) introduces different syllable types in the world's languages, including V, CV, VC, CVC, CCV, CCVC, CVCC, VCC, CCVCC, CVCCC, without discussion of CGVC, the maximal syllable structure of SX as well as all other Chinese dialects. The CGVC syllable structure of SX is different from the CCVC syllable structure of Indo-European languages such as English, because G≠C in SX and G does not belong with either C or V while CC in the latter case belongs to the same constituent.

For syntactic structure, it is claimed that there can be multiple specifiers in an X-bar structure in which there are two X''s dominating X', each X'' having a specifier (see Chomsky 1995: ch.4; Hornstein 1999, 2004: 191), as shown in (36):

(36)


In the multiple-Spec[29] construction in (36), there are two XPs, viz. X^{max} and X'', each having a specifier, and only X is the head.[30] From this syntactic X-bar structure, I just borrow its structural framework for syllable organization without exploiting any parameterization of syntax. However, there is a parallel with the X-bar syllable structure in that every syllable is defined by a unique head or nucleus, and every nucleus defines a unique syllable (Levin 1985). The multiple-Spec X-bar structure in (36) captures the internal organization of the maximal SX syllable structure, CGVC, which can be represented as in (37):

(37)



The maximal SX syllable structure in (37) shows that the onset C is spec of N^{max}, the prenuclear G is spec of N'', and the coda C is comp of N', because C, G and C are all optional in syllabic structure underlyingly. N^0 is the head of N' and also the head of N^{max}, which is the maximal projection, the syllable. The multiple-Spec X-bar syllable structure in (37) also shows that the G is not affiliated directly with either C or V, on the assumption that G belongs to N'', a higher projection than N' and different from N^{max}. With the X-bar structure in (37), I claim that every N (including N' and N'') in (37) is a constituent. Languages with structures more complex than CGVC are very rare because in CGVC each segment belongs with a different constituent. Syllable types such as CCVCC and CVCCC (as in English), can also be subsumed under the $[N''[N'[N]]]$ structure because the prevocalic consonants (if they are in the onset cluster) are both lodged under the N'' and the postvocalic consonants are both (or all) under one N' and/or the N'', depending on whether or not the postvocalic consonants are in the weight domain, according to Levin (1985).

[29] The term 'multiple-Spec' is borrowed from Chomsky (1995: ch.4).
[30] I owe my thanks here to Boban Arsenijević, my colleague in LUCL, for discussing X-bar theory with me.

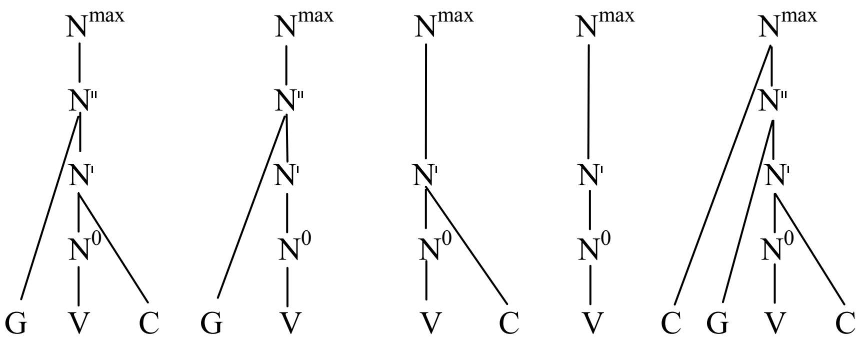
The X-bar syllable structure in (35a) is universal because every language has the CV syllable type (Roca 1994; Blevins 1995), which, however, allows language-specific parameterization of possible syllabic segments (Levin 1985) and different sub-constituents like those in (37). The structures in (35c), (35d) and (37) are all derived from (35a) and are language specific. However, the X-bar template itself is universal.

Harris (1985, 1990) argues that the coda is not needed as a constituent in phonology on assumption that all subsyllabic structure is maximally binary (Fudge 1969; Blevins 1995, among others). In SX syllable structure, complex onsets and complex codas do not exist, so that the onset, medial glide and coda are all terminal slots, not constituents, which is explicitly represented by the multiple-Spec X-bar structure in (37), obviously better so than in the structure in (30). The main difference between the multiple-Spec X-bar syllable structure in (37) and the conventional syllable structure in (30) can be summarized as follows:

(a) The onset and coda are not constituents, for they are the terminals in syllable structure;
(b) The medial glide is not the head but specifier of N'', for G is optional in the SX syllables;
(c) Only N^0 is a head in the syllable structure, for only N^0 is an obligatory element.

In (37), there are three constituents under the syllable node. Each constituent is a projection. The maximal projection is Syllable (N^{max}). There is also the projection of Final (N'') and the projection of Rhyme (N'). The Final is a constituent which can branch into G and N'. The Rhyme is a constituent which can branch into N^0 and C. The Nucleus is also a constituent which can branch into VV (in SX) (or VG in Mandarin) or two moras. Every entity of CGVC in SX can be parsed into the immediately higher constituent in the structure in (37). The existence of the Final constituent suggests that syllables in SX are maximally composed of two parts: Onset and Final. If a syllable has no onset underlyingly, like GVC, GV, VC, or V, its X-bar structure also obligatorily has an N^{max} and an N^0 as CGVC. These structures can be illustrated as follows:

(38)



The X-bar syllable structures in (38), all derived from (35a), have the same N^{max} realization as those in (35), which Levin calls N-placement. The structures in (38) show that G cannot link directly to N^{max} even if there is no onset consonant in the syllable (spec of N^{max}) in underlying form, which strongly suggests that G cannot be the onset in SX, so that a phonetic onset must be inserted as the specifier of N^{max}. In short, the $[N^{max}[N^0]]$ X-bar syllable structure is universal, which allows intermediate nodes N' and/or N'', differently from language to language. I claim that in syllable typology, there should be a different CGVC syllable type like $[N^{max}[N''[N'[N^0]]]]$ which holds for SX and which does not have the problems associated with the proposals that I described in the previous sections. The multiple-Spec X-bar syllable structure in (37) allows three different internal hierarchical constituents, which can behave as a whole, or independently of each other, thus satisfying the different phonological rules involved in poetic rhyming, language games, speech errors, the *Fanqie* system, etc. In addition, I hope that the syllable structure proposed for SX in (37) will also be useful for an analysis of the syllable structure of Mandarin.

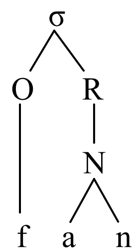
4.4 The Coda in Shaoxing

According to the SX syllable structure in (37), there is a coda position, which seems well motivated by the SX data in such examples as in (2) and (7) above. However, there is a strict constraint on the syllable coda in SX in that only a glottal stop or velar nasal is acceptable in the coda (see CODA-COND (18), ch.2). This leads to an extremely asymmetrical picture that in the onset all the 29 consonants in SX can occur while in the coda only [ʔ] or [ŋ] are acceptable. This phenomenon gives rise to the natural question whether it is really necessary to postulate a full coda position in the SX syllable. In this section, I will discuss the phonetic and phonological properties of the coda consonants in SX. I claim that the glottal stop [ʔ] is the result of debuccalization of syllable final stops and the velar nasal [ŋ] is what has remained as a coda when the other two nasals (bilabial nasal [m] and alveolar nasal [n]) were debuccalized, resulting in the nasalization of the preceding vowels. The debuccalization of syllable final consonants is caused by diachronic attrition, which will be discussed in this subsection.

4.4.1 Previous argumentations for different dialects

Because of the extremely limited number of the coda consonants, some researchers have claimed that there is no real coda in the underlying syllable-final position in Chinese dialects such as Beijing (on which Mandarin is mainly based) (Wang 1993, 1999; Duanmu 1999) and Shanghai (which is one of the Wu dialects, just like SX) (Duanmu 1999). Wang argues that there is no coda in Beijing because the five postnuclear segments in the syllables are [j], [w], [n], [ŋ] and [r], all of which are sonorant. Wang (1999: 226) assumes that the Beijing syllable finals [n], [ŋ] and [r] are actually all glides and vocalic, just like [j] and [w], so that all Beijing syllables are open, without a final C. Wang's syllable structure of Beijing, taking [fan^{55}] 'turn over' as an example, is represented as follows in the OR model:

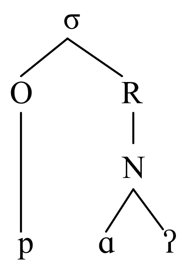
(39)



Duanmu (1999) argues that Shanghai has no coda consonants underlyingly and all syllables have a CV structure. Duanmu (1999: 494) assumes that the VN combinations such as [ən in yn oŋ] can be phonemically analyzed as nasalized vowels [ə̃ ĩ ỹ õ] and all [Vʔ] combinations are underlyingly laryngealized vowels (written as [V']). In his view, the syllable-final C in Shanghai only occurs in surface representation and it will be deleted when unstressed.

Chan (1997) proposes that in Fuzhou (one of the Min dialects) the syllable-final glottal stop [ʔ] is placed in the nucleus with a plain vowel preceding, e.g. [pɑʔ^5] 'white', whose syllable structure can be represented in the OR model as in (40):

(40)



Chan (1997: 279) assumes that [ʔ] in Fuzhou [pɑʔ^5] is disappearing glottal stop and that it occupies the same position as an off-glide in the nucleus, forming an open syllable so that it has the same length as an open syllable ending in a vowel. What Chen means by "disappearing glottal stop" is, as I understand, that the glottal stop is never like a common syllable-final consonantal stop with specific place articulation and it has no robust phonetic realization. In Shanghai and Fuzhou, the syllable-final [ʔ] is assumed to be dropped, for the sake of the syllable length, either resulting in shorter duration, as in Shanghai (Duanmu 1999), or in longer duration, as in Fuzhou (Chan 1997). The motivation of the two proposals is to assume that there is no coda consonant in Mandarin or in Shanghai or in Fuzhou. However, I assume that both the final nasal and the final glottal stop are in coda position in SX, which will be discussed in the following subsections.

4.4.2 Debuccalization in SX

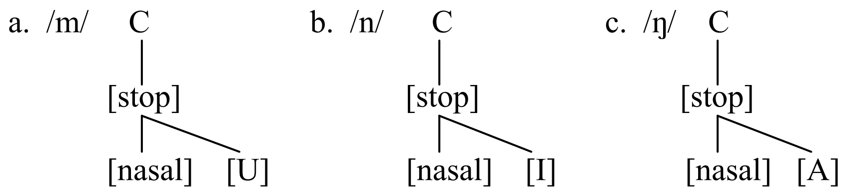
As is well known, the syllables in Middle Chinese allowed a number of coda consonants, including the nasals /m/, /n/ and /ŋ/ and the stops /p/, /t/ and /k/ (Wang 1963, 1985; Zhan 1991; Liu 2001; among others). Wang (1963, 2003) and Liu (2001) divided the Middle Chinese syllables into three categories: *yin rhyme, yang rhyme* and *ru rhyme*. *Yin rhyme* refers to syllables ending with a vowel, also called open syllables; *yang rhyme* refers to syllables ending with nasals; and *ru rhyme* refers to syllables ending in stops. Through diachronic attrition, the syllable-final nasals and stops were gradually dropped, with differences from dialect to dialect. Mandarin nowadays only has /n/ and /ŋ/, missing all other consonants, and SX has only /ŋ/ and /ʔ/ in syllable-final position. What happened to the SX syllable-final consonants? I assume that /m/ and /n/ and /p/, /t/ and /k/ have not really disappeared but that they have debuccalized.

Debuccalization, also sometimes referred to as de-oralization, is a phonological process in which a consonant segment loses its oral articulation (Trask 1996). Debuccalization is found with stops, fricatives, and nasals cross-linguistically (Humbert 1995). Humbert describes a number of processes of syllable attrition in which a syllable-final consonant is debuccalized. In such processes, the segmental place component is lost, but the manner component often remains and is interpreted as a (new) segment. A number of authors (Chomsky & Halle 1968; Clements 1985; Steriade 1987) claim that /ʔ/ and /h/ have no (oral) place component. Based on a unary feature system (van de Weijer 1994, 1996), the segment structures of /p t k ʔ/ (which were claimed to be the syllable codas in Middle Chinese) could be represented as follows:[31]

(41)	a. /p/	b. /t/	c. /k/	d. /ʔ/
	C	C	C	C
	|	|	|	|
	[stop]	[stop]	[stop]	[stop]
	|	|	|	
	[U]	[I]	[A]	

In the segment structures in (41), the three place elements [U], [I] and [A] indicate labial, coronal and velar articulations, respectively. Cross-linguistic evidence suggests that debuccalization processes indeed involve deletion of the entire place component (Clements 1985; Humbert 1995). As for the syllable-final stops, the information that closure takes place in the oral cavity is encoded in the presence of the c-place component. The place elements then specify which articulators are involved. The manner component representing complete closure without a place component implies obstruction in a location outside the oral cavity. The only remaining physiological possibility for closure, then, is the glottis, resulting in a glottal stop /ʔ/, as shown in (41d). Thus /ʔ/ is the new segment resulting from the three stops when debuccalizing through historical attrition in SX. A parallel course of events can be envisaged for the syllable-final nasals /m/, /n/ and /ŋ/, whose segment structures are represented below:

[31] Not shown in (41) are Laryngeal features: oral voiceless consonants have [–voice], while the glottal stop has [+constricted glottis].

(42)


When debuccalizing, nasals lost the place components, with manner components still remaining, just as the final stops did. Nasality, however, cannot be realized in a segment without a place component (Humbert 1995), which leads to a default Place interpretation involving coronality, as was discussed in chapter 2 and 3. Cross-linguistic evidence shows that the coronal Place feature has the most unmarked status in different aspects. (cf. Paradis & Prunet 1989; Rowicka & van de Weijer 1994).

As was discussed above, in SX there is only one velar nasal [ŋ] in the coda position and there are three nasalized vowels in surface representation. I have discussed vowel nasalization in chapters 2 and 3 and assumed that the underlying forms of the nasalized vowels are /VN/. It is also believed that SX had three nasals, [m], [n] and [ŋ] in the coda in the Middle Chinese times just like other Wu dialects (Zhan 1991). I assume that in SX, of the three nasals mentioned above, /m/ and /n/ underwent debuccalization and lost the place components of labial and coronal, spreading the feature of nasality to the preceding vowel, resulting in vowel nasalization, as was discussed in chapter 3, with only [ŋ] remaining as a coda. The phenomenon that the underlying /m/ and /n/ were lost in syllable-final position, triggering the preceding vowel nasalization and /ŋ/ remains a coda nasal in surface representation suggests that the syllable final [ŋ] has the velar place component, which will be discussed further in §4.4.3. Thus, I assume that [m] and [n] were debuccalized but [ŋ] remained as a coda nasal in SX. This assumption gives rise to the question why [m] and [n] underwent debuccalization but [ŋ] did not. This question will be addressed in the next section.

4.4.3 Nasal debuccalization

Debuccalization can be either a diachronic process or a synchronic rule. For example, in some Spanish dialects, [s] debuccalizes into [h] unless it precedes a vowel, e.g. *mismos* [mihmoh] 'similar'. Such a phenomenon of debuccalization is essentially the loss of the supraglottal articulation with retention only of the open glottis gesture (McCarthy 1988: 88). Nasal debuccalization in SX is caused by historical attrition in the same way as debuccalization of syllable-final stops. Through debuccalization, nasals lost their place compenent, with nasality remaining. Since nasality cannot be realized in a form of segment without articulation component, debuccalized nasals get deleted, resulting in the nasalization of the preceding vowel.

In markedness theory,[32] there are at least two separate markedness hierarchies that play a role in phonological processes such as assimilation, epenthesis, lenition and deletion (van der Torre 2003). For example, stops are highly unmarked in the onset while they are the most marked in the nucleus. The places of articulation of [m], [n] and [ŋ] are labial, coronal and dorsal, respectively. Cross-linguistic evidence shows that [m] is the most preferred nasal in onset position while [ŋ] is the most preferred nasal in coda position (see Selkirk 1984; Harris 1994). For example, in almost all Chinese dialects, which have a nasal in onset position, it always includes [m], and when in coda position it always includes [ŋ], but the reverse is not true. Therefore, [ŋ] is the most unmarked and [m] is the most marked in coda position. The reverse is true in onset position. In SX, coda consonants are assumed to be moraic (which will be discussed in §4.5) and it is universally true that the more sonorant a segment is, the more likely it is to be moraic. In element theory, [A] (dorsal) is more sonorant than [I] (coronal) and [U] (labial), and [U] is the least sonorant of the three (Botma 2004). Thus, we have a hierarchical ranking of the three elements for the coda of SX as: *LAB ≫ *COR ≫ *DORS.

Although the debuccalization of the SX syllable final consonants has the tendency of losing the place components, the syllable structure still requires a moraic coda after a [−tense] vowel. To satisfy this requirement, there is a constraint MAX-IO(C) which requires the input coda position to be preserved in the output. This constraint should be highly ranked to guarantee preservation of the coda. Now, we have the constraint ranking for the debuccalization of the SX syllable final nasals, as in (43):

(43) Debuccalization constraint ranking:
MAX-IO(C) ≫ *LAB ≫ *COR ≫ *DORS

[32] Markedness theory refers to any of several approaches which attempt to establish a systematic, principled and (usually) universal distinction between marked and unmarked forms. The best-known attempt is that in the last chapter of Chomsky & Halle (1968), which argues that, for every feature in every possible environment, one value will be unmarked. The *SPE* approach is not usually regarded as satisfactory (Trask 1996).

The constraint ranking in (43) captures the fact that a coda is required in the SX syllables if the nucleus vowel is [−tense], e.g. [pɛʔ⁵] 'eight', [zɪʔ³] 'enter', [ɦəʔ³] 'join', [tɕɪŋ⁵²] 'gold', [nəŋ³¹] 'able', [zɒŋ³¹] 'taste', etc. We assume that the preservation of /ŋ/ and /ʔ/ in the SX syllable coda is the result of a compromise between syllable weight requirements and debuccalization, which can be expressed through a constraint-based analysis as in (44) and (45):

(44)

Input /-m, -n, -ŋ/	MAX-IO(C)	*LAB	*COR	*DORS
a. [-m, -n, -ŋ]		*!	*	*
b. [-n, -ŋ]			*!	*
c. ☞ [-ŋ]				*
d. *[C]	*!			

The tableau in (44) shows that candidate (d) does not allow any coda, violating MAX-IO(C), and is ruled out; among candidates (a), (b) and (c), the more nasals are kept, the more violations there are against debuccalization. Thus, candidate (c) is the winner because it violates the least important constraints. As a result, SX still allows syllable-final nasal /ŋ/ as a coda, rather than bilabial or alveolar nasals. As for the stops, the optimal coda /ʔ/ comes out through the same constraint, as shown in (45):

(45)

Input /-p, -t, -k/	MAX-IO(C)	*LAB	*COR	*DORS
a. [-p, -t, -k]		*!	*	*
b. [-t, -k]			*!	*
c. [-k]				*!
d. ☞ [-ʔ]				
e. *[C]	*!			

The tableau in (45) shows that [ʔ] is the optimal output as the syllable coda in SX for it violates no constraint against place of articulation. In short, in nasal debuccalization in SX, the bilabial nasal may have been lost first and the velar nasal remained, perhaps because velar is the least marked place of articulation among coda consonants (cf. van der Torre 2003) and because the coda position was still required for the sake of the weight of the rhyme in SX. As a result of debuccalization, SX has only two consonant segments, /ŋ/ and /ʔ/, as potential syllable codas. SX underwent debuccalization of all three coda stops, which became realised as /ʔ/. As for the nasals, only /m/ and /n/ debuccalized, resulting in nasalization of the preceding vowel, while /ŋ/ remained in the coda, because the underlying nasal(s) in the vowel nasalization in SX cannot be /ŋ/, but usually /n/, which can be supported by the nasal restoration triggered by liaison (which was discussed in chapter 3), in which the nasal is restored from the nasal vowel when followed by a vowel and the restored nasal becomes the onset of the following syllable. Since the nasality in nasal vowels receives the default Place specification, namely, coronal, the restored nasal is usually [n] in surface representation, as shown in (46):

(46)	Monosyllables:		Disyllables:			
	b. [nɛ̃^{31}]	'difficult'	[nɛ̃^{31}a]	→	[nɛ̃^{31}na]	'Difficult!'
	c. [t^{h}jɛ̃^{52}]	'heaven'	[t^{h}jɛ̃^{52}a]	→	[t^{h}jɛ̃^{52}na]	'God!'

The nasal restoration in (46) shows that the nasal vowel is underlyingly a vowel + nasal sequence which, however, disappears in nasal debuccalization.[33] Moreover, /n/ and /ŋ/ are underlyingly distinctive since underlying minimal pairs like /lan^{13}/ 'lazy' and /laŋ^{13}/ 'cold' exist in SX. Cross-linguistically, in Mandarin which has the same linguistic origin as SX, /kan/ 'dry' and /kaŋ/ 'steel' are phonetically realized by way of different nasalized vowels as [kæ̃] and [kɑ̃], respectively (Wang 1993). Through the analysis above, I assume that the syllable-final nasal [ŋ] in SX does not debuccalize because I assume that the remaining coda nasal [ŋ] still has an articulatory place, which can be explained by the velar nasal gemination in liaison (which was discussed in 4.2), as shown in (47):

(47)	[ts^{h}oŋ^{52}a]	→	[ts^{h}oŋ^{52}ŋa]	'dash ahead'
	[ɲɪŋ^{31}a]	→	[ɲɪŋ^{31}ŋa]	'(how many) people!'

The examples in (47) show that the first syllable-final nasal is ambisyllabified to be the onset of the second syllable through nasal gemination. When the velar nasal [ŋ] becomes the onset, it must have an articulatory place to be checked by certain constraints on the distribution between onset consonants and the following vowels for aspects of Place features. However, once debuccalization has taken place, no articulator remains accessible in the representation, and the feature tree of the seg-

[33] The geminated nasal [n] from the data in (46) is consistent with the reconstructed rhyme [jɛn] when preceded by onset coronal consonants in Middle Chinese, according to *Zhongyuan Phonology* (a book of ancient Chinese phonology, 1324 AD) (see Wang 1963).

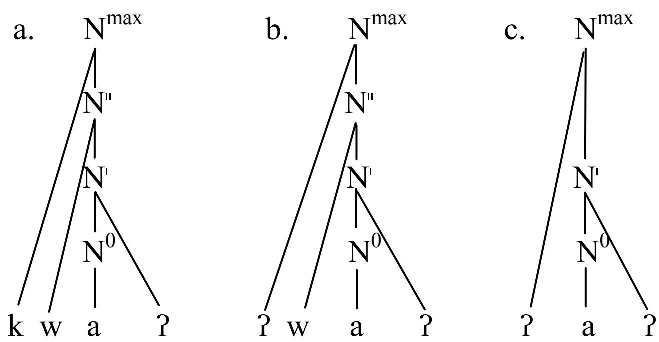
ment is deleted (Halle 1995: 214). The difference between nasal debuccalization and plosive debuccalization is that when all place components get lost, nasality is usually realized by way of nasalization of the preceding vowel and the debuccalized nasal is lost while plosive debuccalization results in a new segment, the glottal stop [ʔ], which has no articulatory place.

4.5 Syllable Weight

Bearing in mind that SX still kept a coda for the sake of potential syllable weight (recall that [−tense] vowels cannot occur in the syllable-final position as the whole weight unit), every syllable in SX is phonetically heavy enough to be stressed or to realize full tones, which was already discussed in chapters 2 and 3. Since the syllable weight in SX is decided by stress, not by structure, a CV syllable can be heavy when stressed and a CGVC syllable can be light when unstressed. In this section, I will present my analysis of the SX syllable structure in terms of weight, focusing on the difference in weight between syllable types. I claim that N' is the weight domain in the SX syllables.

4.5.1 Weight-irrelevance of CG

As was argued in chapter 1, the advantage of the mora model in analyzing syllable structure is to offer a clear-cut picture of the weight of syllables. However, the X-bar syllable structures as shown in (38) also explicitly explains the weight relevance or weight irrelevance of syllabic constituents or non-constituent elements. As was discussed previously in this chapter, the SX syllable structure has three internal constituents: Final, Rhyme and Nucleus, represented as $[N^{max}[N''[N'[N^0]]]]$ in X-bar framework, within which the onset is specifier of N^{max} (Syllable), the prenuclear glide is specifier of N'' (Final), and the coda is complement of N' (Rhyme). The syllable is an N-centered structure in the X-bar framework so that N is an obligatory weight contributor to the syllable. Most phonologists hold that syllables are maximally bimoraic (Hyman 1985; Davis 1999, among others), although there are some arguments for trimoraic structure over superheavy rhymes which, however, do not occur in SX. Weight-by-stress ((47), ch.2) demands that stressed syllables are heavy and bimoraic and unstressed syllables are light and monomoraic for all the syllable types in SX. Some researchers (e.g. Duanmu 1999) claim that heavy syllables must be CVV or CVC and light syllables are CV, assuming that the syllable weight determines the syllable types or the reverse. Moreover, it is well known that the onset is weight-irrelevant in the syllables of the world's languages (McCarthy & Prince 1986; Hayes 1989; Goedemans 1998, among others). Some phonetic experiments (Howie 1976; Lin 1995) show that the prenuclear glides in Chinese are also weight-irrelevant. Thus, as for a CGVC syllable in SX, both C and G before V are non-moraic, i.e. a specifier is weight-irrelevant in syllable structure. Accordingly, such syllables as [kwaʔ^{5}] 'scrub', [ʔwaʔ^{5}] 'dig' and [ʔaʔ^{5}] 'duck' in SX are of the same weight. Their syllable structures can be presented as in (48):

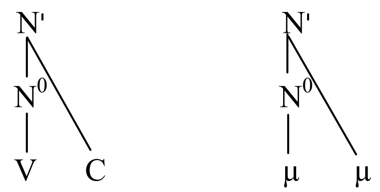
(48)


In (48) the prevocalic consonants, [k] and [w], are adjoined to N^{max} and N'', respectively, which suggests that neither N^{max} nor N'' is the weight domain of the syllable. However, the fact that both C and G are weightless does not indicate that C and G are both in the onset or even in the same constituent. In terms of weight, syllables like CGVC, CVC, GVC, and VC have the same weight as long as VC has a weight the same heavy. Both moraic structure and X-bar structure are consistent with each other, capturing the organizational facts of the SX syllables.

4.5.2 The weight domain in SX

The weight domain is the constituent that directly dominates the segments contributing to weight. The elements of the weight domain can be different from language to language. For example, a glide before the peak vowel can be in the weight domain if it is vocalic, like [j] in English [tjuːn] 'tune'; whereas [j] in SX [tjɤ^{52}] 'lose' is not in the weight domain, as discussed above. Consonants can also contribute to weight according to weight-by-position (Hayes 1989). In the onset they never do,[34] while in the coda they can, depending on the language in question or even on the position in the word.[35] If the coda (which is adjoined to N') is weightless, it is non-moraic, which suggests that N' is not the weight domain of the language in question. Cross-linguistic evidence shows that in most of the world's languages the weight of a syllable depends on the properties of its rhyme (Hyman 1985). In the X-bar structure of SX, $[N'[N^0]]$ is the weight domain which contains both nucleus vowels and coda consonants, because the coda in SX contributes to weight and is therefore moraic. In moraic theory, the approach to distinguishing weight differences is that short vowels project one mora whereas long vowels are necessarily bimoraic and thus heavier (van der Hulst & Ritter 1999). In SX, vowels are unspecified for length. Instead, they are specified as tense vs. lax. All six phonemic vowels are [+tense] and can occur in syllable-final position, forming the whole weight unit. Hence, [+tense] vowels in open syllables (CV) are potentially bimoraic; [−tense] vowels, which cannot occur in open syllables, are unexceptionally monomoraic, so that a syllable with only a [−tense] vowel as the whole weight unit is unacceptable. As it is, every [−tense] vowel must be followed by a coda consonant, making the syllable bimoraic. Thus, the coda consonant in SX contributes to weight. In terms of moras, the VC tier in SX is equal to the moraic tier in the syllable structure, as shown in (49):

(49)

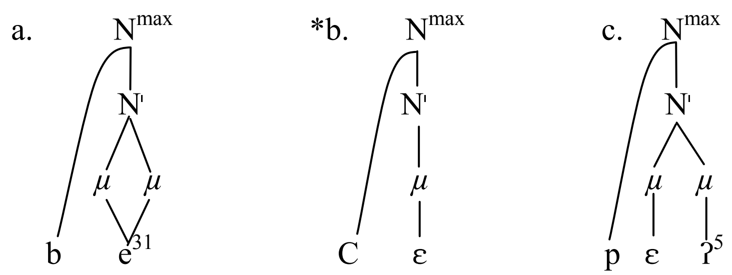


The structure in (49) shows that the syllable final C is moraic in SX. Duanmu (1999) assumes that all syllables of Shanghai Chinese[36] are weak syllables because (i) Shanghai has no diphthongs and vowels are not distinct for length, like SX. (ii) if Shanghai has heavy syllables, there are no

[34] Some assume that in some languages like Pirahã, the onset can also be moraic (see Davis 1985; Everett 1988; Topintzi 2004). Cf. Goedemans (1998) for discussion.
[35] Hayes (1989) assumes that an obstruent in the coda is usually not moraic, but if it is in the position of ambisyllabicity it is moraic.
[36] Shanghai is also one of the Wu dialects like SX and its syllable structure is very similar to that of SX, with no diphthongs. However, it has an alveolar nasal and a velar nasal in coda position as well as the glottal stop [ʔ] (Duanmu 1999).

weak syllables, because all the syllables in Shanghai have the same structure; (iii) there is no phonological evidence that all rhymes in Shanghai are heavy (Duanmu 1999: 493). Weak syllables, however, are all monomoraic (Duanmu 2000a). However, I assume that all syllables in SX are heavy when stressed and can be weak when unstressed. The most important phonological evidence that the rhymes in SX are basically heavy is that contour tones should be realized on two moras or by bimoraic rhymes, because a contour has two tone features, either [hl] or [lh], and one mora can bear at most one tone feature. The bimoraic status is realized by tense vowels or VC combinations if the nucleus vowels are [−tense]. Thus, the two syllable-final consonants in SX, [ŋ] and [ʔ], also contribute to weight. For example, the syllables like [be^{31}] 'compensate' and [pɛʔ^{5}] 'eight' are well-formed because the glottal stop is also moraic. But *[Cɛ][37] is ill-formed in SX for its lack of enough weight when stressed. This can be represented in terms of the weight domain as in (50):

(50)



Syllables like (50b) are not acceptable in SX, as stipulated by the segment filter (see (45) in ch.3). Although syllables ending with the glottal stop is phonetically short and realized in one-digit tone pitch, phonologically, the tones [5] on [pɛʔ] and [31] on [be] are both born by bimoraic syllables, as shown in (50).

Some scholars think that the syllable-final glottal stop /ʔ/ in some Chinese languages is weightless, e.g. Shanghai (Duanmu 1999), Fuzhou (Chan 1997), Cantonese (Yip 2002), etc. However, the two syllable-final consonants, /ŋ/ and /ʔ/, both contribute to weight in SX. Nasals are sonorant enough to be moraic. Obstruents can also be moraic according to weight-by-position (though some (e.g. Hyman 1985) think that an obstruent within the rhyme does not have weight). Trigo (1988) argues, within a binary valued Feature Geometry framework, that /ʔ/ and nasality lose the property of being consonantal when the place component is lost in that

[37] [C] here refers to any existing consonant in SX.

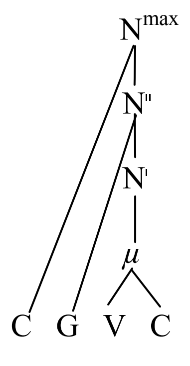
the [+cons] specification is realized by the c-place component and the debuccalized /ʔ/ and nasality become defective for [cons] when the c-place component is lost. The syllable-final /ʔ/ is a defective segment in SX, resulting from the debuccalization of the final stops, during which the place component is lost, as was discussed in the previous section. As a result the syllable-final /ŋ/ and /ʔ/ in SX are both phonetically and phonologically well motivated to be moraic, making the rhyme heavy enough to be bimoraic and get stressed.

Since all [−tense] vowels are allophones of [+tense] vowels, which was discussed in chapter 3, underlyingly all syllables are heavy and bimoraic and light syllables are only phonetically realized as monomoraic in surface representation in SX. As was discussed in chapter 2, the syllable weight is decided by stress and heavy syllables are realized as full tones in SX, viz. any full-tone syllable is heavy and bimoraic, regardless of the syllable types. Not only can a CV syllable be heavy (like (50a)), a CGVC syllable can also be light and monomoraic when unstressed for lexical or syntactic reasons. This phenomenon can be captured by a constraint as in (51):

(51) WSP ≫ MAX-μ[38]

The constraint in (51) stipulates that a syllable is monomoraic unless it is stressed, or vice versa (a syllable is bimoraic unless it is unstressed), which is unexceptionally observed by all the SX syllables. Thus when unstressed, the syllable of CGVC can also be monomoraic, as shown in (52):

(52)



[38] In (51), WSP (Prince 1990) requires a stressed syllable to be bimoraic; MAX-μ (after Kager 1999) requires an unstressed syllable to be monomoraic.

The structure in (52) shows that a maximal CGVC syllable type in SX can be monomoraic when unstressed. Bearing in mind that stress in tone languages is different from that in a stress language, as was discussed in chapter 2, cross-linguistic evidence shows that stress is strongly associated with a number of acoustic features. The stressed syllable is usually marked by a change in pitch, increased length, increased amplitude and changes in the quality of the vowel and the relative absence of spectral tilt (Fry 1958; Lieberman 1960; Sluijter & van Heuven 1996). Unstressed syllables may be signaled acoustically by less intensity, shorter duration, the absence of a spectral change, vowel reduction, steeper spectral tilt or some combination of these. However, the coda is still moraic even in an unstressed monomoraic syllable, as shown in (52). Every syllable in SX should be potentially bimoraic in that stress in SX is decided by the lexical meaning or grammatical function of the syllable, rather than by the syllable type. Stress plays a crucial role in the phonological and morphological structure of languages cross-linguistically. The structures in (50) and (52) also suggest that the moras in the weight domain are the TBUs in SX in that only stressed syllables are heavy enough to bear full tones.

4.6 Phonotactics

4.6.1 Simplicity of segment sequences in SX

Phonotactic constraints are phonological constraints on segment sequences in the syllable domain. Any formulation of generalizations at the segmental level and higher prosodic levels is awkward without reference to the syllable constituent (Féry & van de Vijver 2003). Cross-linguistic evidence shows that there are some universal principles of phonotactics in the world's languages, such as the Sonority Sequencing Principle (SSP) and the Syllable Contact Law (SCL)[39] (see e.g. Butskhrikidze 2002). However, co-occurrence restrictions are to a smaller or larger extent language-specific. For example, *ts-* is perfectly acceptable in syllable-initial position in SX but never syllable-finally, while the reverse is true in English. I have already discussed some phonological constraints on the segmental arrangement in the SX syllable structure in previous chapters. The topic of phonotactics in SX in this section serves as a summary of

[39] The SCL (Syllable Contact Law): In any sequence Ca.Cb, there is a preference for Ca to exceed Cb in sonority (Vennemann 1988).

this chapter since it is always the case that it is hardly possible to analyze syllable constituency in Chinese without going into phonotactics.

As was discussed above, the X-bar structure in (37) very well captures the segmental organization of the syllable structure in SX, which tells us that the maximal segment sequences of a SX syllable is CGVC, each segment belonging to a different hierarchical constituent. Except for [ɑɒ] as in [hɑɒ^{35}] 'good', which is a combination of vowels (VV) in the nucleus, there is no two-segment sequence under a minimal sub-constituent in a syllable. In other words, the segment sequence is rather simple in terms of phonotactics. The simplicity of SX segment sequences can be briefly summarized as follows:

(a) Onset clusters are not allowed;
(b) Complex codas are not allowed;
(c) The minimal sub-constituent has only one terminal segment (except [ɑɒ]) in the nucleus.

4.6.2 Distribution of the Initials and the Finals

There are constraints governing the segment sequences between the onset consonant and the prenuclear glide or the nucleus vowel, although they are not in the same sub-constituent. As was discussed in §4.3.6, the syntactic structure in (37) just captures the sub-constituent relations of the SX syllable structure. In X-bar structure, the onset is the specifier of the syllable (N^{max}), the prenuclear glide is the specifier of the Final (N$''$) and the coda is the complement of the rhyme (N$'$), as shown in (37). The sequences of different segments between constituents are based on certain constraints which have been discussed in previous chapters. The constraints for the coda segments are very simple, for there are only two segments in the coda: /ŋ/ and /ʔ/ in SX, both contrastive with each other as a coda, which is stipulated by the CODA-COND (see (16), ch.2). All the constraints and rules concerning the phonotactics in SX will be presented again in this subsection.

As was discussed previously in this chapter, syllables in SX are typically parsed into two parts: Initial (onset consonant) and Final (all that is left after the onset). As a sub-constituent, the Final always appears as one integral unit in the forms of VC, GV and GVC, as was discussed in chapter 2. There are no segment sequence constraints involved between V and C (except for the constraint that any [−tense] nucleus vowel has to be followed by a coda), or between G and V, or G and V and C, each combination functioning as a sort of closed element, which means their number is limited so that an increase of the number of lexical syllables is only made possible by combinations of Initials, Finals and tones. As was discussed in chapter 2, there are 29 consonants, 48 finals (in five different types) and eight tones. Roughly speaking, there should be logically 620 syllables (excluding the shaded cells for non-existing syllables and '?' cells for undecided status, as shown in the tables of (54), (64), (70) and (73) below, and disregarding the different tones) and when different tones are added, the number of syllables could be about three or four times bigger (considering four tones are in complementary distribution with the other four in obstruent-initial syllables, and some undecided possibilities and accidental gaps. However, no stock has been taken of the existing SX syllables, most of which have more than one lexical meaning for each, symbolized by different Chinese characters.

As for the phonotactic segment sequences, disregarding tones, the main issue involved is the combination of onset consonants with Finals. There are 29 onset consonants, including [p p^h b m f v t t^h d n l ts ts^h dz s z tɕ $tɕ^h$ dʑ ȵ ɕ ʑ k k^h g ŋ h ɦ ʔ], which were discussed in chapter 2 (see the table of 29 consonants in (35), ch.2). There are 48 finals, which can be presented in the following table:

(53) 48 Finals in SX:

C̩	V(V)	Ṽ	GV(V)		GṼ		VC		GVC	
m̩	ɿ	ẽ	ja	wa	jɛ̃	wɛ̃	ɪŋ	ɪʔ	jaŋ	jaʔ
n̩	i	ɛ̃	je	we	ɥɵ̃	wɵ̃	əŋ	əʔ	joŋ	joʔ
ŋ̍	y	ɵ̃	jo	wo	jẽ		aŋ	aʔ	waŋ	waʔ
l̩	e		jɤ				oŋ	oʔ	woŋ	woʔ
	a		jɑɒ				ɒŋ	ɛʔ	wɒŋ	
	ɤ									
	o									
	u									
	ɑɒ									

In (53), 'C̩' refers to syllabic consonants; VV is grouped together with V, for this VV only acts as a long vowel; Ṽ is separate from V, for underlyingly Ṽ is /VN/. The table in (53) displays the full picture of 48 Finals in SX, divided into seven groups. The construction of syllables in SX is simply a combination between a Final and an onset consonant. In this subsection, I will present my data-based analysis of the construction of such combinations (syllables), focusing on the distribution of the 29 consonants, when combined with the different Finals.

Since the phonotactic constraints or rules of segment sequences have been largely discussed in chapters 2 and 3 and previous sections in this chapter, I will present my analysis through the tables in (54), (64), (70) and (73), in which, 'C' (onset consonant) is listed down and Finals across; in the cells of all these four tables, the blank cells stand for existing syllables; the shaded cells for non-existing syllables; '–' for accidental gap; '?' for undecided status. Although the matrixes were largely based on published data and completed through a careful consultation with native speakers and agree with my own intuition, possibly still some uncommonly used syllables exist where I may have marked '–' or even '?'. Consider the cross-tabulation in (54):

(54) Cross-tabulation of C against V/VV/Ṽ

	ɿ	y	i	e	a	ɤ	o	u	ɑɒ	ẽ	ɛ̃	ə̃
p												
pʰ												
b												
m												
f									–			?
v									–			?
ts												
tsʰ												
dz												
s												
z												
n										?		
l										?		
t												
tʰ												
d												
tɕ												
tɕʰ												
dʑ												
ɕ												
ʑ												
ȵ												
k												?
kʰ												?
g												?
ŋ												?
h												?
(ɦ)												?
(ʔ)												?

In the cross-tabulation in (54), V, VV and Ṽ are put together, because the consonants have a very similar distribution when preceding the same vowels, whether oral or nasalized. The glottal fricative [ɦ] and stop [ʔ] are put in brackets because there is no constraint on their distribution when they occur in vowel-initial syllables. According to the table in (54), the segment distribution in the syllable combinations between the onset consonants and the Final V, VV or Ṽ can be formalized as follows:

(i) [ʅ] only occurs after [ts], [tsʰ], [dz], [s], and [z]. In fact, [ʅ] is an allophone of /i/, which is formulated in the rule in (55):

(55) $$/i/ \rightarrow [ʅ] \;/\; \begin{bmatrix}+\text{cons}\\ +\text{apical}\end{bmatrix} \text{—}$$

Rule (55) says that /i/ becomes [ʅ] when preceded by a [+cons, +apical] consonant, which is made possible by sharing the feature of [+apical]. There is also a constraint which stipulates the distribution of [ʅ]:

(56) AGREECV[apical]
An apical consonant must agree with the following high front vowel in value for the status of apical.

The details of the distribution of [ʅ] were discussed in chapter 2 and 3 (see §2.3.3.1; §3.3.1).

(ii) [y] cannot occur after a [+ant, –lateral] or [+dors] consonant, which can be stipulated by the constraint in (57):

(57) *{+ant, –lateral/+dors}-[y]

(iii) [ts], [tsʰ], [dz], [s] and [z] never occur before [i] because of the rule in (55). There is also a constraint to block [k], [kʰ], [g] and [ŋ] from occurring before [i], as shown in (58):

(58) *[+dors]-[–back, +high]
Dorsal consonants cannot be followed by a front high (semi-) vowel.

(iv) [tɕ], [tɕʰ], [dʑ], [ɕ], and [ʑ] only occur before [i] or [y] just like [ȵ] (which is treated as an allophone of /n/), formulated by a palatalization rule:

(59) $$\begin{bmatrix}+\text{nas}\\ +\text{cor}\end{bmatrix} \rightarrow [ȵ] \;/\; \text{—} \begin{bmatrix}-\text{back}\\ +\text{high}\end{bmatrix}$$

There is also a constraint to formalize the distribution of the alveolo-palatal consonants and front high vowels as follows:

(60) AGREECV[+H, −B]
A [+high, −back] consonant must agree with the following vowel in value for the features of [+high] and [−back].

(v) No consonant can occur before all the vowels in the table (54).
(vi) [ŋ] cannot occur before [i] or [u], which suggests the constraint in (61):

(61) *[ŋ][+high]

(vii) [h] cannot occur before any high front vowel, which suggests the constraint in (62):

(62) *[h][+high, −back]

(viii) [ə̃] does not occur after [+back] consonants, which suggests the constraint in (63):

(63) ?{+back, +con}-[ə̃][40]

From the cross-tabulation in (54) and the formalization above, [tɕ], [tɕʰ], [dʑ], [ɕ], [ʑ], and [ɲ] (which are alveolo-palatal affricates, fricatives and nasal) are completely in complementary distribution with [ts], [tsʰ], [dz], [s], [z], and [n] (which are dental affricates and fricatives and the alveolar nasal). The alveolo-palatal consonants can only occur before [i] (or [j]) while the dental affricates, fricatives and alveolar nasal cannot occur before [i]. Are the alveolo-palatal affricates, fricatives and nasal the allophones of dental affricates, fricatives and alveolar nasal, respectively? It may be too early to draw any conclusion, even though there is the fact that [ɲ] is an allophonic variant of /n/ (see (35), ch.2). I will answer this question later in this subsection. Now let us examine the distribution of the consonants before other Finals such as VC, as shown in (64):

[40] Two kinds of constraints are stipulated: one marked by ?, which signals undecided status, indicating no existing word was available though systematically possible, and the other marked by *, which means it is systematically impossible.

(64) Cross-tabulation of C against VC

	ɪŋ	əŋ	aŋ	ɒŋ	oŋ	ɪʔ	əʔ	aʔ	ɛʔ	oʔ
p		?					?			
p^h		?					?			
b		?					?			
m		?					?			
f	?	?				–	?			
v	?	?					?			
ts										
ts^h										
dz										
s										
z										
n										
l										
t										
t^h										
d										
tɕ										
$tɕ^h$										
dʑ										
ɕ										
ʑ										
ȵ										
k										
k^h										
g			–							
ŋ		–			–					
h										
(ɦ)										
(ʔ)										

The cross-tabulation in (64) shows the distribution of the 29 consonants with the ten VC combinations. Since the vowels in VC combinations are similar to those in (54), the distribution shown by (54) and that in (64) bear a lot of similarities, and also a few differences. The distribution of the consonants with VC can be formalized below:

(i) Labial consonants (including bilabial and labial-dental ones) do not seem to occur before [əŋ] or [əʔ], which suggests the constraint in (65):

(65) ?{+labial, +cons}-[ə][41]

(ii) Labial-dental fricatives and dental affricates and fricatives do not occur before [ɪŋ], while the latter can occur before [ɪʔ], which can be formalized by the constraint in (66):

(66) *{+ant, +cont}-[ɪŋ]

(iii) Alveolo-palatal affricates and fricatives can only occur before [ɪŋ] and [ɪʔ], which also follows the constraint AGREECV/G[+H, −B] in (60).
(iv) [n] cannot occur before [ɪŋ] or [ɪʔ] due to the rule in (59).
(v) Dorsal consonants cannot occur before [ɪŋ] or [ɪʔ] because of the constraint in (58).
(vi) [h] cannot occur before [ɪŋ] or [ɪʔ], which is also dominated by the constraint of *[h][+high, −back] in (62).

Table (64) and the formalization of (ii) and (iii) show that [tɕ], [tɕʰ], [dʑ], [ɕ], and [ʑ] are not actually in complementary distribution with [ts], [tsʰ], [dz], [s], and [z] because all can occur before [ɪʔ]. Consider the following examples:

(67)	[tsɪʔ^5]	'fold'	[tɕɪʔ^5]	'urgent'
	[tsʰɪʔ^5]	'exit'	[tɕʰɪʔ^5]	'eat'
	[dzɪʔ^3]	'nephew'	[dʑɪʔ^3]	'fetch (water)'
	[sɪʔ^5]	'brush'	[ɕɪʔ^5]	'snow'
	[zɪʔ^3]	'enter'	[ʑɪʔ^3]	'medicine'

The examples in (67) show that dental affricates and fricatives contrast with alveolo-palatal affricates and fricatives when preceding [ɪʔ] and they make up perfect minimal pairs. Thus, neither is the allophone of the other. However, dorsal stops are indeed in complementary distribution with alveolo-palatal affricates, as shown in the tables in (54), (64), (70) and (73).

---

[41] [vɪʔ] or [vəʔ] can be phonetically realized in urban SX and some rural SX dialects, respectively. My analysis is mainly based on the data of urban SX.

In fact, there is an argument over the phonological properties of alveolo-palatal affricates and fricatives in Mandarin (Chao 1934, 1968; Yip 1996; Duanmu 1999). Chao (1968: 21) claims that "historically the (alveolo-)palatals[42] have come from the dentals and the velars, as reflected in many present-day dialects. The 'feeling of the native' seems to favour the velar". One of the strong reasons for the claim that the alveolo-palatals are likely to be the allophones of the velars is onomatopoeia in Mandarin (Chao 1934, 1968; Yip 1996, among others). Yip (1996: 771) explains with the examples in (68) that the data (a) show the general pattern, and the (b) data show the palatalization. The underlying form shows up in the third syllable, and the palatalizing environment /i/ is in the first syllable:

(68) CV → Ci li CV lV

a.	p^hi li p^ha la	'noise of fire crackers'
	ti li ta la	'sound of rain drop'
	t^hi li t^hu lu	'slurping'
b.	tɕi li k(^{w})a la	'chattering noise'
	$tɕ^h$i li k^h(^{w})a	'noise of falling objects'
	ɕi li xu lu	'eating fast'

The data in (68b) suggest that [tɕ $tɕ^h$ ɕ] are allophones of /k k^h x/ in Mandarin when preceding the high front vowel. It is also true in SX that [tɕ $tɕ^h$ ɕ] and [k k^h h] are in complementary distribution, as shown in (54) and (64). However, there is another example of onomatopoeia in both Mandarin and SX where the alveolo-palatals do not occur complementarily with the velars when followed by [ɪŋ] as in (69):

(69) CV → Cɪŋ lɪŋ CVŋ lVŋ

a.	pɪŋ lɪŋ paŋ laŋ	'noise of knocking objects with each other'
b.	tɪŋ lɪŋ kwaŋ laŋ	'noise of falling metal objects' (Mandarin)
	tɪŋ lɪŋ kwɒŋ lɒŋ	'noise of falling metal objects' (SX)
c.	*tɕɪŋ lɪŋ kwaŋ laŋ	(Mandarin)
	*tɕɪŋ lɪŋ kwɒŋ lɒŋ	(SX)

The data in (69) show that in the pattern of Cɪŋ lɪŋ CVŋ lVŋ for onomatopoeia, [t], instead of [tɕ], occurs for the onset of the first syllable when [k] is the onset of the third syllable in (69b), although [tɕɪŋ] is a well-formed

[42] Chao (1968) identifies [tɕ $tɕ^h$ dʑ ɕ ʑ] as palatals; I classify them as alveolo-palatals according to IPA (revised to 1993) (see Ladefoged & Maddieson (1996: 426).

syllable in both Mandarin and SX. Also according to Baxter (1992), both [tɕ] and [k] could precede the prenuclear glide [j] which is also [+high] and [−back] in Old Chinese (11th to 7th centuries B.C.). However, it is not necessary that of alveolo-palatal affricates and velar stops in SX, one of them has to be an allophone of the other, even though they are in complementary distribution. I argue that both the alveolo-palatals and the velars are phonemic consonants and their complementary distribution is decided by the phonotactic constraints in SX, just like /h/ and /ŋ/ in English. In addition to the analysis I presented above, there are two more reasons why both the alveolo-palatal affricates and velar stops are underlying forms:

(i) There is much phonetic difference between these groups of sounds, not only in place of articulation but also in manner of articulation.

(ii) If alveolo-palatal affricates are allophones of the velar stops, what is the phonemic status of [ɕ] and [ʑ], which have the same distribution as [tɕ], [tɕ^h] and [dʑ]?

The distribution of the 29 consonants when combined with GV and when with GVC should be the same, differing only in the case of the two glides: [j] and [w], which can be represented in (70):

(70) Cross-tabulation of C against GV/GṼ

	ja	je	jo	jɤ	jɑɒ	jɛ̃	jẽ	ɥɵ̃	wa	we	wo	wɛ̃	wɵ̃
p	?	?	?	?		?							
p^h	?	?	?	?		?							
b	?	?	?	?		?							
m	?	?	?	?		?							
f		?	?	?		?							
v	–	?	?	?	–	?							
ts													
ts^h													
dz													
s													
z													
n									?	?	?	?	
l	?	?	?			?		?	?	?	?	?	
t		?	?			?		?	?	?	?	?	
t^h		?	?			?		?	?	?	?	?	
d	–	?	?			?		?	?	?	?	?	
tɕ		?				?							
$tɕ^h$		?				?							
dʑ		?				?							
ɕ		?	–			?							
ʑ		?	–			?							
ɲ		–											
k													
k^h													
g													
ŋ													
h													
(ɦ)													
(ʔ)													

The cross-tabulation of C against GV and GṼ in (70) shows quite a different picture, with only about 31% of combinations which are attested and possible syllables,[43] including blank cells and '–'. About 49% of the combinations are systematic gaps and about 20% of the combinations are undecided, though systematically possible. I assume that the main reason for such a small proportion of possible syllables is the general tendency of complementary distribution of the consonants combined with the two pre-nuclear glides, [j] and [w]. As was mentioned in chapter 2, [je] and [jo] are mainly applied to words borrowed from Mandarin. The general distribution of the 29 consonants with the 14 GV and GṼ combinations in (70) can be formalized as follows:

(i) Systematically, [+labial] consonants cannot precede [wV] combinations and [+back] consonants cannot precede [jV] combinations. The latter still follows the constraint in (58) and the former involves an OCP constraint specified as in (71):

(71) OCP(labial)
Two adjacent prevocalic segments identical in [+labial] are not acceptable within one syllable (CV is irrelevant).

(ii) No GV combination can occur after [ts ts^h dz s z], which can be formalized by a constraint like in (72):

(72) *[+apical]-GV

(iii) [n] cannot occur before [jV], due to the rule in (59).
(iv) The alveolo-palatal affricates and fricatives cannot occur before [wV], which also follows the constraint AGREECV/G[+H/−B] in (60).
(v) [ŋ] cannot occur before any GV combination, due to the constraint *[ŋ][+high] in (61).
(vi) [h] cannot occur before a [jV], because of the constraint *[h][+high, −back].

Now let us see if there is any difference in the distribution of the consonants with GVC combinations, as shown in (73):

[43] The calculation of possible syllables does not consider the syllables with different tones.

(73) Cross-tabulation of C against GVC

	jaŋ	jaʔ	joŋ	joʔ	waŋ	wɒŋ	woŋ	waʔ	woʔ
p									
p^h									
b									
m									
f									
v									
ts									
ts^h									
dz									
s									
z									
n									
l			?	?					
t									
t^h									
d									
tɕ									
$tɕ^h$									
dʑ									
ɕ									
ʑ									
ȵ									
k									
k^h									
g									
ŋ									
h									
(ɦ)									
(ʔ)									

The cross-tabulation in (73) shows an even smaller proportion of the possible syllables combining the 29 consonants and the nine GVC Finals in SX, with only 24.5% of all the combinations being existing syllables. The distribution shows that all GVC combinations seem only to yield possible syllables when preceded by alveolo-palatal, dorsal and glottal consonants, and that all [ant] consonants are irrelevant to the syllable formation with GVC, except the lateral [l], which can occur before [jaŋ] and [jaʔ], as shown in (74):

(74) [ljaŋ^{22}] 'two'
[ljaʔ^{3}] 'omit'

The combinations between the onset consonants and GVC can be formalized as follows:

(i) No [+ant, –lateral] consonant can combine with GVC, which suggests a constraint in (75):

(75) *{+ant, –lateral}GVC

(ii) Alveolo-palatal affricates and fricatives cannot occur before [wVG], which still follows the constraint of AGREECV/G[+H, −B] in (60).
(iii) Velar stops cannot occur before [jGC], due to the constraint *[+dors]-[–back, +high] in (58).
(iv) [ŋ] cannot occur before any GVC combination, due to the constraint *[ŋ][+high] in (61).
(v) [h] cannot occur before any [jVC] combination due to the constraint *[h][+high, −back].

In brief summary, all the four tables with the 29 consonants and the 48 Finals in SX present an overall distribution of all these segments in SX, from which we have found that there is still a lot of room for more syllables (even more so when combined with different tones, though the Finals are fixed and limited in number. The distribution of all the segments in SX as shown in (54), (64), (70) and (73) also proves that the prenuclear glides are not in the onset, and that GV and GVC occur as a subsyllabic constituent in the formation of syllables in SX. The distribution of the segment sequences as shown in the above four tables involve different phonotactic constraints which were discussed above. Some consonants are in complementary distribution with others; some can occur before more different vowels or Finals than others. The four cross-tabulations also show the phenomenon of segment distribution, viz. different onset consonants are decided by the following glides (if there is one) or by the nucleus vowel (when there is no medial glide) while the allophonic vowels in the Finals are decided by either the preceding onset consonants or the following coda. However, the phonotactic constraints of SX, as a tonal language, also involve other constrains in terms of the tonal system. This will be the topic of the next chapter.

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 6^{th} and 10^{th} 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

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, [bɛ̀lɛ̀] '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

a. Lo

bɛ lɛ

b. * Lo Lo

bɛ lɛ

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.

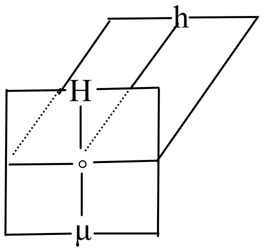
5.4 The Geometry of Tone

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)



Register tier

Tonal tier

Tonal root node (TRN) tier

Tonal-bearing unit tier

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)

h	h	l	l	Register tier
H	L	H	L	Tonal tier
∘	∘	∘	∘	Tonal root node tier
μ	μ	μ	μ	TBU tier
Hi	Mid_2	Mid_1	Lo	

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)

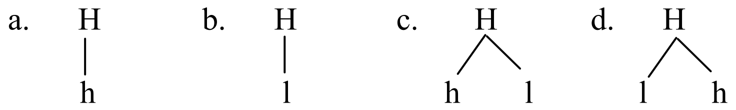
σ		
H	Tonal Node	Register
l h		Tone

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]

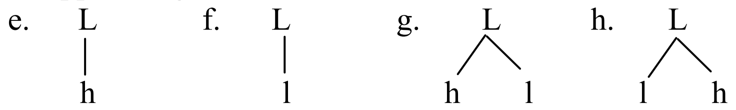
[11] It is still a matter of controversy whether concave and convex tones exist underlyingly in the languages which are claimed to have them (see Yip 2002).

[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



[−Upper] Register

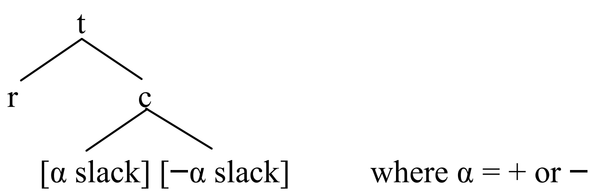


Yip’s [±Upper] approach to Register allows eight distinctive tones, which quite precisely captures the tonal system of SX, whose tones have a clear register division, four in the [+Upper] register and the other four in the [−Upper] register, strictly divided. On Yip’s assumption that register features 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

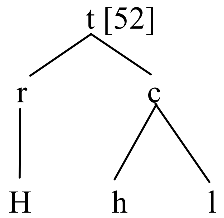
Following Yip (1980, 1989) and Halle and Stevens (1971), let us assume that tone allows a register division ([+Upper] and [−Upper]) and is specified by either [+stiff] (H) or [−stiff] (L). Bao (1999) proposes a geometry of tone in which a tonal root node (t) dominates both a register (r) and a contour (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)



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):

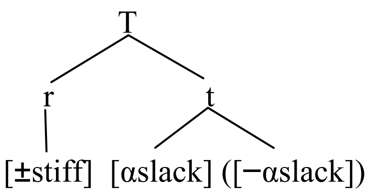
(14)



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.

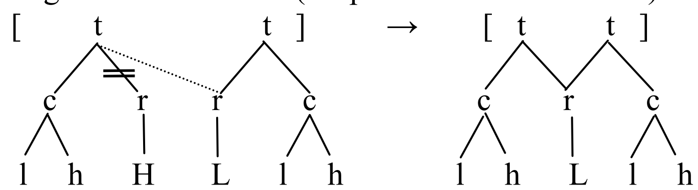
5.4.4 Register feature spreading

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):

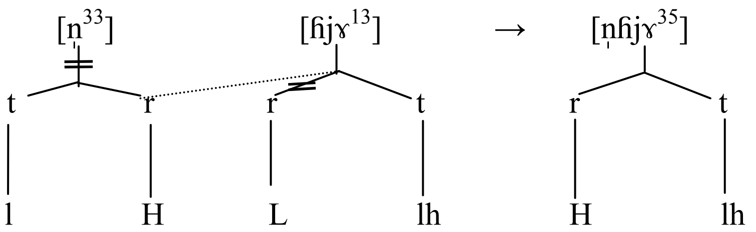


In (17), the register feature [L] of the second tone spreads regressively to the tonal node and the register feature [H] of the first tone is delinked, changing from [H] to [L], while the contour features remain unchanged, and thus the two tones have the same register feature [L]. Rule (17) says that 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 [ŋ̍jɤ^{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).

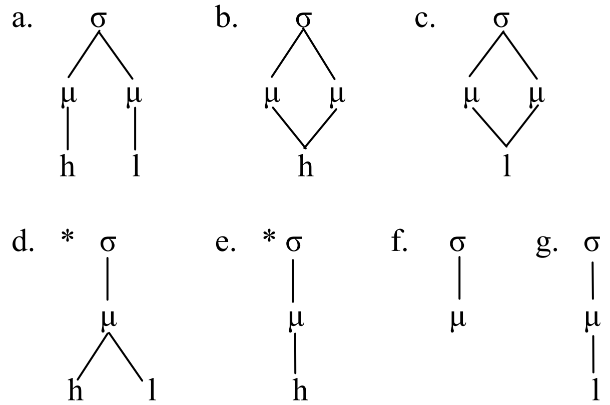
5.5 The TBU in SX

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) Stressed:

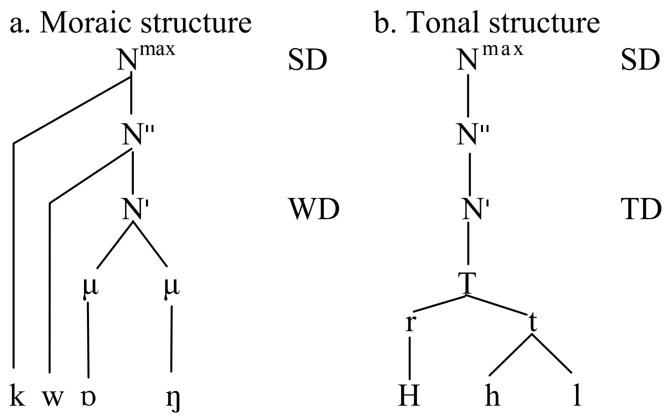
Unstressed:



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).

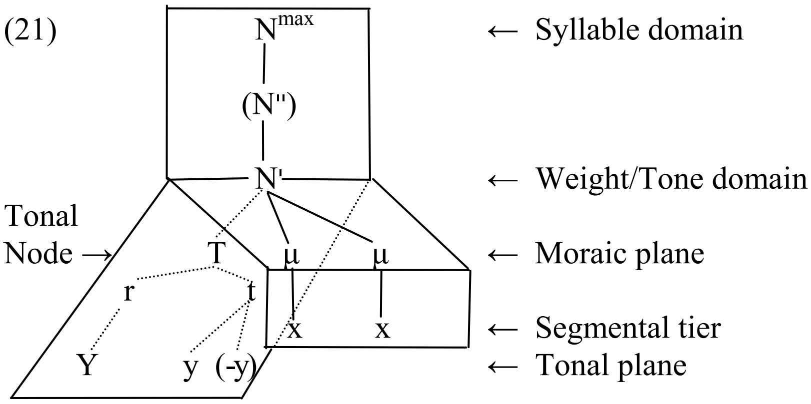
As was discussed in chapter 4, the nucleus and coda are moraic in SX, so that the weight domain is N' in the X-bar structure, as was shown in ((36), ch.4). Thus, N' is also the tone domain in tonal structure. The moraic structure and the tonal structure are two aspects of one and the same syllable structure.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

(20)



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'
[$t^h ɪʔ^{5}$]	'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'

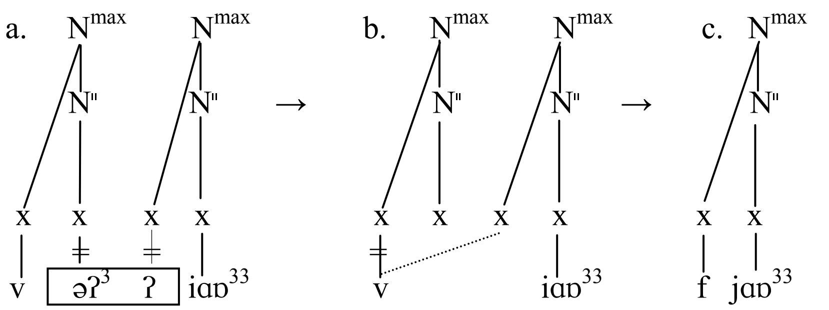
However, the phrase [vəʔ^{3}ʔjɑɒ^{33}] 'don't want (to)' in (28a) always appears in a merged syllable [fjɑɒ^{33}], in which the negator [vəʔ^{3}] 'not', as a clitic, merges phonetically and phonologically into the host syllable [ʔjɑɒ^{33}] 'want', resulting in a new syllable [fjɑɒ^{33}] 'don't want (to)'. 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		ʔ	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. $*\mathrm{C}\ \mathrm{V}^{\mathrm{L}}$ [+stiff] b. $*\mathrm{C}\ \mathrm{V}^{\mathrm{H}}$ [+slack] c. $*\mathrm{o}\ \mathrm{V}^{\mathrm{H/L}}$ [+stiff/+slack] d. $*\ \mathrm{C}\ \ \mathrm{V}^{\mathrm{H/L}}$ [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$^{\mathrm{H}}$/ (underlyingly a high-register tone has no onset consonant) has the surface form in (42):

(42)

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

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.

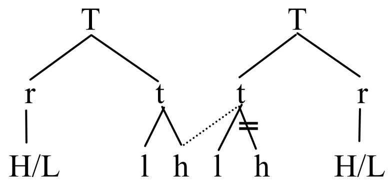
[23] All the examples of tone sandhi in SX in this chapter are based on Yang & Yang (2000), with some alternations and additions which were made after consultation of SX native speakers.

(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])

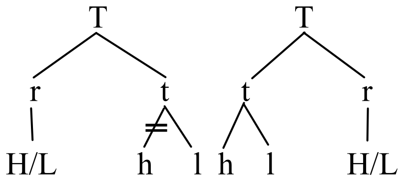
[ku^{52}ȵjaŋ^{52}]	→	[ku^{33}ȵjaŋ^{52}]	‘girl’
[ɕjaŋ^{52}ʔjẽ^{52}]	→	[ɕjaŋ^{33}ʔjẽ^{52}]	‘cigarettes’

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])

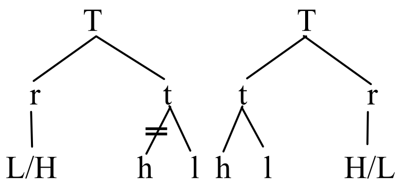
[dɑɒ^{31}hwo^{52}]	→	[dɑɒ^{22}hwo^{52}]	'peach blossom'
[dɒŋ^{31}tɕɪŋ^{52}]	→	[dɒŋ^{22}tɕɪŋ^{52}]	'glucide'

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):

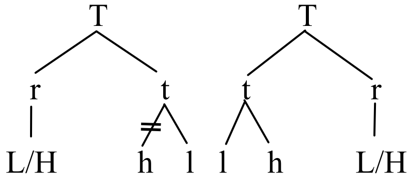
(56) a. [31.13] → [22.13] ([L.hl][L.lh] → [L.l][L.lh])

[nuɵ̃^{31}ȵy^{13}]	→	[nuɵ̃^{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):

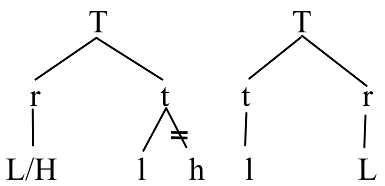
(57)


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.).

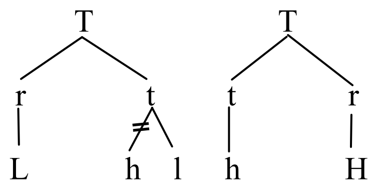
However, contour simplification in SX does not necessarily result from an OCP violation, because the [h] delinking in (57) does not change the [l][l] identical feature adjacency. Sometimes the application of contour simplification will in fact result in what is called a partial OCP violation (Yip 1989;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

(59) [13.22] → [22.22] ([L.lh][L.l] → [L.l][L.l])
[lɑɒ^{13}do^{22}] → [lɑɒ^{22}do^{22}] 'eldest brother/sister'

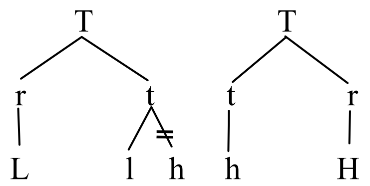


The example in (59) shows that the delinking of the [h] feature has caused a violation of partial OCP″ (58). This suggests that the partial OCP″ constraint does not account for the data regarding tone sandhi in SX. However, there are many other examples of contour simplification in SX, all of which are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

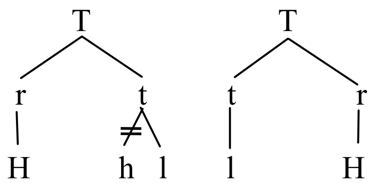
(60) [31][5] → [22][5] ([L.hl][H.h] → [L.l][H.h])
[dʑjaŋ^{31}pɪʔ^{5}] → [dʑjaŋ^{22}pɪʔ^{5}] 'wall'
[ɦɥẽ^{31}tsoʔ^{5}] → [ɦɥẽ^{22}tsoʔ^{5}] 'round table'



(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])

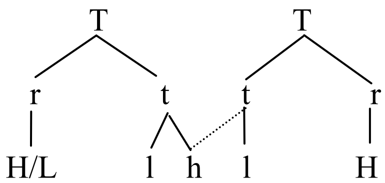
[doŋ^{13}tɕi^{33}]	→	[doŋ^{13}tɕi^{52}]	‘intention’
[lɛ̃^{13}hẽ^{33}]	→	[lɛ̃^{13}hẽ^{52}]	‘lazy man’

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):

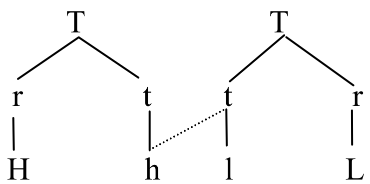
(67)



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][_]#

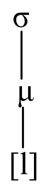
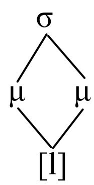
The rule in (69) says that a low level tone will become a falling contour when it is at a right-hand lexical or phrasal boundary and is preceded by a rising contour or a high level tone (i.e. preceded by a high pitch). The contour formation rule in (69) excludes a situation in which [l] is preceded by 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.  b. 

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* $\gg$ DEFAULT-T(R/L) $\gg$ IDENT-TT(R) $\gg$ *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. [$k^h\tilde{ẽ}^{33}$]	'look'	→	[$k^h\tilde{ẽ}^{33}k^h\tilde{ẽ}^{33}$]	'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 [k^{h}ẽ^{33}]	OCP (contour)	IDENT -BR
a. ☞ [k^{h}ẽ^{33}k^{h}ẽ^{33}]		
b. [k^{h}ẽ^{33}k^{h}ẽ^{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])

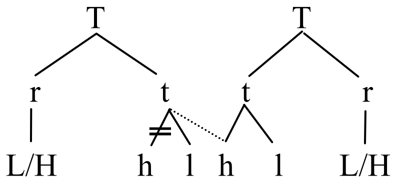
[ɦja^{31}ɦja^{31}]	→	[ɦja^{13}ɦja^{31}]	‘grandfather’
[ȵjaŋ^{31}ȵjaŋ^{31}]	→	[ȵjaŋ^{13}ȵjaŋ^{31}]	‘grandmother’
[bo^{31}bo^{31}]	→	[bo^{13}bo^{31}]	‘grandmother’

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])

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

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

[dʑjɤ^{13}dʑjɤ^{13}]	→	[dʑjɤ^{13}dʑjɤ^{31}]	'uncle'
[di^{13}di^{13}]	→	[di^{13}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.

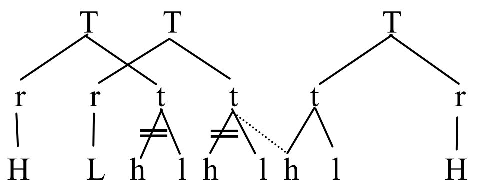
As was discussed previously, IDENT-TT(R) in (55) serves the same aim as IDENT-BR[C] in (79), viz. it preserves two contours as tone type in sandhi forms for the underlying identical contours, which means that two adjacent identical rising contours have the same sandhi rule whether they are in the form 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

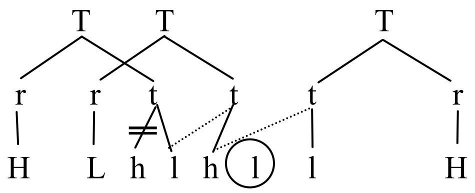
When tone sandhi occurs in disyllabic lexical compounds or phrases, only one tone in the disyllabic unit undergoes some change, either by contour dissimilation, contour simplification or contour formation. These changes are phonologically realized by spreading or delinking the [h] feature and they are 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]



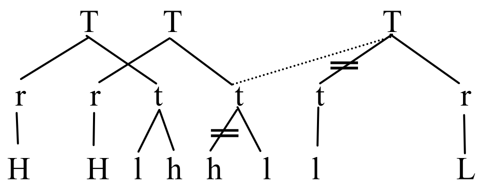
In (82), first, the leftmost contour delinks the [h] feature and becomes a low level tone; secondly, the middle tone delinks its original [h] feature when receiving the [h] feature from the following tone, resulting in the sandhi form: [l.lh.hl].[27]

(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].

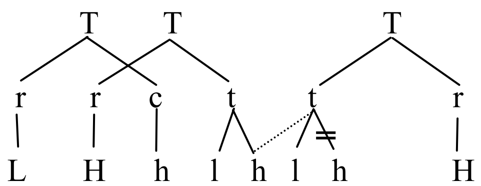
(84) [fu^{35}ts^{h}o^{52}lu^{22}] → [fu^{35}ts^{h}o^{33}lu^{31}] ‘railway’
[H.lh][H.hl][L.l] → [H.lh][H.l][L.hl]



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].

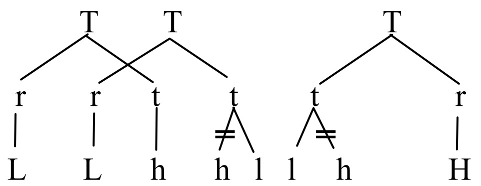
[27] Since the register feature is irrelevant in tone sandhi in SX, it is omitted in the sandhi form, in which a dot ‘.’ indicates a syllable boundary.

(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]



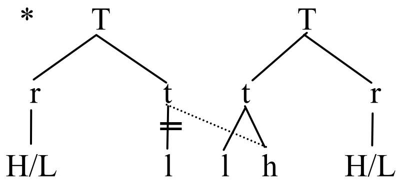
In (86), only contour simplification occurs in both the middle tone and the rightmost tone, resulting in the sandhi form [h.l.l].

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:

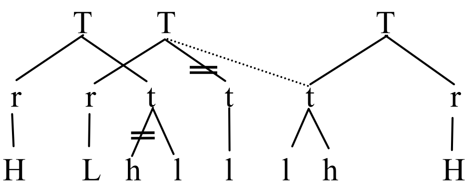
(87) Principles of tone sandhi in SX:
   a. Tone sandhi in SX is phonologically realized by tone feature spreading or/and delinking; no new features are ever inserted.
   b. Feature spreading can be progressive or regressive.
   c. The sequence of application of feature spreading or delinking in trisyllabic 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):

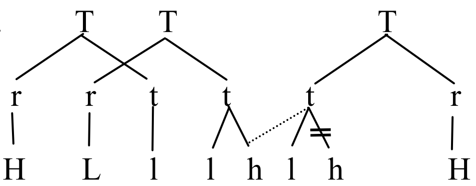
(88)


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}sɿ^{35}] 'toilet water' changing into [huo^{33}lu^{13}sɿ^{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^hɑɒ.k^hɑɒ]	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}nə̃^{31})he^{35}]	→	[(tsoŋ^{33}nə̃^{13})he^{52}]
	middle south sea		[(l.lh).hl]
	'Middle South Sea'		
b.	[fi^{52}(nə̃^{31}he^{35})]	→	[fi^{33}(nə̃^{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}dzɤ^{33}]	→	[po^{33}dzɤ^{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 ([dɒŋ³¹lɒŋ³¹])	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] ([ɦia³¹ɦia³¹] '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] ([dɒŋ³¹lɒŋ³¹] '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 [tɕi³⁵tɕi³⁵] 'elder sister' and non-reduplicated [dɑɒ¹³tsʰɑɒ³⁵] 'straw' in tableaux (109) and (110), respectively:

(109) Reduplication

Input: lh.lh ([tɕi³⁵tɕi³⁵])	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 [kʰẽ³³kʰẽ³³] 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 [tɕi³³ȵjɛ̃²²] '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. [pɛʔ⁵pɛʔ⁵] 'pat', whose sandhi form is also [pɛʔ⁵pɛʔ⁵], 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 [$k^{h}\tilde{e}^{33}$$\text{ɕ}y^{52}$]. The same holds for [l.lh]. For example, [$ma^{22}mi^{23}$] '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, [$\text{sɪʔ}^{5}ba^{22}$] '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.

6 Conclusion

6.1 Summary

This dissertation has attempted to present a detailed analysis of the phonology of Shaoxing (SX) Chinese, including (i) the surface inventory of consonants and vowels and their distribution, (ii) the underlying vowel system, (iii) the syllable structure, focusing in particular on the status of the prenuclear glide in the syllable, (iv) the phonotactics of SX, in particular the possible Finals and their combination with possible Initials in the syllable, (v) the tonal inventory of SX, (vi) the consonant-tone interactions, and (vii) the tone sandhi rules. The most important conclusion that can be drawn from the analysis I have presented is that SX has a phonological system that works in comparable ways to other languages of the world, in the sense that much of its behaviour follows common linguistic principles, albeit with specific variations on universal themes. I will briefly report on general conclusions with respect to specific phonological phenomena in SX in the following section.

6.2 Main Conclusions

6.2.1 The vowel and consonant inventory

Languages differ considerably regarding their inventories of consonants and vowels, and also with regard to the phonological grammar that specifies how these sounds can be combined to form words and utterances (Kay 1989). What makes SX different from other languages is its rather large inventory of consonants and vowels. As was discussed in chapter 2, SX has 29 initial consonants and 14 surface vowels, which form 48 Final combinations. Among the Chinese languages, the Wu language family (including SX) has more initial consonants than others. This is related to the fact that it still retains the historical voiced obstruents, while all the other six Chinese language families have (more or less) lost the voiced obstruents, resulting in inventories of around 20 initial consonants or even fewer (e.g. Hakka has 17 initial consonants and Cantonese has 16, according to Campbell 2003). There are three remarkable characteristics in the SX consonant system. First, the language displays a distinction between voiceless aspirated, voiceless unaspirated and voiced obstruents. Secondly, SX has a relatively symmetric system of eight voiced and voiceless fricatives, while there are fewer fricatives in other Chinese language families. Thirdly, the “filler” onset consonants [ʔ] and [ɦ], correlated with the register division, are also features of SX that call out for analysis.

Vowels in SX form a more complicated system in SX than in other Chinese languages. Vowels display a large number of surface variants, according to phonetic environment. I assume in chapter 3 that of the 14 surface vowels in SX, there are only six underlying segments (/i u e ɤ o a/), i.e. quite close to a basic five-vowel system (/i u e o a /) which is found more often in the world’s languages. The fact that there is a large difference in the number of surface vowels and underlying vowels makes it necessary to postulate a number of phonetic realisation rules.

The analysis I presented of the underlying vowel system of SX is intended to present a clear picture of the overall distribution of the 14 surface vowels in SX. I have also presented all the phonological rules and constraints concerning the vowel distribution in SX, which shows that the phonological principles are the same, although phonological behaviour may be different from language to language.

A number of claims that were made in the course of the discussion of vowels and consonants are the following:

(i) Affricates are single segments;
(ii) All vowels in open syllables must be [+tense];
(iii) Glide-Vowel combinations cannot share [+high];
(iv) Vowel-Glide combinations are not permitted;
(v) The two parts of a diphthong cannot agree for both [high] and [back].

6.2.2 Syllable structure

Syllable structure has attracted phonologists’ attention for many decades; the syllable allows the formulation of many generalizations both at the segmental level and at higher prosodic levels. There have been many approaches to the internal structure of the syllable. This dissertation proposes a syntactic approach to internal syllable structure, especially to shed light on the status of prenuclear glides in SX. Following Levin (1985) and Chomsky (1995), I have adopted a multiple-Spec X-bar syllable theory, which allows a syllable maximally to have three sub-constituents: N'' (Final), N' (Rhyme) and N^0 (Nucleus). In this X-bar schema, a syllable is maximally parsed into Onset and Final, instead of Onset and Rhyme (as in the classical OR models), and the problematic prenuclear glide is located in the specifier position of N". The multiple-Spec X-bar structure not only solves the problem of the controversial syllabic position of the prenuclear glide, but also accounts for other data in the Chinese languages, including language games, the poetic rhyming system, loanword phonology and the traditional *Fanqie* system.

My multiple-Spec X-bar syllable structure is based on Levin's (1985) X-bar structure, but is also different from hers. In Levin's proposal, every consonant belongs to an independent N", which, however, is not a subconstituent, while in my proposal every binary node is a constituent projection. This follows the general linguistic principle that every constituent is binary (see Radford 1997). With the multiple-Spec X-bar schema in hand, I assume that the following generalizations hold for the syllable structure of all Chinese languages, including SX:

(i) Onset clusters are not allowed.
(ii) Coda clusters are not allowed.
(iii) Onsetless syllables are permitted underlyingly.
(iv) A syllable is maximally parsed into Onset and Final.
(v) Prenuclear glides are in the Spec position of N".
(vi) The rhyme domain is the weight domain.

In future work, I will explore the consequences of the X-bar model for other Chinese languages.

6.2.3 Consonant-tone correlation

Although it is well documented that voiceless initial obstruents induce high tones and voiced initial obstruents induce low tones cross-linguistically, consonant-tone interaction is still a controversial issue, both phonetically and phonologically. The fact that voiceless obstruents and voiced obstruents and high-register tones and low-register tones, respectively, occur together in SX may lead to a chicken-and-egg situation: which determines which? In this dissertation, I have attempted to present an objective analytic description of consonant-tone interaction in SX, with evidence from, for instance, syllable merger in cliticization, phonetic onset insertion, sonorant initials, etc., and reached the conclusion that both voiced initial obstruents and low-register tones occur in underlying representation in SX.

Some of the dialects in the Wu language family, which is claimed to be the only Chinese language family that still retains the voiceless vs. voiced distinction in the obstruent system, are losing the voiced obstruents as well as the low-register tones (see Cao 2002). SX still preserves a clear-cut register division and the historical voiced obstruents. In some Wu dialects (e.g. Jinhua, Lanxi, Longquan, etc.), the original low-register tones are phonetically realized as high-tone pitches, and voiced and voiceless initial obstruents are beginning to appear with high and low register tones, respectively (Cao 2002). This phenomenon may also throw some light on the issue of consonant-tone interaction, and requires further study.

6.2.4 Tone sandhi rules

Tone sandhi is a common phenomenon in all tonal languages; yet its complexity and the sheer variety of sandhi rules have prevented a fully-fledged analysis in contemporary linguistic studies. SX is a typical tone language with eight tones, equally divided into high and low registers. Besides a systematic analysis of tone feature specifications and a discussion of the geometry of tone, I have made an effort to formalize the intricacies of tone sandhi in SX. We found that tone sandhi in SX is realized by tone feature spreading and delinking, and does not involve register features at all. In this dissertation I have presented an overview of disyllabic sandhi forms. Although not much has been done about systematic explicit formulation, either rule-based or constraint-based, of a tone sandhi system for any of the Chinese languages so far, I have presented a metrically-based analysis, claiming that SX is a right-prominent language and that the stress foot is the tone sandhi domain. Careful study and systematic analysis reveals the following generalizations with respect to tone sandhi:

(i) Tone sandhi in SX is phonologically realized by tone feature spreading or/and delinking.
(ii) Feature spreading can be progressive or regressive.
(iii) Feature spreading cannot cause association lines to cross.
(iv) Register features are never affected by tone sandhi in SX.
(v) In SX, identical contours are never allowed in a foot domain.
(vi) A *ru* tone never changes in SX tone sandhi.
(vii) No falling contour can occur on the left-hand syllable in SX.

Based on these generalizations, I have presented the following constraint ranking that regulates all the 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.

This constraint ranking precisely captures the tonal behaviour in sandhi and accounts for all the disyllabic sandhi rules in SX, as was discussed in Chapter 5. I hope that similar constraint rankings may be appropriate for the formalization of tone sandhi phenomena in other Chinese languages, which, however, requires more investigation.

Tone sandhi is still a linguistically mysterious issue: why are tones allowed to change so much, even in a tone language where lexical meaning is partly determined by pitch? I have made a systematic and hopefully exhaustive analysis of disyllabic tone sandhi in SX. Yet the sandhi rules of trisyllabic and polysyllabic tone sandhi remain unclear, in addition to the fact that the way in which tone sandhi operates differs considerably across linguistic areas (Gandour 1978). However, as long as there are rules, there is a way to formulate them in general. This is certainly a task for further studies.

6.3 Further Study

Some issues I have discussed in my dissertation still require further study, as I have mentioned above. Besides, there are some more issues which are very interesting in my view but are not discussed in this dissertation. For example, with respect to the consonant-tone correlation, the question can be raised whether the nucleus vowel can assign a register feature. How would this bear on the consonant-tone correlation? This question requires more phonetic and phonological investigation. Secondly, how do we account for (Chinese) languages in which there are counterexamples against the consonant-tone correlation? For example, [so^{221}] 'sit', [kɤ^{221}] 'dyke', [tɕye^{221}] 'column' in Qingyuan[1] (Cao 2002). These also require both diachronic and synchronic studies.

There is a Chinese saying "to cast a brick to attract jade". I hope my work on the phonology of Shaoxing is a brick which may cause jade to appear in future work.

[1] Qingyuan is a southern Wu dialect which also has eight tones, among which [33] and [221] are a pair of even tones in high and low registers, respectively, according to Cao (2002: 100).

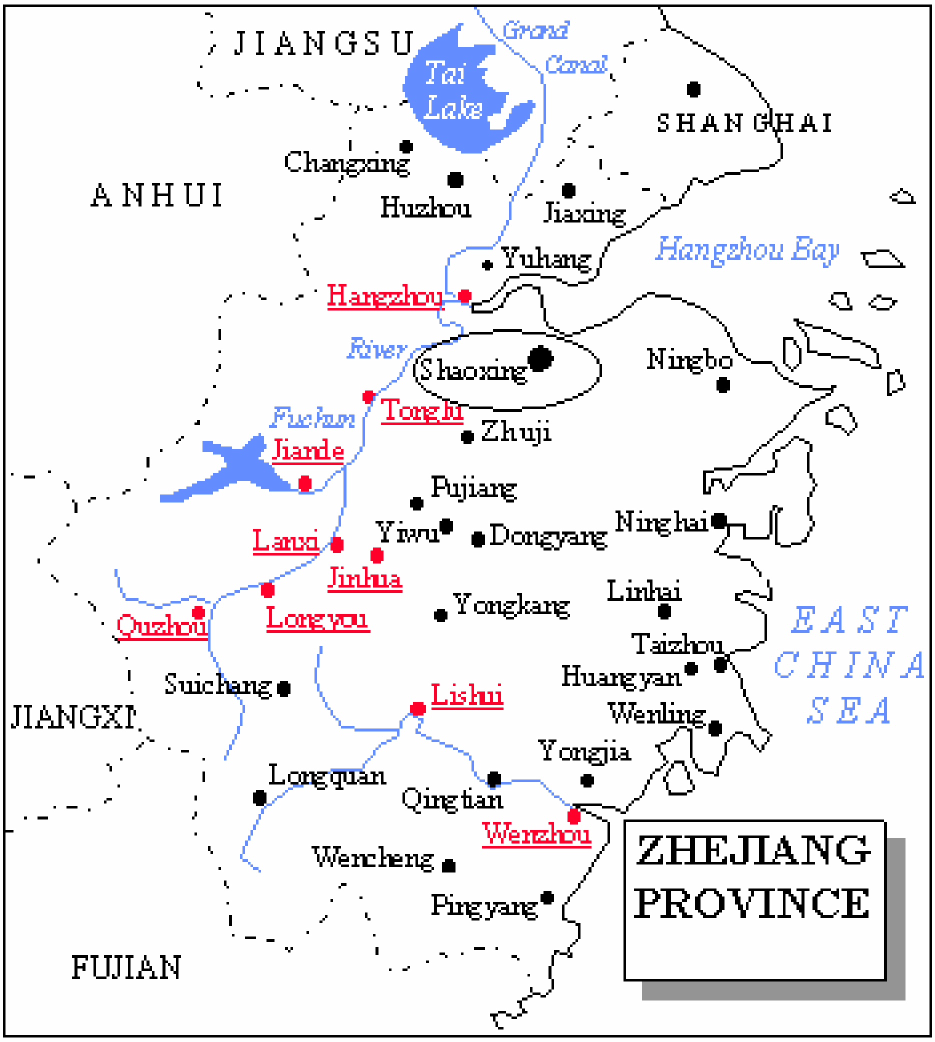
Appendix I

Distribution of Chinese Linguistic Groups

Source: http://www.glossika.com/en/dict/dialectmap.php

Appendix II

Map of Zhejiang Province



Source: http://www.magma.ca/~mtooker/cities/zhejiang.htm

Appendix III

Location of Shaoxing

Source: http://www.multimap.com/wi/12989.htm

References

Anderson, Stephen R. 1978. Tone features. In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey*. New York: Academic Press.

Anderson, John M. & Colin J. Ewen 1987. *Principles of Dependency Phonology*. Cambridge: CUP.

Bagemihl, Bruce 1991. Syllable structure in Bella Coola. *Linguistic Inquiry* 22, 589-646.

Ballard, William Lewis 1969. Phonological History of Wu. PhD dissertation, University of California in Berkeley.

Bao, Zhiming 1990. Fan-Qie languages and reduplication. *Linguistic Inquiry* 21.

----- 1999. *The Structure of Tone*. New York, Oxford: OUP.

----- 2000. Syllabic constituency and sub-syllabic processes. *Journal of East Asian Linguistics* 9, 287-313.

Baw, Howshing 1988. Hangzhou fangyan yu Beifanghua [Hangzhou dialect and northern Mandarin]. In *Wuyu Luncong*, 282-288.

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.

Bhat, D.N.S. 1978. A general study of palatalization. In J.H. Greenberg (ed.) *Universals of Human Language*, 47-92. Stanford: Stanford University Press.

Bickmore, Lee 1995. Tone and stress in Lamba. *Phonology* 12, 307-341.

Blevins, Julie 1995. The syllable in phonological theory. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. 206-244. Cambridge: Blackwell.

Boer, Bart de 2001. *The Origins of Vowel Systems*. Oxford: OUP.

Borowsky, Toni J. 1986. Topics in the Lexical Phonology of English. PhD dissertation, University of Massachusetts at Amherst.

Botma, Bert 2004. Phonological Aspects of Nasality: An Element-Based Dependency Approach. PhD dissertation. Utrecht: LOT.

Bradshaw, Joel 1979. Obstruent harmony and tonogenesis in Jabêm. *Lingua* 49, 189-205.

Bradshaw, Mary M. 1995. Tone on verbs in Suma. In Akinbiyi Akinlabi (ed.) *Theoretical Approaches to African Linguistics*, 255-271. Trenton: Africa World Press, Inc.

Bright, William 1978. Sibilants and naturalness in aboriginal California. *Journal of California Anthropology, Papers in Linguistics* 1, 39-63.

Brooks, M.Z. 1965. On Polish affricates. *Word*, 20, 207-210.

Broselow, Ellen 1995. Skeletal positions and moras. In John Goldsmith (ed.), *The Handbook of Phonological Theory*. 175-205. Cambridge, MA: Blackwell.

Butskhrikidze, M. 2002. Consonant phonotactics of Modern Georgian. Doctoral dissertation, University of Leiden.

Calabrese, Andrea 1999. Sievers' law in Vedic. In Harry van der Hulst & Nancy A. Ritter (ed.), *The Syllable Views and Facts*. 673-751. Berlin: Mouton de Gruyter.

Campbell, James 2003. Glossika Chinese Dialects, www.glossika.com.

Cao Zhiyun 1998. Hanyu fangyan shengdiao yanbian de liangzhong leixing [Two types of tone change in Chinese dialects]. *Yuyan Yanjiu* 1.

----- 2002. Nanbu Wuyu Yuyin Yanjiu [Sound Change in the Southern Wu Dialects]. Beijing: Shangwu Yinshuguan.

----- 2004. Yanyu fangyan zhongde yunwei fendiao xianxiang [The phenomenon of tone split with different coda consonants in Chinese dialects]. *Zhongguo Yuwen* 1, 40-45.

Casali, Roderic F. 1996. Resolving Hiatus. PhD dissertation, University of California in Los Angeles.

Chan, Marjorie K. M. 1985. Fuzhou Phonology: A Non-linear Analysis of Tone and Stress. PhD dissertation, University of Washington.

---- 1991. Contour-tone spreading and tone sandhi in Danyang Chinese. *Phonology* 8.2, 237-259.

----- 1995. An autosegmental analysis of Danyang tone sandhi: some historical and theoretical issues. In E. Zee (ed.), *Studies of the Wu Dialects*. New Asia Academic Bulletin, vol. 11. 145-184. Hong Kong.

----- 1997. Fuzhou glottal stop: floating segment or correlation of close contact?. In J. Wang & N. Smith (ed.), *Studies in Chinese*, 275-289. Berlin: Mouton de Gruyter.

Chang, Laura Mei-chih 1992. A Prosodic Account of Tone, Stress, and Tone Sandhi in Chinese Languages. PhD. Dissertation, University of Hawaii.

Chao Yuen Ren 1928. *Studies in the Modern Wu-Dialects*. Peking: Tsing Hua College Research Institute.

----- 1929. Nanjing inshih [The phonology of Nanjing Chinese]. *Keshyue* 3, 1005-1036.

----- 1930. A system of tone letters. *Le Maître Phonétique* 45, 24-27.

----- 1931. Fanqie yu ba zhong [Eight types of Fanqie languages]. *Bulletin of the Institute of History and Philology*, Academia Sinica 2, 312-354.

----- 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* 4, 363-397.

----- 1941. Distinctions within Ancient Chinese. *Harvard Journal of Asiatic Studies* 6, 203-233.

----- 1967. The Contrastive Aspects of the Wu Dialects, *Language* 43:1, 92-101.

----- 1968, 1985 (3rd edition.). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chen, Matthew Y. 1973. Cross-Dialectal comparison: a case study and some theoretical considerations. *Journal of Chinese Linguistics* 1, 38-63.

----- 2000. Tone Sandhi: Patterns across Chinese Dialects. Cambridge: CUP.

Chen Qiaoyi 1982. *Shaoxing Shihua* [History of Shaoxing]. Shanghai: Shanghai Renmin Chubanshe.

----- 1994. *Guoji Baiyue Wenhua Yanjiu* [Study on Shaoxing Culture]. Beijing: Zhongguo Shehui Kexue Chubanshe.

Cheng, Chin-Chuan 1978. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

----- 1992. *Syllable-Based Dialect Classification and Mutual Intelligibility*. Taibei: Zhongguo Jingnei Yuyan ji Yuyanxue. 1, 145-177.

Cheng, R.L. 1966. Mandarin phonological structure. *Journal of Linguistics* 2, 135-58.

Cho, Y.-m. Yu & T. H. King 2003. Semi-syllables and universal syllabification. In C. Fery and R. van de Vijver (eds.) *The Syllable in Optimality Theory.* Cambridge: CUP.

Chomsky, Noam 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam & Morris Halle 1968. *The Sound Pattern of English*. New York: Harper and Row.

Chung, Raungfu 1996. The Segmental Phonology of Southern Min in Taiwan. Taipei: The Crane Publishing Co., Ltd.

Clements, G.N. 1985. The geometry of phonological features. *Phonology Yearbook* 2, 223-252.

----- 1986. Compensatory lengthening and consonant gemination in Luganda. In W.L.M. Wetzels and E. Sezer (eds.) *Studies in Compensatory Lengthening*, 37-77. Foris Publications, Dordrecht.

----- 1990. The role of the sonority cycle in core syllabification. In J. Kingston and M.E. Beckman (eds.), *Papers in Laboratory Phonology I*, 283-333. Cambridge: CUP.

Clements, G.N. & K. Ford 1979. Kikuyu tone shift and its synchronic consequences, *Linguistic Inquiry* 10, 179-210.

Clements, G.N. & S.J. Keyser 1983. *CV-Phonology: A Generative Theory of the Syllable*. Massachusetts: MIT Press.

Crothers, John 1978. Typology and universals of vowel systems. In Joseph H. Greenberg *et al* (eds.) *Universals of Human Language*. Vol.2, Phonology, 93-152. Stanford: Stanford University Press.

Davis, S. 1985. Topics in Syllable Geometry. PhD dissertation, University of Arizona. (Published in 1988, New York: Garland).

----- 1999. On the moraic representation of underlying germinates: evidence from prosodic morphology. In R. Kager, H.G. van der Hulst & W. Zonneveld (eds.), *The Prosody-morphology Interface*. Cambridge: CUP.

De Lacy, Paul 1999.Tone and prominence. ROA #333.

----- 2002. The interaction of tone and stress in Optimality Theory. *Phonology* 19, 1-32.

DeFrancis, John 1990. *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawaii Press.

Ding Bangxin 1992. Hanyu fangyanshi he fangyan quyushi de yanjiu [Study in historical and geographical records of Chinese dialects]. Taibei: Zhongguo Jingnei Yuyan ji Yuyanxue. 1.23-39.

Ding Shengshu 1984. Hanyu yinyun jiangyi [Lectures on Chinese Phonology]. Shanghai: Shanghai Jiaoyu Chubanshe.

Downing, Laura J. 1990. Local and metrical tone shift in Nguni. *Studies in African Linguistics* 21, 261-317.

----- 2003. Stress, tone and focus in Chichewa and Xhosa. In Rose-Juliet Anyanwu (ed.), *Stress and Tone – the African Experience*. Köln: Rüdiger Köppe Verlag.

Dragunov, A.A. 1930. *The HPhags-pa Script and Ancient Mandarin*. Tang Yuzhong (tran.) 1959. Beijing: Science Publishing House.

Duanmu San 1990. A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages. Doctoral dissertation, MIT, Cambridge, Mass. Distributed by MIT Working Papers in Linguistics.

----- 1992. An autosegmental analysis of tone in four Tibetan languages. *Linguistics of the Tibeto-Burman Area* 15.1, 65-91.

----- 1993. Rime length, stress, and association domains. *Journal of East Asian Linguistics* 2.1, 1-44.

----- 1994. Against contour tone units. *Linguistic Inquiry* 25.4, 555-608.

----- 1995. Metrical and tonal phonology of compounds in two Chinese dialects. *Language* 1.2, 225-259.

----- 1999. The Syllable in Chinese. In Harry van der Hulst & Nancy A. Ritter (ed.) *The Syllable: Views and Facts*. 477-500. Berlin: Mouton de Gruyter.

----- 2000a. Phonology of Standard Chinese, Oxford: UOP.

----- 2000b, Tone: An overview. In Lisa Lai-Shen Cheng and Rint Sybesma (eds.). *Studies in Generative Grammar 48*. 251-286. Berlin: Mouton de Gruyter.

Durand, Jacques 1990. Generative and Non-Linear Phonology. Essex: Longman.

Everett, Daniel 1988. On Metrical Constituent Structure in Piraha Phonology. *Natural Language & Linguistic Theory* 6, 207-246.

Ewen, Colin J. & Harry van der Hulst 2001. *The Phonological Structure of Words*: an Introduction. Cambridge: CUP.

Féry, Caroline & Ruben van de Vijver 2003. *The Syllable in Optimality Theory*. Cambridge: CUP.

Fry, D. B. 1958. Experiments in the perception of stress, *Lang. Speech* 1, 126-152.

Fu Guotong 1986. Wuyu de Fenqu [Classification of the Wu dialects] *Fangyan*, 1, 1-7.

Fudge, Erik C. 1969. Syllables. *Journal of Linguistics* 5, 253-286.

----- 1987. Branching structure within the syllable. *Journal of Linguistics* 23, 359-377.

Gandour, Jackson T. 1978. The perception of tone. In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey*. New York: Academic Press.

Geudens, A. & D. Sandra 2001. Beyond Children's Implicit Phonological Awareness: No Evidence for Onset and Rime Units? Ms, University of Antwerp, Belgium.

Ghini, Mirco 2001. Place of articulation first. In T. Alan Hall (ed.) *Distinctive Feature Theory*. Berlin: Mouton de Gruyter.

Goedemans, R.W.N. 1998. Weightless Segments: A Phonetic and Phonological Study Concerning the Metrical Irrelevance of Syllable Onsets. LOT Dissertations No 9. The Hague: Holland Academic Graphics.

Goldsmith, John A. 1976. Autosegmental phonology. (Unpublished Ph.D dissertation MIT. Distributed Bloomington, Indiana: Indiana University Linguistics Club.)

----- 1982. Accent systems. In van der Hulst and Smith, pp.47-64.

----- 1987. Tone and accent, and getting the two together. *BLS* 13, 88-104.

----- 1990. *Autosegmental and Metrical Phonology*. Basil Blackwell, Oxford.

----- 1995. Phonological theory. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. 1-23. Cambridge: Blackwell.

Grønnum, N. & H. Basbøll 2001. Consonant length, stød and morae in Danish. *Working Papers* 19, 46-49. Lund University.

Gussenhoven, Carlos & Haike Jacobs 1998. *Understanding Phonology*. London: Arnold.

Gussenhoven, C. & G. Bruce 1999. Word prosody and intonation. In H. van der Hulst (ed.) *Word prosodic systems in the languages of Europe.* 233-271. Berlin: Mouton de Gruyter.

Hall, Tracy Alan 1992. Syllable Structure and Syllable-Related Processes in German. Tübingen: Niemeyer.

----- 2001 (ed.). *Distinctive Feature Theory*. Berlin: Mouton de Gruyter.

Halle, Morris 1995. Feature geometry and feature spreading. In Morris Halle (2002), *From Memory to Speech and Back: Papers on Phonetics and Phonology 1954-2002*. Berlin/New York: Mouton de Gruyter.

----- 2005 (forthcoming). Morris Halle's Comment on Clements's paper (to be revised). In Charles Cairns (ed.), *CUNY*.

Halle, M. & K. Steven 1971. A note on laryngeal features. In Morris Halle (2002). Berlin/New York: Mouton de Gruyter.

Halle, M. & G. N. Clements 1983. *The Problem Book in Phonology*. Cambridge, Massachusetts: MIT Press.

Harris, John 1983. Syllable Structure and Stress in Spanish: A Non-linear Analysis. Cambridge, Massachusetts: MIT Press.

----- 1985. Phonological Variation and Change. Cambridge: CUP.

----- 1990. Segmental complexity and phonological government. *Phonology* 7, 255-300.

----- 1994. *English Sound Structure*. Oxford, UK: Blackwell.

Harris, J. & G. Lindsey 1995. The elements for phonological representation. In Durand & Katamba, 34-79.

Hashimoto, Ann Oikan 1987. Tone sandhi across Chinese dialects. In *The Wang Li Memorial Volumes*. 445-474. Chinese Language Society of Hong Kong.

Haudricourt, A. G. 1954. De l'origine des tons en Vietnamien. *Journal Asiatique* 242, 69-82.

Hawkins, S. & K.N. Stevens 1985. Acoustic and perceptual correlates of the non-nasal-nasal distinction for vowels. *Journal of the Acoustical Society of America* 77, 1560-75.

Hayes, B.P. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253-306.

----- 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Heffner, R-M.S. 1949. *General Phonetics*. Madison: The University of Wisconsin Press.

Heuven, Vincent J. van 1994. What is the smallest prosodic domain? In P. Keating (ed.), *Papers in Laboratory Phonology III: phonological structure and phonetic form*. 76-98. London, Cambridge: CUP.

----- 2004. Phonetic or phonological contract in Dutch boundary tone? In Leonie Cornips & Jenny Doetjes (eds.) *Linguistics in the Netherlands*. Amsterdam: John Benjamins.

Holtman, A. 1996. A Generative Theory of Rhyme: An Optimality Approach, PhD dissertation,Universiteit Utrecht. Utrecht: LEd (OTS dissertation series).

Hombert, Jean-Marie 1978. Consonant types, vowel quality, and tone. In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey*. New York: Academic Press.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan 1979. Phonetic explanations for the development of tones. *Language* 55, 37-58.

Hooper, J.B. 1976. An Introduction to Natural Generative Phonology. New York: Academic Press.

Hope, E.R. 1953. Karlgren's Glottal Stop Initial in Ancient Chinese. Ottawa.

Hornstein, Norbert 1999. Minimalism and quantifier raising. In Epstein, Samuel David & Norbert Hornstein (eds.), *Working Minimalism*. Mass: MIT Press.

Howie, J.M. 1976. An Acoustic Study of Mandarin Tones and Vowels. London: CUP.

Hulst, Harry van der & Nancy A. Ritter 1999. Theories of the syllable. In *The Syllable: Views and Facts*. 13-52. Berlin: Mouton de Gruyter.

----- 1999. Head-driven phonology. In *The Syllable: Views and Facts*. 113-168. Berlin: Mouton de Gruyter.

Hulst, Harry van der & Keith Snider 1993. Issues in the representation of tonal register. In *The Phonology of Tone*. Berlin, New York: Mouton de Gruyter.

Humbert, Helga 1995. *Phonological Segments: Their Structure and Behaviour*. The Hague: Holland Academic Graphics.

Hyman, Larry M. 1973. (ed.). Consonant types and tone. *Southern California Occasional Papers in Linguistics*, 1.

----- 1978. Historical tonology. In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey*. New York: Academic Press.

----- 1983. Are there syllables in Gokana? In Jonathan Kaye *et al*. (eds.), *Current Issues in African Linguistics (vol. 2)*, 171-179. Dordrecht: Foris.

----- 1985. *A Theory of Phonological Weight*. Dordrecht: Foris Publications.

----- 1986. The representation of multiple tone heights. In K. Bogers, H. van der Hulst, & M. Mous (eds), *The Phonological Representation of Suprasegmentals*. 109–152. Dordrecht: Foris.

----- 1989. Accent in Bantu: an appraisal. *Studies in the Linguistic Sciences* 19:111-128.

----- 1990. Non-exhaustive syllabification: evidence from Nigeria and Cameroon. In *Papers from the Parasession on the Syllable in Phonetics and Phonology*. Chicago Linguistic Society 26.175-195.

----- 1993. Register tones and tonal geometry. In Harry van der Hulst & Keith Snider (eds.), *The Phonology of Tone: the Representation of Tonal Register*. Dordrecht: Foris.

----- 2001. Tone system. In Martin Haspelmath et al (eds.), *Language Typology and Language Universals*: An International Handbook, vol. 2, 1367-1380. Berlin & New York: Water de Gruyter.

----- 2003. *A Theory of Phonological Weight*. Stanford: CSLI Publications.

International Phonetic Association, Corporate Author International Phonetic 1999. *Handbook of the IP Association*. Cambridge: CUP.

Itô, Junko 1986. Syllable Theory in Prosodic Phonology. PhD dissertation, University of Massachusetts, Amherst.

Itô, Junko & Armin Mester 2003. On the sources of opacity in OT: coda processes in German. In Caroline Féry and Ruben van de Vijver (eds.) *The Syllable in Optimality Theory*. Cambridge: CUP.

Jacobson, L. C. 1978. Dholuo Vowel Harmony: A Phonetic Investigation (UCLA) *Working Papers in Phonetics 43*). PhD thesis. University of California, Los Angeles.

Jakobson, Roman, Gunnar Fant and Morris Halle 1952. Preliminaries to speech analysis. The distinctive features and their correlates. *Acoustics Laboratory, Massachusetts Inst. of Technology, Technical Report No*. 13, 58. Massachusetts: MIT press.

Jansen, Wouter 2004. Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian, and Dutch. Groningen: Groningen Dissertations in Linguistics 47.

Jones, Daniel 1975. *An Outline of English Phonetics*. (9th edition) Cambridge: CUP.

Jongenburger, Willy & Vincent van Heuven 1991. The distribution of (word initial) glottal stop in Dutch. In *Linguistics in the Netherlands*. Frank Drijkoningen & Ans van Kemenade (eds.) Amsterdam: John Benjamins Publishing Company.

----- 1993. Sandhi processes in natural and synthetic speech. In Vincent J. van Heuven & Louis C. W. Pols (ed.) *Analysis and Synthesis of Speech: Strategic Research towards High-Quality Text-to-Speech Generation*. Berlin: Mouton de Gruyter.

Kager, René 1994. Ternary rhythm in alignment theory. Ms., Utrecht University. ROA-35.

----- 1999. *Optimality Theory*. Cambridge: CUP.

Karlgren, Bernhard 1915-1926. *Etudes sur la Phonologie Chinoise*. Stockholm: P.A. Norsted och Soener.

----- 1954. Compendium of Phonetics in Ancient Chinese and Archaic Chinese. Stockholm: s.n. (Reprinted from The Museum of Far Eastern Antiquities 26, 211-367).

Katamba, Francis 1993, *Morphology*. Hampshire: Palgrave.

Kay J.D. 1989. *Phonology: A Cognitive View*. LEA: Hillsdale, NJ.

----- 1990. Coda licensing. *Phonology* 7, 301-330.

Kaye, J.D., J. Lowenstamm and J.R. Vergnaud 1985. The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* 2, 305-328.

----- 1990. Constituent structure and government in phonology. *Phonology* 7, 193-232.

Kenstowicz, Michael J. 1994. *Phonology in Generative Grammar*. Cambridge: Blackwell.

Kiparsky, P. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry* 10, 421-441.

Kratochvil, Paul 1968. *The Chinese Language Today: Features of an emerging standard*. London: Hutchinson.

Kuryłowicz, J. 1948. Contribution à la théorie de la syllabe. *BPTJ* 8, 80-114.

Labov, W. 1994. Principles of Linguistic Change. Volume 1: Internal Factors. Oxford: Blackwell.

Ladefoged, Peter 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.

----- 1975. *A Course in Phonetics*. Florida: Harcourt Brace & Company.

----- 1983. The linguistic use of different phonation types. In D. Bless & J. Abbs (eds.) *Vocal Fold Physiology*: *Contemporary Research and Clinical Issues*, 351-360. San Diego: College Hill Press.

Ladefoged, Peter & Ian Maddieson 1990. Vowels of the world's languages, *Journal of Phonetics* 18, 93-122.

----- 1996. *The Sounds of the World's Languages*. Oxford: Blackwell Publishing Ltd.

----- 2003. *Phonetic Data Analysis*. Oxford: Blackwell Publishing Ltd.

Lass, Norman J. Mosby 1995. *Principles of Experimental Phonetics*. St. Louis: Mosby.

Lass, Roger & John M. Anderson 1975. *Old English Phonology*. Cambridge: CUP.

Laver, John 1994. *Principles of Phonetics*. Cambridge: CUP.

Leben, William 1973. Suprasegmental Phonology. (Unpublished Ph.D dissertation, MIT.)

Lehiste, I. 1970. *Suprasegmentals*. Cambridge, Mass.: MIT.

Levin, Juliette 1985. A Metrical Theory of Syllabicity. PhD dissertation. Massachusetts: MIT.

Li Aijun, Yu Jue, Chen Juanwen & Wang Xia 2004. A contrastive study of Standard Chinese and Shanghai-accented Standard Chinese. 251-288. In G. Fant, H. Fujisaki, J. Cao & Y. Xu (eds.), *From Traditional Phonology to Modern Speech Processing*. Beijing: Foreign Language Teaching and Research Press.

Li, Rong 1956. Qieyun Yinxi [The Phonology of Qieyun Languages]. Beijing: Kexue Chubanshe.

Lieberman, P. 1960. Some acoustic correlates of word stress in American English. JASA 32, 451-454; Fry kap. 27. B.

Liljencrants, J. & B. Lindblom 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48, 839-62.

Lin, Yen-Hwei 1989. Autosegmental Treatment of Segmental Processes in Chinese Phonology. PhD dissertation. The University of Texas at Austin. Austin.

----- 1990. Prenuclear glides in Chinese. Paper presented at the *Mid-America Linguistics Conference*. University of Kansas.

Lin, Maocan 1995. Beijinghua shengdiao fenbuqu de zhijue yanjiu [Acoustic study in the tone distribution in Beijing Mandarin]. *Shengxue Xuebao* 6, 437-445.

Lindau, Mona 1979. The feature expanded. *Journal of Phonetics 7*, 163-176.

Lindblom, B. *et al*. 1984. Self-organizing processes and the explanation of phonetic universals. In B. Butterworth et al. (ed.) *Explanations for Language Universal.* pp. 181-203. New York: Mouton.

Liu Lunxing 2001. Yinyunxue jichu jiaocheng [Fundamental Course of Chinese Phonology]. Beijing: Zhongguo Shehuikexue Chebanshe.

Lombardi, Linda 1994. *Laryngeal Features and Laryngeal Neutralization*. New York: Garland Publishing, Inc.

----- 2001. Segmental Phonology in Optimality Theory: Constraints and Representations. Cambridge: CUP.

Lu, Dora Hsin-yi 2005. The duration of CGVX and CVX syllable in Mandarin. UIUC.

Maddieson, Ian 1972. Tone system typology and distinctive features. *Proceedings of the Seventh International Congress of Phonetic Sciences*: 958-961.

----- 1974. A note on tone and consonants. *UCLA Working Papers in Phonetics* 27, 18-27.
----- 1978. Universal of tone. In J.H. Greenberg (ed.), *Universals of Human Languages*, vol. II, pp.335-366. Stanford University Press.
----- 1984a. *Patterns of Sounds*. Cambridge: CUP.
----- 1984b. The effects on F_0 of a voicing distinction in sonorants and their implications for a theory of tonogenesis. *Journal of Phonetics* 12, 9-15.
McCarthy, John & Alan Prince 1993. Generalized alignment. 79-153. In G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology* 1993. Dordrecht: Kluwer.
Manfredi, Victor 1993. Spreading and downstep: prosodic government in tone languages. In Snider, Keith & Harry van der Hulst (eds.), *The Phonology of Tone: The Representation of Tonal Register*. Berlin: Mouton de Gruyter.
Mannell, Robert 2002. Modelling of the segmental and prosodic aspects of speech intensity in synthetic speech, *Proceedings of the Ninth Australian International Conference on Speech Science and Technology*, Melbourne.
Matisoff, J.A. 1973. Tonogenesis in Southeast Asia. In L. Hyman (ed.), *Consonant Types and Tone*, Southern California Occasional Papers in Linguistics 1.
McCarthy, John J. 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17, 207-263.
----- 1988. Feature geometry and dependency: a review. *Phonetica* 45, 84-108
McCarthy, J. & A. Prince 1986. *Prosodic Phonology*. Ms, University of Massachusetts, Amherst.
----- 1995. Faithfulness and reduplicative identity. In Beckman, Walsh Dickey, and Urbanczyk (eds.), 249-384. *Papers in Optimality Theory*. University of Massachusetts Occasional Papers in Linguistics 18. Amherst, Mass: Graduate Linguistic Student Association.
McCawley, James 1970. Some tonal systems that come close to being pitch accent systems but don't quite make it. *CLS* 6, 526-532.
----- 1978. What is a tone language? In Victoria A. Fromkin (ed.) *Tone: A Linguistic Survey*. New York: Academic Press.
Mohr, B. 1971. Intrinsic variations in the speech signal. *Phonetica* 23, 65-93.
Morelli, Frida 1999. The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory. Ph.D. dissertation, University of Maryland at College Park.
Myers, S. 1987. Vowel shortening in English. *NL&LT* 5, 485-518.
Nartey, Jonas N.A. 1982. On Fricative Phones and Phonemes: Measuring the Phonetic Differences Within and Between Languages. *UCLA Working Papers in Phonetics* 55.
Neugebauer, Moritz 2002. Featural alternation and formal representation. DGfS/LSA Summer School '*Formal and Functional Linguistics*', Student Evening Session. Heinrich-Heine-Universität Düsseldorf.

Nevin, Bruce 1993. Harris the revolutionary: phonemic theory (revision of representations of contrast: features and phonemes), to be presented at *ICHOLS* VI 8/14/93 at Georgetown.

Norman, Jerry 1988. *Chinese*. Cambridge: CUP.

Odden, David 1988. Predictable tone systems in Bantu. In H. van der Hulst and N. Smith (ed.) *Autosegmental Studies on Pitch Accent*. 225-251. Dordrecht: Foris.

----- 1995. Tone: African languages. In Goldsmith (ed.), 1995. 444–475.

Ohde, R. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustic Society of America* 75, 224-230.

Oostendorp, Marc van 1995. Vowel Quality and Phonological Projection. PhD dissertation, Tilburg University.

Padgett, Jaye 2001. Contrast dispersion and Russian palatalization. In E. V. Hume and K. Johnson (eds.), *The Role of Speech Perception in Phonology*. 187-218. San Diego, CA: Academic Press.

----- 2003. Russian voicing assimilation, final devoicing, and the problem of [v] (*or*, The mouse that squeaked). University of California, Santa Cruz.

Paradis, Carole, and Jean-Francois Prunet 1989 On coronal transparency. *Phonology* 6.317-348.

Parkinson, Stephen 1983. Portuguese nasal vowels as phonological diphthongs. *Lingua* 61, 157-77.

Pater, Joe 1995. On the nonuniformity of weight-to-stress and stress preservation effects in English. Ms. McGill University. ROA-107.

Pike, Kenneth L. 1949 (1982, the 10th printing). *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion.* Ann Arbor: the University of Michigan Press.

Pike, K. & E. Pike 1947. Immediate constituents of Mazateco syllables. *International Journal of American Linguistics* 13, 78-91.

Prince, Alan 1983. Hierarchy without constituency in stress theory. *First U. Mass. Conference on Hierarchy and Constituency in Phonology*. Amherst.

----- 1990. Quantitative consequences of rhythmic organization. In Karen Deaton, Manuela Noske, and Michael Ziolkowski (eds.), *CSL 26-II*: *Papers from the Parasession on the Syllable in Phonetics and Phonology*, 355-398. Chicago: Chicago Linguistics Society.

Prince, Alan & Paul Smolensky 1993. Optimality Theory. Technical Report No. 2, Rutgers Center for Cognitive Science, Rutgers University.

Pulleyblank, Douglas 1962. The consonantal system of Old Chinese, Part II. *Asia Major*, 9(58), 206-265.

----- 1986. *Tone in Lexical Phonology*. Dordrecht: Reidel.

----- 1996. Optimality Theory and learning algorithms: The representation of recurrent featural asymmetries. In Durand and Laks (eds).

Radford, Andrew 1997. Syntactic Theory and the Structure of English: a Minimalist Approach. Cambridge: CUP.

Roca, Iggy 1994. *Generative Phonology*. London: Routledge.

Rosenthall, Samuel 1997. *Vowel/Glide Alternation in a Theory of Constraint Interaction*. New York: Garland Publishing, Inc.

Rosner, B.S. & J.B. Pickering 1994. *Vowel Perception and Production*. Oxford: OUP.

Rowicka, Grażyna J. 1999. *On Ghost Vowels: A Strict CV Approach*. The Hague: Holland Academic Graphics.

Rowicka, Grażyna J. & Jeroen M. van de Weijer 1994. Prosodic constraints in the lexicon of Polish: the case of derived imperfectives. *The Linguistic Review* 11, 49-76.

Sagey, E.C. 1986a. The Representation of Features and Relations in Non-Linear Phonology. Doctoral dissertation, Cambridge, Massachusetts: MIT.

----- 1986b. The timing of contour segments. *MIT Working Papers in Linguistics* 8, 208-20.

Sampson, Geoffrey 1969. A note on Wang's phonological features of tone. *International Journal of American Linguistics* 35, 62-66.

Selkirk, Elisabeth 1982. The Syllable. In Harry van der Hulst & Norval Smith (eds.), *The Structure of Phonological Representations*. Part II. 337-384. Dordrecht: Foris Publications.

----- 1984. *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge, Massachusetts: MIT Press.

Shi, Feng, Lin Shi and Rongrong Liao 1987. An experimental analysis of the five level tones of the Gaoba Dong language. *Journal of Chinese Linguistics* 15, 335-361.

Shih, Chilin 1986. The Prosodic Domain of Tone Sandhi in Chinese. PhD. Dissertation, University of California, San Diego.

----- 1997. Mandarin third tone sandhi and prosodic structure. In J. Wang and N. Smith (eds.), *Studies in Chinese Phonology*, pp. 81-124. Berlin and New York: Mouton de Gruyter.

Shryock, Aaron Michael 1995. Investigating laryngeal contrasts: an acoustic study of the consonants of Musey. *UCLA Working Papers in Phonetics* No. 89.

Simmons, Richard van Ness 1999. *Chinese Dialect Classification*. Amsterdam: John Benjamins Publishing Co.

Singh, Sadan & Kula S. Singh 1976. *Phonetics: Principles and Practices*. Maryland: University Park Press.

Sluijter, A.M.C. & V.J. van Heuven 1996. Spectral balance as an acoustic correlate of linguistic stress. *JASA*, in press.

Smith, Norval 1999. A preliminary account of some aspects of Leurbost Gaelic syllable structure. In *The Syllable Views and Facts*. Berlin: Mouton de Gruyter. 577-630.

Snider, Keith 1988. Towards the representation of tone: a three-dimensional approach. In Harry van der Hulst and Norval Smith (eds.), *Features, Segmental Structure and Harmony Processes*, part 1, pp. 237-267. Dordrecht: Foris Publications.

----- 1999, *The Geometry and Features of Tone*. Arlington: The Summer Institute of Linguistics, Inc.

Spencer, Andrew, 1991. *Morphological Theory*. Oxford: Blackwell.

Steriade, Donca 1987. Redundant Values. In A. Bosch and E. Schiller (eds.) *Parasession on Autosegmental and Metrical Phonology*, Chicago Linguistic Society vol. 23, University of Chicago.

Stevens, K. N. *et al*. 1964. Crosslinguistic study of vowel discrimination. *Journal of the Acoustical Society of America*, 36, 1989.

----- 1969. Crosslinguistic study of vowel perception. *Language and Speech*, 12, 1-23.

Stewart, J. M. 1967. Tongue root position in Akan vowel harmony. *Phonetica* 16, 185-204.

Szigetvári, Peter 1997. On affricates. In Agnes B. Farkas (ed.), *Proceedings of the First Symposium of Doctoral Students in Linguistics*. Budapest: theoretical Linguistics Programme.

Thurgood, Graham 1996. Language contact and the directionality of internal drift: the development of tones and registers in Chamic. *Language* 72, 1-31.

----- 2000. Voice quality differences and the origin of diphthongs. *Berkeley Linguistics Society* 25, 295-303.

Topintzi, Nina 2004. Moraic onsets and WSP partition in Pirahã. A paper for *the 2nd Cambridge Postgraduate Conference*, UCL.

Torre, Erik Jan van der 2003. Dutch Sonorants: The Role of Place of Articulation in Phonotactics. Utrecht: LOT.

Traill, A. 1985. Phonetic and Phonological Studies of !Xóõ Bushman. Hamburg: Buske.

Trask, R.L. 1996. A Dictionary of Phonetics and Phonology. London: Routledge.

Trigo, Lauren 1988. On the Phonological Deviation and Behavior of Nasal Glides. PhD dissertation. MIT.

Trommelen, Mieke 1984. *The Syllable in Dutch*. Dordrecht: Foris Publications.

Tung, Chau-Hui 1983. *Chinese Syllables and English Syllables*. Student Book. Taipei.

Veer, Bart van der 2005. Italian Mobile Diphthongs: A Phonetic and Phonological Approach. Unpublished PhD dissertation (forthcoming).

Vennemann, Theo 1988. *Preference laws for syllable structure*. Berlin, New York, Amsterdam: Mouton de Gruyter.

Walli-Sagey, Elisabeth 1985. On the representation of complex segments and their formation in Kinyarwanda. In Wetzels, Leo & Engine Sezer (eds.), *Studies in Compensatory Lengthening*. Dordrecht: Foris Publications.

Walton, A. Ronald 1983. Tone, Segment, and Syllable in Chinese: A Polydimensional Approach to Surface Phonetic Structure. New York: The Cornell University East Asia Papers.

Wan, I-Ping 1999. Mandarin Phonology: Evidence from Speech Errors. PhD dissertation. SUNY-Buffalo.

----- 2002. *Alignment of Prenuclear Glides in Mandarin*. Taipei: The Crane Publishing Co., Ltd.

Wang, Futang 1959. Shaoxing jiyin [Sounds of Shaoxing]. Yuyanxue Luncong, 3. Shanghai: Shanghai Jiaoyu Chubanshe.

Wang, Jenny Zhijie 1993. The Geometry of Segmental Features in Beijing Mandarin, PhD dissertation, University of Delaware.

----- 1999. Beijinhua de Yinjie yu yinxi [Syllable structure and others phonological issues in Beijing Mandarin]. In Xu Liejiong *et al.* (eds.), Gongxing yu gexing---hanyu yuyanxuezhong de zhengyi [Commonality and Individuality: Arguments in Chinese Linguistics]. Beijing: Beijing Yuhan Wenhua Daxue Chubanshe.

Wang, Jialing & Norval Smith 1997. *Studies in Chinese Phonology*. Berlin: Mouton de Gruyter.

Wang, Li 1980. Nanbeichao shiren yongyunkao [Study in application of rhyming by poets in Nanbei Dynasty]. *Tsinghua Journal*, 11-3. In Longchong Bingdiao Zhaiwenji. No.1 Beijing: Zhonghua Shuju.

----- 1963 (reprinted in 2003). Hanyu yinyun [The Phonology of Chinese]. Beijing: Zhonghua Shuju.

----- 1985. Hanyu yuyinshi [History of Chinese Phonology]. Beijing: Zhongguo Shehuikexue Chubanshe.

Wang, William S.-Y. 1967. Phonological features of tone. *IJAL* 33.2, 93-105.

----- 1968. The basis of speech. *Project on Linguistic Analysis Reports*, University of California at Berkeley. Reprinted in 1971, by C. E. Reed (ed.), *The Learning of Language*.

Wang, H Samuel & Chih-ling Chang 2001. On the status of the prenuclear glides in Mandarin Chinese. *Language and Linguistics*. 2.2, 243-260.

Weijer, Jeroen M. van de 1994. *Segmental Structure and Complex Segments*. Doctoral dissertation. HIL. Second edition in 1996. Tübingen: Max Niemeyer Verlag.

Wells, J.C. 1995. New syllabic consonants in English. In J. Windsor Lewis (ed.), *Studies in General and English Phonetics*. Essays in honour of Prof. J.D. O'Connor. London: Routledge.

Wiese, Richard 1997. Underspecification and the description of Chinese vowels. In Wang & Smith (ed.), *Studies in Chinese Phonology*, 219-249. Berlin: Mouton de Gruyter.

Williamson, Kay 1977. Multivalued features for consonants. *Language* 53, 843-871.

Woo, Nancy 1969. Prosody and phonology. Doctoral dissertation, MIT, Cambridge, Mass.

Wright, J.T. 1980. The behavior of nasalized vowels in the perceptual vowel space. In *Reports of the Phonology Laboratory*, Univ. California, Berkerley 5, 127-163.

Wright, Martha S. 1983. A Metrical Approach to Tone Sandhi in Chinese Dialects. PhD. Dissertation, University of Massachusetts, Amherst.

Xu Baohua & Tang Zhenzhu 1988. Shanghai Shiqu Fangyanhzi [Dialect of Shanghai]. Shanghai: Shanghai Jiaoyu Chubanshe.

Xu Yunyang 1995. Wuyu yanjiu [Study in the Wu dialects]. Hong Kong: Xin Ya xueshu jikan. 11.

Xue Fengsheng 1975. *Phonology of Old Mandarin*. Den Haag. Lu Guoyao & Shi Jianguo (tran.), 1990. Beijing: Beijing Yuyanxueyuan Chubanshe.

----- 1986. Guoyu yinxi jiexi [Analysis of Chinese Phonology]. Taibei: Xuesheng Shuju.

Yan Yiming 1994. Wuyu gaishuo [Survey of the Wu dialects]. Shanghai: East China Normal University Publishing House.

----- 1994. Wuyu fenqu lunhang [Classification of the Wu dialects]. In Xu Yunyang (ed.), *Studies of the Wu Dialects*. Hong Kong: University of Chinese.

Yang Jianqiao 1996. Hanyu xiandai yinyunxue [Modern Phonology of Chinese]. Shanghai: Fudan University Press.

Yang Wei & Yang Naijun 2000. *Shaoxing Fanyan* [Shaoxing Dialect]. Guoji Wenhua Chuban Gongsi.

Yin, Yun-Mei 1989. Phonological Aspects of Word Formation in Mandarin, Chinese. PhD dissertation. The University of Texas at Austin. Austin.

Yip, Moira 1980. The Tonal Phonology of Chinese. Doctoral dissertation, Cambridge, Massachusetts: MIT.

----- 1982. Reduplication and CV skeletal in Chinese secret languages. *Linguistic Inquiry* 13, 637-662.

----- 1988. The Obligatory Contour Principle and phonological rules: a loss of identity. *Linguistic Inquiry* 19, 65-100.

----- 1989. Contour tones. *Phonology* 6.1, 149-74.

----- 1993. Cantonese Loanword Phonology and Optimality Theory. *Journal of East Asian Linguistics* 2, 261-291.

----- 1996. Lexical optimization in languages without alternations. In J. Durand & B. Laks (eds.), *Current Trends in Phonology: Models and Methods*. 757-788. CNRS. Paris-X and University of Salford: University of Salford publications.

----- 1997. Consonant-vowel interaction in Cantonese. In Wang & Smith (ed.), *Studies in Chinese Phonology*. Berlin: Mouton de Gruyter.

----- 2001. Non-arguments for sub-syllabic constituents [C]. Proceeding of the 7th International and 19th National Conference on Chinese Phonology. 125-144.

----- 2002. *Tone*. Cambridge: CUP.

----- 2003. Casting doubt on the Onset-Rime distinction. In *Lingua* 113, 779-816.

Yu, John 2000. *Optimization of Wu Dialect Classification*. Shanghai: Fudan University Press.

Zhan Bohui 1991. Hanyu fangyan ji fangyan diaocha [Chinese Dialects and the Dialect Investigation]. Wuhan: Hubei Jiaoyu Chubanshe.

Zhang, Jie 2002. The Effects of Duration and Sonority on Contour Tone Distribution—A Typological Survey and Formal Analysis. Routledge, New York.

Zhang Jisheng 2000. *A Study on English Phonological System*. Zhongqing: Zhongqing Chubanshe.

----- 2003. Optimality theory and syllable acceptability of Chinese loanwords. In *Journal of PLA Foreign Language Institute* 4, 7-12.

----- 2005. A phonological approach to clitics in Shaoxing Chinese. In *Leiden Papers in Linguistics* 2.2, 69-79.

Summary

This dissertation presents an analysis of central aspects of the phonology of Shaoxing Chinese (SX), one of the Wu dialects of the Chinese language, from a synchronic perspective, with the secondary goal of casting some light on current issues in Modern Chinese (Mandarin). The analysis mainly covers (i) the surface inventory of consonants and vowels and their distribution, (ii) the underlying vowel system, (iii) the syllable structure, focusing in particular on the status of the prenuclear glide in the syllable, (iv) the phonotactics of SX, in particular the possible Finals and their combination with possible Initials in the syllable, (v) the tonal inventory of SX, (vi) the consonant-tone interactions, and (vii) the tone sandhi rules.

Languages differ considerably with regard to their inventories of consonants and vowels, and also with regard to the phonological grammar that specifies how these sounds can be combined to form words and utterances. What makes SX different from other languages is its rather large inventory of consonants and small number of phonemic vowels. There are three remarkable characteristics in the SX segmental system. First, the language displays a distinction between voiceless aspirated, voiceless unaspirated and voiced obstruents. Secondly, SX has the "filler" onset consonants [ʔ] and [ɦ] in the surface presentation, which are correlated with a register division. Thirdly, vowels display a large number of surface variants: 14 allophonic vowels are derived from six underlying segments (/i u e ɤ o a/), which makes it necessary to postulate a number of phonetic realisation rules, as presented in the dissertation.

Syllable structure has attracted the attention of phonologists for many decades; the syllable allows the formulation of many generalizations both at the segmental and at higher prosodic levels. There have been many approaches to the internal structure of the syllable. This dissertation argues in favour of a multiple-specifier X-bar syllable structure, which allows a syllable maximally to have three sub-constituents: N'' (Final), N' (Rhyme) and N^0 (Nucleus). In this X-bar schema, a syllable is maximally parsed into Onset and Final, instead of Onset and Rhyme (as in the classical OR models), and the problematic prenuclear glide is located in the specifier position of N''.

It is well documented cross-linguistically that voiceless initial obstruents induce high tones and voiced initial obstruents induce low tones. However, consonant-tone interaction is still a controversial issue, both phonetically and phonologically. This dissertation presents an objective analytic description of consonant-tone interaction in SX, with evidence from, for instance, syllable merger in cliticization, phonetic onset insertion, sonorant initials, etc. We conclude that both voiced initial obstruents and low-register tones occur in underlying representation in SX.

Tone sandhi is a common phenomenon in tonal languages; yet its complexity and the sheer variety of sandhi rules have prevented a fully-fledged analysis in contemporary linguistic studies. The dissertation makes an effort to formalize the intricacies of tone sandhi in SX. We argue that tone sandhi in SX is realized by tone feature spreading and delinking, and does not involve register features. In this dissertation I present an overview of disyllabic sandhi forms. Although neither rule-based nor constraint-based accounts have given a complete and explicit formulation of a tone-sandhi system for any of the Chinese languages so far, I present a metrically-based analysis, claiming that SX is a right-prominent language and that the stress foot is the tone sandhi domain. I present a hierarchical constraint ranking that precisely captures the tonal behaviour in sandhi and accounts for all the disyllabic sandhi rules in SX.

Samenvatting (Summary in Dutch)

Deze dissertatie bevat een analyse van de belangrijkste aspecten van de synchrone fonologie van het Chinees van Shaoxing (SX), een van de Wu-dialecten van de Chinese taalfamilie. Als tweede doel heeft de dissertatie enig licht te werpen op verschillende aspecten van de fonologie van het Mandarijn. De analyse betreft vooral (i) het oppervlakte-inventaris van consonanten en vocalen en hun distributie, (ii) het onderliggende vocaalsysteem, (iii) de syllabestructuur, en vooral de positie van de pre-nucleaire glijklank in de syllabe, (iv) de fonotactische regels voor SX, in het bijzonder de mogelijke combinaties in de *Final* en combinaties van *Finals* met *Initials* in de syllabe, (v) het inventaris van tonen in SX, (vi) de interactie tussen tonen en consonanten, en (vii) de regels voor tone sandhi.

Talen verschillen voor wat betreft hun inventarissen van consonanten en vocalen, en ook wat betreft de fonologische grammatica die specificeert hoe deze klanken gecombineerd kunnen worden om woorden en langere uitingen te vormen. SX verschilt van andere talen door het betrekkelijk grote aantal consonanten en kleine aantal fonemische vocalen. Er zijn drie opvallende aspecten aan het segmentsysteem van het SX: in de eerste plaats laat de taal een contrast zien tussen stemloos geaspireerde, stemloos ongeaspireerde en stemhebbende plosieven. In de tweede plaats heeft de taal de “opvulconsonanten” [ʔ] en [ɦ], die gecorreleerd zijn met de twee verschillende toonregisters. In de derde plaats laten de klinkers een vrij groot aantal allofonische varianten zien: van zes onderliggende klinkers (/i u e ɤ o a/) worden aan de oppervlakte 14 allofonen afgeleid. Hiervoor zijn een aantal fonetische realisatieregels noodzakelijk, die in deze dissertatie gepresenteerd worden.

Syllabestructuur is al vele jaren een onderwerp van gesprek in de fonologie; de syllabe laat de formulering van vele generalisaties toe zowel wat betreft het segmentele als het prosodische niveau. Er zijn vele benaderingen wat betreft de interne structuur van de syllabe. In deze dissertatie wordt betoogd dat deze structuur het beste te beschrijven is in termen van een X-bar structuur waarin meerdere *specifier* posities mogelijk zijn. Hierin heeft een syllabe maximaal drie subconstituenten: N'' (*Final*), N' (Rijm) en N^0 (Nucleus). In dit model kan een syllabe maximaal geparseerd worden in Onset en Final, in plaats van Onset en Rijm (zoals in de klassieke OR-modellen), zodat de problematische prenucleaire glijklank in de specifier positie van N'' gesitueerd worden.

Het is welbekend dat stemloze initiële obstruenten een hoge toon in een volgende vocaal kunnen veroorzaken en dat stemhebbende obstruenten een tegenovergesteld effect kunnen hebben. De interactie tussen consonanten en toon blijft echter een controversieel onderzoeksterrein, zowel uit als fonetisch als fonologisch oogpunt. Deze dissertatie biedt een objectieve beschrijving en analyse van deze interactie in het SX, met evidentie van onder andere syllabevorming in cliticizering, fonetische consonant epenthese, initiële sonorante medeklinkers, etc. De conclusie is dat zowel stemloze als stemhebbende initiële medeklinkers in onderliggende representatie aanwezig moeten zijn.

Toon sandhi is een frequent verschijnsel in toontalen; huidige studies hebben echter geen geheelomvattend analyse kunnen geven door de complexiteit en variëteit van het verschijnsel. In deze dissertatie is een poging gedaan om de verschillende aspecten van toon sandhi in het SX te analyseren. Wij stelden vast dat toon sandhi gerealiseerd wordt door kenmerkspreiding en -deletie, en dat registerkenmerken daarbij niet betrokken zijn. Ik heb een overzicht gepresenteerd van sandhi regels in vormen van twee syllaben, en een analyse op grond van metrische kenmerken gegeven, waarin geclaimd wordt dat SX een rechts-prominente taal is waarin de voet het domein voor toon sandhi vormt. Ik heb een hiërarchische verzameling constraints gepresenteerd die het gedrag van de tonen in toonsandhi-vormen van twee syllaben precies beschrijven.

概要

本論文運用當代音系學各種理論，對漢語紹興方言（南部吳方言）的音系進行了系統的共時性研究分析，其中也對比分析了漢語普通話的一些音系要點。論文的主要內容包括：①紹興方言的輔音和母音及其分佈；②底層母音系統；③音節結構，尤其是介音的音節結構成分；④音段的序列規則，主要是聲母與韻母的組合規則；⑤聲調結構；⑥聲母-聲調的相互關係；⑦連續變調規則。

一種語言區別與另一種語言的要素之一是一種語言所特有的一套輔音母音系統，以其闡述這些音素如何組合表義的音系法則。紹興方言的音段系統有三大特徵：第一，紹興方言仍然完整地保留著古漢語特有的“幫滂並”、“端透定”、“見溪群”聲母（即送氣清阻塞音、不送氣清阻塞音和濁阻塞音）三分格局。第二，紹興方言的表層表達式有“填補式”聲母[ʔ]和[ɦ]與聲調的陰陽調關聯。第三，紹興方言的底層結構只有六個母音（/i u e ɤ o a/），而其表層表達式有十四個變體母音（[i y ɪ ʅ u e ɛ ə ɵ ɤ o a ɑ ɒ]），這需要一系列語音實現規則進行推導，或一系列制約條件進行優化選擇。

長期以來，音節結構一直是音系學家們所討論的課題，音節可以用音段層面和更高的韻律層面的許多概括進行公式化表述。就音節的內部結構而言，存在著許多不同的分析，尤其對漢語音節中韻核前介音的歸屬，一直沒有統一定論。本論文提出了一個多指示語語杠（multiple-specifier X-bar）音節結構，該結構允許音節結構最大可以有三個內部結構成分：N'' (韻母), N' (音韻) 和 N^0 (韻核)。運用這一語杠組合式，一個音節可最大劃分爲聲母和韻母（Onset-Final），而不是經典音節理論主張的聲母和音韻（Onset-Rhyme）。論文認爲，紹興方言韻核前介音既不屬於聲母，也不屬於音韻，而是屬於N'' (韻母)指示語位置。語杠音節結構具有生成語法的普遍性，它同樣適合於漢語所有其他方言的音節結構。

有大量跨語言語料證明，聲母清阻塞音與聲調的陰調匹配，濁阻塞音與陽調匹配。然而，聲母-聲調的相互關係一直存有爭議。本論文根據紹興方言的一些實證語料，如附著語素形式的合音、“填

補式”語音聲母、聲母響音與聲調的匹配等，對聲母-聲調的相互關系進行了客觀分析，認爲，儘管紹興方言的聲母濁阻塞音和陽調處于完全互補分佈狀況，但兩者都是底層形式。

連續變調是聲調語言中的普遍現象，然而，由於其複雜性和多樣性，當代語言學對連續變調的研究還沒有完全成熟的理論和系統的分析。本論文對紹興方言連續變調的錯綜複雜進行了形式化分析，認爲，紹興方言連續變調是通過聲調特徵的擴展與脫離來實現的，這一過程不涉及調域特徵（register features）。論文根據韻律結構，全面地分析了紹興方言雙音節連續變調形式，認爲紹興方言是右突顯語言，並且其重音音步就是連續變調域。論文提出了一個有等級的制約條件序列，這個制約條件序列準確地抓住了連續變調中聲調的行爲表現，透徹地解釋了所有紹興方言雙音節連續變調的規則。

Curriculum Vitae

Zhang Jisheng was born on 12 March 1955 in Shaoxing, China. He finished his junior middle school in 1971 and in the same year he was sent to Inner Mongolia and began work there as a bricklayer. In 1978 (two years after the Cultural Revolution finished), he entered the Inner Mongolia Teachers' University, majoring in English Language in the Department of Foreign Languages. After his graduation from the university in 1982, he began to teach English in a college in Inner Mongolia and in 1988 he started a job as an English teacher in Ningbo University, which he held until recently. From 1992 to 1993, he studied linguistics, especially phonology, in Queensland University, Australia. In 2001 he visited ULCL, Leiden University, as a visiting scholar for half a year, sponsored by the Chinese Scholarship Council and meanwhile he became a Professor of English Language Teaching in Ningbo University, China. Since 2002, he has worked on his PhD project in LUCL (previously ULCL). He is currently a Professor in East China Normal University in Shanghai.

Index of Authors

Index of Languages

Index of Subjects