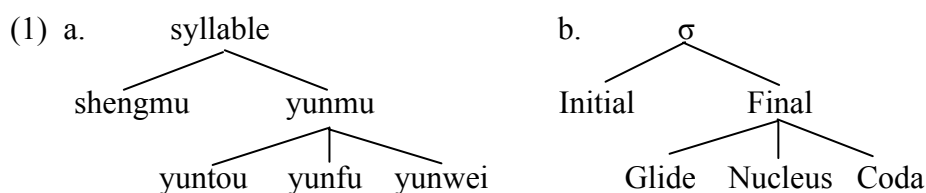


## 2 The Consonants and Vowels of Shaoxing: Surface Representations

### 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):



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,<sup>1</sup> “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.

<sup>1</sup> 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<sup>2</sup> (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 <sup>h</sup>	t <sup>h</sup>	k <sup>h</sup>
	voiced unaspirated	b	d	g

---

<sup>2</sup> Middle Chinese is the language of the Sui, Tang, and Song dynasties (7<sup>th</sup>-10<sup>th</sup> 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 <sup>33</sup> ]	‘soldier’	[p <sup>h</sup> ɪŋ <sup>33</sup> ]	‘marry’	[bɪŋ <sup>22</sup> ]	‘ill’
	[tɪŋ <sup>33</sup> ]	‘book’	[t <sup>h</sup> ɪŋ <sup>33</sup> ]	‘listen’	[dɪŋ <sup>22</sup> ]	‘decide’
	[koŋ <sup>33</sup> ]	‘supply’	[k <sup>h</sup> oŋ <sup>33</sup> ]	‘free time’	[goŋ <sup>22</sup> ]	‘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<sup>3</sup> 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,

<sup>3</sup> 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 <sup>33</sup> ]	‘cloth’	[p <sup>h</sup> u <sup>33</sup> ]	‘berth’
	[tɿŋ <sup>33</sup> ]	‘book’	[t <sup>h</sup> ɿŋ <sup>33</sup> ]	‘listen’
	[kɛʔ <sup>5</sup> ]	‘cut’	[k <sup>h</sup> ɛʔ <sup>5</sup> ]	‘block’
but	[bu <sup>22</sup> ]	‘step’	*[bu <sup>33</sup> ]	
	[dɿŋ <sup>22</sup> ]	‘decide’	*[dɿŋ <sup>33</sup> ]	
	[gɛʔ <sup>3</sup> ]	‘squeeze’	*[gɛʔ <sup>5</sup> ]	

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<sup>h</sup>] 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.	[bɿŋ <sup>22</sup> ]	‘ill’	[bɿŋ <sup>13</sup> ]	‘judge’	[bɿŋ <sup>31</sup> ]	‘level’
	b.	[doŋ <sup>22</sup> ]	‘cave’	[doŋ <sup>13</sup> ]	‘pail’	[doŋ <sup>31</sup> ]	‘same’
	c.	[gɑŋ <sup>22</sup> ]	‘a tool’	[gɑŋ <sup>13</sup> ]	‘confuse’	[gɑŋ <sup>31</sup> ]	‘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]<sup>4</sup>, 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 <sup>33</sup> ]	‘love’	[ʔɛʔ <sup>5</sup> ]	‘duck’
	[ʔjɛ <sup>52</sup> ]	‘smoke’	[ʔɥθ <sup>52</sup> ]	‘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 13<sup>th</sup> and 14<sup>th</sup> centuries, as is shown by the hPhags-pa script, and in nearly all the modern Wu dialects” (cited by Hope 1953).<sup>5</sup> 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

<sup>4</sup> 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.

<sup>5</sup> “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 [ʉẽ]. 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	[ʔáRm]	[áRm]	‘poor’
	elf	[ʔélf]	[élf]	‘eleven’
	oft	[ʔóft]	[óft]	‘often’
	Uhr	[ʔú:ʌ]	[ú:ʌ]	‘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)

$\emptyset \rightarrow [ʔ] / \text{F}[ \_ [-\text{cons}]$

Jongenburger & van Heuven (1991) carried out a phonetic experiment in Dutch with the example of [dat ən ʔɑ:ntal] ‘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:ntal]. 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 <sup>h</sup>	tɕ <sup>h</sup>
	voiced unaspirated	dz	dʒ

Like the SX stops in (2), there are voiceless unaspirated, voiceless aspirated and historically voiced unaspirated affricates in SX<sup>6</sup>. 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:

---

<sup>6</sup> 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 <sup>52</sup> ]               | ‘end’  | [tɕi <sup>33</sup> ]               | ‘post’ |
|      | [ts <sup>h</sup> oŋ <sup>52</sup> ] | ‘dash’ | [tɕ <sup>h</sup> i <sup>33</sup> ] | ‘air’  |
|      | [dzoŋ <sup>31</sup> ]               | ‘worm’ | [dʒi <sup>31</sup> ]               | ‘ride’ |

Middle Chinese had nine affricates in initial position, including [ts], [ts<sup>h</sup>], [dz], [tʃ], [tʃ<sup>h</sup>], [dʒ], [tɕ], [tɕ<sup>h</sup>] and [dʒ], according to the *Guangyun*<sup>7</sup> (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/, /tʃ<sup>h</sup>/, /dʒ/, /tʂ/, /tʂ<sup>h</sup>/, and /dz/.

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. [traɪ]  | ‘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.<sup>8</sup> 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:

<sup>7</sup> *Guangyun* (1008) is an ancient book of traditional Chinese phonology that describes the pronunciation of the Chinese syllables and rhyming system.

<sup>8</sup> 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) [pjɑp<sup>35</sup>] ‘watch’  
 [ɕjɑp<sup>35</sup>] ‘small’  
 [kwɔŋ<sup>52</sup>] ‘light’  
 [fɪwɔŋ<sup>31</sup>] ‘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) [dzjɑp<sup>31</sup>] ‘bridge’                      [tswɔ<sup>33</sup>] ‘do’  
 [tɕ<sup>h</sup>joŋ<sup>5</sup>] ‘lack’                              [ts<sup>h</sup>wo<sup>33</sup>] ‘wrong’  
 [dzjoŋ<sup>31</sup>] ‘poor’                                [dzwo<sup>31</sup>] ‘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)  
 \*<sub>[σ</sub> 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<sup>h</sup>loun] ‘clone’ will be realized as [k<sup>h</sup>ɿloŋ], 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]<sup>9</sup>.

<sup>9</sup>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.<sup>10</sup> For example, the English *clone* and *tank* are [k<sup>h</sup>ɿloŋ] and [t<sup>h</sup>ẽk<sup>h</sup>ɿ], respectively, in SX loan words. The tableau deriving the optimal candidate is given in (18):

(18)

input	k <sup>h</sup> loun	*COMPLEX(ONS)	CODA-COND	MINWD
a.	k <sup>h</sup> loun	*!	*	*
b.	k <sup>h</sup> loŋ	*!		*
c.	☞ k <sup>h</sup> ɿloŋ			
d.	k <sup>h</sup> 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 [tʃi?pu] in SX. The selection of this candidate is represented in the following OT tableau:

<sup>10</sup> 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.		tʃi?p		*!	*
b.	☞	tʃi?pu			
c.		tʃi?			*!

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]	<i>spring</i>	[sʃbi?lɪŋ]	‘spring lock’
b.	[brændi]	<i>brandy</i>	[pa?lɛdi]	‘Brandy wine’
c.	[fræns]	<i>France</i>	[fa?lɛçi]	‘France’
d.	[tʃɔkəlɪt]	<i>chocolate</i>	[tʃ <sup>h</sup> ʃaɔk <sup>h</sup> ə?li?	‘chocolate’
e.	[tʃə:tʃɪl]	<i>Churchill</i>	[tʃ <sup>h</sup> ʃɪtʃi?ɬ]	(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<sup>11</sup>, 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	ç	h
	voiced	v	z	ʒ	fi

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<sup>12</sup> (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 <sup>33</sup> ]	‘pay’	[vu <sup>22</sup> ]	‘attach /enclose’
	[sɑp <sup>35</sup> ]	‘little’	[zɑp <sup>13</sup> ]	‘make/invent’
	[çy <sup>33</sup> ]	‘need’	[zy <sup>22</sup> ]	‘tree’
	[ho <sup>52</sup> ]	‘shrimp’	[fi <sup>31</sup> ]	‘river’

Most languages in the world have no contrast between voiced and voiceless glottal fricatives (/h/ vs. /fi/). In fact, Maddieson (1984a) lists only two languages in the world which have a contrast between /h/ and /fi/, 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 /fi/ (e.g. whether they are fricatives or laryngeals, vowels or approximants; see Maddieson 1984a), /h/ and /fi/ in SX are undoubtedly glottal fricatives

<sup>11</sup> Nartey's (1982) analysis includes Amharic, Arabic, English, Hebrew, Hopi, Japanese, Korean, Navajo, Papago, Pima, Polish, Swedish, Yoruba and Zuni.

<sup>12</sup> 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 <sup>52</sup> ]	‘shrimp’	[fio <sup>31</sup> ]	‘river’
	[hεɾ <sup>5</sup> ]	‘blind’	[fεɾ <sup>3</sup> ]	‘narrow’

According to Maddieson (1984a), about 63.7% of the languages in UPSID<sup>13</sup> have voiceless /h/, while only 4.1% of the languages have voiced /f/. In SX, voiced /f/ 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	
	[fiɑɳ <sup>31</sup> ]	‘roar’	[hɑɳ <sup>52</sup> ]	‘spend (time)’
	[fɛŋ <sup>31</sup> ]	‘stable’	[hɛŋ <sup>52</sup> ]	‘groan’
	[fiɔɾ <sup>3</sup> ]	‘learn’	[hoɾ <sup>5</sup> ]	(surname)
	[fwo <sup>22</sup> ]	‘speech’	[hwo <sup>33</sup> ]	‘spend’
	[fiwe <sup>31</sup> ]	‘return’	[hwe <sup>52</sup> ]	‘ash’
	[fwiẽ <sup>31</sup> ]	‘play’	[hwiẽ <sup>52</sup> ]	‘well-behaved’
	[fiwɔŋ <sup>31</sup> ]	‘king’	[hwɔŋ <sup>52</sup> ]	‘nervous’
	[fiu <sup>22</sup> ]	‘unclear’	[hu <sup>33</sup> ]	‘call’
	[fiɔŋ <sup>31</sup> ]	‘line’	[hɑŋ <sup>52</sup> ]	‘ram’
	[fiwa <sup>22</sup> ]	‘bad’	?	(undecided possibility)
	[fijẽ <sup>22</sup> ]	‘hate’	*	(systematic
	[fijɾ <sup>31</sup> ]	‘oil’	*	impossibility)
	[fijoɾ <sup>3</sup> ]	‘moon’	*	
	[fija <sup>22</sup> ]	‘night’	*	
	[fiqẽ <sup>31</sup> ]	‘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 [fi] followed by an identical final vowel or combination. The reverse is not true, however. In short, [fi] 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 [fi] can be followed by all different vowels, as

<sup>13</sup> 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 <sup>31</sup> ]	‘move’	[ʔi <sup>52</sup> ]	‘clothes’	*[hi <sup>52</sup> ]
	[ɦjɿ <sup>31</sup> ]	‘oil’	[ʔjɿ <sup>52</sup> ]	‘low voice’	*[hjɿ <sup>52</sup> ]
	[ɦjẽ <sup>31</sup> ]	‘salt’	[ʔjẽ <sup>33</sup> ]	‘smoke’	*[hjẽ <sup>33</sup> ]
	[ɦyẽ <sup>31</sup> ]	‘cloud’	[ʔyẽ <sup>33</sup> ]	‘complain’	*[hyẽ <sup>33</sup> ]

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<sup>52</sup>], \*[a<sup>33</sup>], \*[u<sup>31</sup>] and \*[o<sup>22</sup>] 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	ɲ	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 <sup>22</sup> ]	‘grind’	[mi <sup>13</sup> ]	‘rice’	[mu <sup>22</sup> ]	‘grave’
	[n]:	[no <sup>22</sup> ]	‘glutinous’	*[ni]	—	[nu <sup>22</sup> ]	‘anger’
	[ɲ]:	*[ɲo]	—	[ɲi <sup>22</sup> ]	‘two’	*[ɲu]	—
	[ŋ]:	[ŋo <sup>22</sup> ]	‘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<sup>22</sup>] ‘younger sister’, [ne<sup>22</sup>] ‘patient’ and [ɲe<sup>31</sup>] ‘fool’, and [ma<sup>13</sup>] ‘buy’, [na<sup>13</sup>] ‘milk’ and [ɲa<sup>13</sup>] ‘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 \*/ɲi/, \*/ɲja/, \*/ɲu/ and \*/ɲwa/ in SX. The alveolar nasal [n] cannot appear before the front high vowel [i] or the glide [j] while the alveolo-palatal nasal [ɲ] can only appear before [i]/[j].<sup>14</sup> This indicates that [ɲ] is in complementary distribution with [n]. Thus, [ɲ] 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):

$$(30) \quad [n] \rightarrow [ɲ] / \text{ — } \begin{bmatrix} +\text{high} \\ -\text{back} \end{bmatrix}$$

or

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

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

<sup>14</sup> 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)	[n̥i <sup>22</sup> ]	‘two’	*[ni]
	[n̥jɿ <sup>31</sup> ]	‘cow/ox’	*[njɿ]
	[n̥ɪŋ <sup>31</sup> ]	‘silver’	*[niŋ]
	[n̥iʔ <sup>3</sup> ]	‘hot’	*[niʔ]

From the analysis above, we have observed that in SX, there are six alveolo-palatal consonants: [tɕ], [tɕ<sup>h</sup>], [dʒ], [ç], [ʒ] and [ɲ]. 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 ɲ/, which strongly suggests asymmetric behaviour in the class of the nasals (Botma 2004), as shown in (33):

(33)	steen+bok	stee[mɸ]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 /ɾ/. /l/ is a common onset in SX syllables. For example:

(34)	[li <sup>13</sup> ]	‘inside’	[lɛŋ <sup>13</sup> ]	‘cold’
	[le <sup>31</sup> ]	‘come’	[laŋ <sup>13</sup> ]	‘cool’
	[la <sup>52</sup> ]	‘pull’	[loŋ <sup>31</sup> ]	‘dragon’
	[lɔɸ <sup>13</sup> ]	‘old’	[loʔ <sup>3</sup> ]	‘green’
	[ljɿ <sup>31</sup> ]	‘flow’	[ljaŋ <sup>22</sup> ]	‘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:

Place Manner		Place						
		bilabial	labio- dental	dental	alveolar	alveolo- palatal	velar	glottal
stop	-asp	p			t		k	ʔ
	+asp	p <sup>h</sup>			t <sup>h</sup>		k <sup>h</sup>	
	+voice	b			d		g	
affricate	-asp			ts		tɕ		
	+asp			ts <sup>h</sup>		tɕ <sup>h</sup>		
	+voice			dz		dʒ		
fic.	-voice		f	s		ɕ		h
	+voice		v	z		ʒ		ɦ
nasal		m			n	ɲ	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];
- Stricture 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	p <sup>h</sup>	b	m	f	v	ts	ts <sup>h</sup>	dz	s	z	t	t <sup>h</sup>	d	n	l	tc	tc <sup>h</sup>	dz	c	z	n	k	k <sup>h</sup>	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 [ $\pm$ 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, [ $\pm$ strident] is used to distinguish [f] ([+strid]) and [ϕ] ([−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.<sup>15</sup> For example:

---

<sup>15</sup> 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	
[vi <sup>31</sup> ]	[bi <sup>31</sup> ]	‘grease’
[vi <sup>22</sup> ]	[bi <sup>22</sup> ]	‘taste’
[tɕjɑŋ <sup>52</sup> ]	[kɑŋ <sup>52</sup> ]	‘hand in’
[da <sup>22</sup> ]	[do <sup>22</sup> ]	‘big’
[zəŋ <sup>22</sup> ]	[ɲŋ <sup>22</sup> ]	‘recognize’
[tso <sup>35</sup> ]	[tɕja <sup>33</sup> ]	‘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<sup>16</sup> of SX, which can be summarized in the following table:

<sup>16</sup> 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<sup>17</sup>

	<i>Kaikouhu</i>	<i>Qichihu</i>	<i>Hekouhu</i>	<i>Cuokouhu</i>
In open syllables	ɿ	i, ij	u	yj, yɿ
	e	ie	ue	
	a	ia	ua	
	ɤ, ə, ɐ	iɤ, iɤ		
	o	io	uo	yo
	æ, A		uæ	
	ɑɒ	iɑɒ		
With nasal coda	en, əŋ, əŋ	ŋ, ɤŋ	uən, uəŋ	
	aŋ	iaŋ	uaŋ	
	ɒŋ	iɒŋ	uɒŋ	
	uŋ	iuŋ		
With nasalized vowels	æ̃	iæ̃	uæ̃	
	ẽ		uẽ	yẽ
	ẽ	ĩ	uẽ	
Syllabic consonants	m̩, ɲ̩, 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 [uẽ], [uən] and [uəŋ], [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 [yɿ], 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:

<sup>17</sup> 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<sup>18</sup>

	<i>Kaikouhu</i>	<i>Qichihu</i>	<i>Hekouhu</i>	<i>Cuokouhu</i>
In open syllables	ɿ	i	u	
	e	ie	ue	yɿ
	a	ia	ua	
	o	io	uo	
	ɤ	iɤ		
	ɑɒ	iɑɒ		
With nasal coda	əŋ	ɨŋ		
	ɒŋ		uɒŋ	
	aŋ	iaŋ	uaŋ	
	uŋ	iuŋ	uuŋ	
With nasalized vowels	æ̃	iã	uã	
	ẽ		uẽ	yẽ
		ĩ		
With stop coda	əʔ	ɨʔ		
	Aʔ	iAʔ	uAʔ	
	oʔ		uoʔ	yoʔ
	eʔ			
Syllabic consonants	m̩ n̩ ɨ̩ 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 [yɿ] which sounds unusual. Yang & Yang (2000) present a similar Final inventory of SX to Campbell's in (40):

<sup>18</sup> 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<sup>19</sup>

	<i>Kaikouhu</i>	<i>Qichihu</i>	<i>Hekouhu</i>	<i>Cuokouhu</i>
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	ẽ	iẽ	uẽ	
	ỹ		uỹ	yỹ
	ẽ	iẽ		
With stop coda	əʔ	iəʔ		
	aʔ	iaʔ		
	oʔ	ioʔ	uoʔ	
	ɛʔ		uɛʔ	
Syllabic consonants	m̩ n̩ ɳ̩ 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 Campbell'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, [Aʔ] and [ǣ] in (39) refer to [aʔ] and [ẽ] 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

<sup>19</sup> 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<sup>20</sup>

	<i>Kaikouhu</i>	<i>Qichihu</i>	<i>Hekouhu</i>	<i>Cuokouhu</i>
In open syllables	ɿ	i	u	y
	e	je	we	
	a	ja	wa	
	o	jo	wo	
	ɤ	jɤ		
	ɑ	jɑ		
With nasal coda	ən	ɲ		
	ŋ		wŋ	
	aŋ	jaŋ	waŋ	
	oŋ	joŋ	woŋ	
With nasalized vowels	ẽ	jẽ	wẽ	
	ẽ̃		wẽ̃	ɥẽ̃
	ẽ̄	jẽ̄		
With stop coda	əʔ	ɿʔ		
	aʔ	jaʔ	waʔ	
	oʔ	joʔ	woʔ	
	ɛʔ			
Syllabic consonants	m̩ n̩ ɲ̩ 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, C,<sup>21</sup> 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:

<sup>20</sup> 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.

<sup>21</sup> Here 'C' refers to syllabic consonants.

(42)

Yangs'	ŋ̣	ĩ	ɒ	iəŋ	iɔŋ	iaŋ	uĩ	yĩ	iəʔ	ueʔ
Zhang's	ŋ	ẽ	ɑɒ	iŋ	jaŋ		wẽ	ɥẽ	íʔ	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

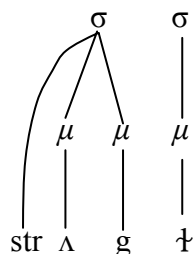
In this subsection, I will examine the syllabic consonants and analyze the phonological rules and constraints concerning the syllabicity of consonants in SX. A syllabic consonant is a segment which has the phonetic characteristics of a consonant but which, in a particular case, functions as a syllabic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

(43)

ṃ	ṇ	ŋ̣	ḷ
----	----	----	----

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) ['strʌgɫ] *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 ['strʌgɫ], in which the syllabic lateral [ɫ] 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 [m], [n], [ŋ] and [l] 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. [m<sup>22</sup>] 'yes'
- b. [n<sup>33</sup>]([~jx<sup>35</sup>]) 'not' (only used together with 'have')
- c. [ŋ<sup>13</sup>] 'five'
- d. [l<sup>22</sup>] '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 [m] 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] / \$ \_ \$

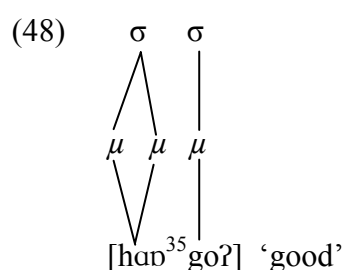
Rule (46) says that any sonorant becomes syllabic when it occurs on its own between syllable boundaries underlyingly. This rule also includes vowels.

### 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 [ $\pm$  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 underlyingly 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.

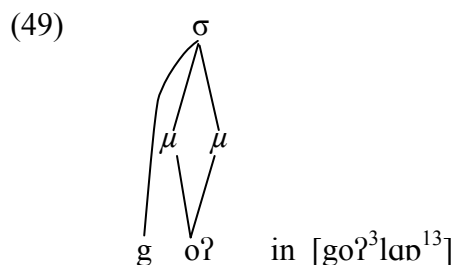
The weight-by-stress principle is largely consistent with constraints like WSP (WEIGHTTOSTRESSPRINCIPLE)<sup>22</sup> (Prince 1990) and SWP (STRESSTOWEIGHTPRINCIPLE)<sup>23</sup> (Prince 1983; Myers 1987; Prince & Smolensky 1993). Neither WSP nor SWP, however, constrains unstressed syllables. Weight-by-stress 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 [hɑp<sup>35</sup>goʔ] ‘good’, the second syllable is unstressed because it is not lexically meaningful. Their moraic syllable structure can be represented as follows:



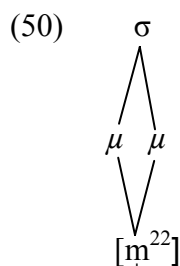
The representation in (48) shows that the first syllable in [hɑp<sup>35</sup>goʔ] is stressed and is heavy and the second syllable is only an adjective marker grammatically and is unstressed. The unstressed syllable is short and always toneless or has a neutral tone because it is monomoraic. Thus, the 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<sup>3</sup>] in [goʔ<sup>3</sup>lɑp<sup>31</sup>] ‘standstill’ is stressed and bimoraic, because it is a lexical word. The moraic structure of the stressed [goʔ] in [goʔ<sup>3</sup>lɑp<sup>31</sup>] is represented as follows:

<sup>22</sup> WSP (WEIGHTTOSTRESSPRINCIPLE): Heavy syllables must be stressed.

<sup>23</sup> SWP (STRESSTOWEIGHTPRINCIPLE): Stressed syllables must be heavy.



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 [n̩²²] ‘yes’ can be illustrated in (50):



Duanmu (1999, 2000a) claims that all heavy syllables (CVV,<sup>24</sup> 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 [n̩³³], 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 out 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

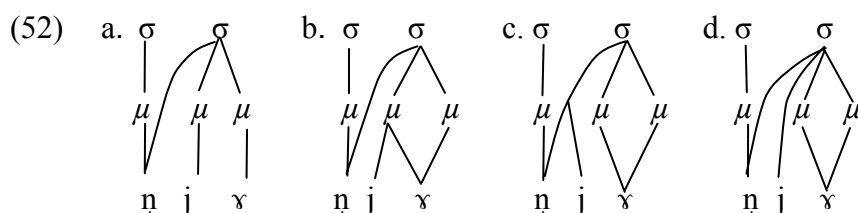
<sup>24</sup> ‘VV’ here refers to either a two-vowel sequence, or a long vowel, or a stressed bimoraic vowel (Duanmu 1999).



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. [lɣ<sup>22</sup>] ‘leak’  
 b. [lɣ<sup>52</sup>] ‘hollow out’  
 c. [mɪɾ<sup>3</sup>] ‘extinguish’  
 d. [mɪɾ<sup>5</sup>] ‘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.<sup>25</sup> 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 [ɲ<sup>33</sup>] for ‘not’ is syllabic only when uttered in the combination [ɲjɣ<sup>35</sup>], which is actually a disyllabic lexical item, meaning ‘not have/haven’t’. A question may arise: is [ɲjɣ<sup>35</sup>] a CGV monosyllabic unit? If not, what is its phonological representation? In fact, the prosodic word [ɲjɣ<sup>35</sup>] deserves a comment. I assume that [ɲ] in [ɲjɣ<sup>35</sup>] is a clitic, which is a toneless syllable by itself and is phonetically and phonologically fused with the host syllable [jɣ]. As a syllabic nasal, [ɲ] is both a lexical syllable by itself and the onset of the second syllable. The moraic structure of [ɲjɣ<sup>35</sup>] can be represented in (at least) four possible ways, given in (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 pre-nuclear glide [j]. It can be assumed that [j] has one mora independently as [ɣ] does, as in

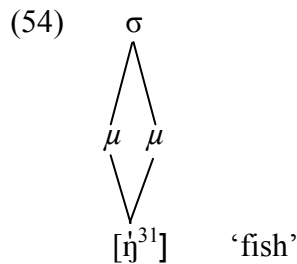
<sup>25</sup> According to the tonal structure of SX, the high register carries high-pitch tones of 52, 35, 33, 5 and the low register carries low-pitch tones of 31, 13, 22 and 3. The details of the tonal system of SX will be discussed in chapter 5.

(52a), suggesting that [j] is in the Nucleus; or [j] is moraic but shares one mora with [ɣ], 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 [ɲjɣ<sup>35</sup>] is like in (52d), which will be discussed in detail in chapter 4. However, there still remain some mysterious questions as to how [ɲ] becomes a clitic; why [ɲ] in [ɲjɣ<sup>35</sup>] is not palatalized as [ɲ̟] according to the nasal palatalization rule in (30); why [ɲjɣ<sup>35</sup>] 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 [ɲ̩] has three different meanings with different tones, representing three different lexical words, as listed in (53):

- (53) a. [ɲ̩<sup>31</sup>]                      ‘fish’  
       b. [ɲ̩<sup>13</sup>]                      ‘five’  
       c. [ɲ̩<sup>22</sup>] ([tẽ<sup>33</sup>~<sup>22</sup>])        ‘a dragon boat festival’

In (53c), the lexical syllable [ɲ̩<sup>22</sup>] 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):



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 [l̥] cannot appear in the same word because SX is a monosyllabic language and has no coda [l]. Like English, the initial [l] and syllabic [l̥] in SX are allophones of the phoneme /l/ because the syllabic [l̥] is only possible when it occurs alone. It is in complementary distribution with the initial [l]. For example:

- (55) [la<sup>52</sup>] ‘pull’  
 [lɔ<sup>13</sup>] ‘old’  
 [lu<sup>22</sup>] ‘road’  
 [ljan<sup>22</sup>] ‘two (people)’  
 [lo<sup>3</sup>] ‘green’  
 [l̥<sup>22</sup>] ‘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  $\alpha$  can be the syllable peak iff  $\alpha$  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).

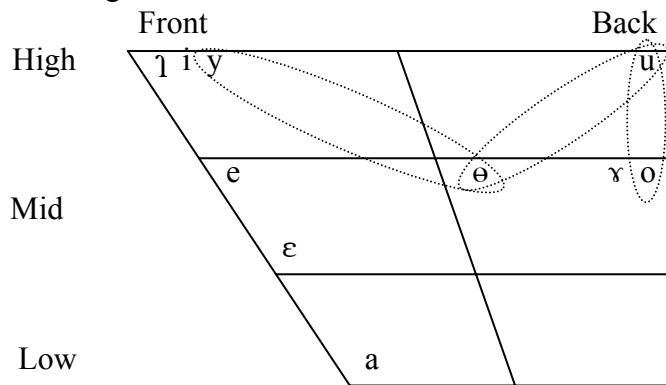
### 2.3.3 Vowels

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<sup>26</sup> 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 [ɿ], which is a very remarkable phone in SX as well as in Mandarin. [ɿ] 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 [ɿ] in Chinese (Karlgren 1915–1926; Chao 1968;

<sup>26</sup> 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 [ɿ], [ə], [a], and [ɔ].

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 ʮ ʮ ʮ].<sup>27</sup> 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 <sup>51</sup> ] <sup>28</sup>	'four'	[sz]
	b. [zɿ <sup>51</sup> ]	'day'	[zz]
	SX		
	c. [sɿ <sup>33</sup> ]	'four'	
	d. [zɿ <sup>22</sup> ]	'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):

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

<sup>27</sup> [ʮ] and [ʮ] are also apical vowels in some Chinese dialects. They are the rounded counterparts of [ɿ] and [ʮ] respectively, differing from the rounded front glide [ʮ].

<sup>28</sup> 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 [ɿ] may be a syllabic fricative. However, he realizes that the filter cannot be correct as stated in (59) because [t] and [t<sup>h</sup>] are also [+cor] and [ti] and [t<sup>h</sup>i] 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<sup>55</sup>] 'close', [pin<sup>55</sup>] 'guest' and [piŋ<sup>55</sup>] 'soldier' while \*[ɿn] or \*[ɿŋ] 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:

(60)	Mandarin		SX	
a.	[t <sup>h</sup> i <sup>55</sup> ]	'shave'	[t <sup>h</sup> i <sup>33</sup> ]	'shave'
	*[t <sup>h</sup> in]		*[t <sup>h</sup> in]	
	[t <sup>h</sup> iŋ <sup>55</sup> ]	'listen'	[t <sup>h</sup> iŋ <sup>33</sup> ] <sup>29</sup>	'listen'
b.	*[si]		*[si]	
	*[sin]		*[sin]	
	*[siŋ]		*[siŋ]	

The examples in (60) show that /i/ and /iŋ/ can occur after /t<sup>h</sup>/, but /in/ cannot, while /i/, /in/ or /iŋ/ cannot occur after /s/<sup>30</sup> in either Mandarin or SX, because of their phonotactics. Wiese claims that the nucleus preceded by /s/ in (58) must be a syllabic fricative [ʒ] because [sʒ] is acceptable, as shown in his argument in (58), while \*[sʒn] or \*[sʒŋ] is not found. This is ill-formed only because of the Sonority Sequencing Principle (SSP).<sup>31</sup> Wiese argues that if the nucleus in (58) were a vowel, [ɿn] or [ɿŋ] 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):

<sup>29</sup> [ɿ] is an allophone of /i/ in SX, which will be discussed in chapter 3.

<sup>30</sup> Here '//' is used instead of '[' ]' for the purpose of indicating that the same underlying phonemes are involved in both Mandarin and SX.

<sup>31</sup> 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 <sup>51</sup> ]	‘mill’	[mo <sup>22</sup> ]	‘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 <sup>31</sup> ]	‘head’	*[dɤŋ]	*[dɤʔ]
	[fɤ <sup>35</sup> ]	‘deny’	*[fɤŋ]	*[fɤʔ]
	[lɤ <sup>22</sup> ]	‘leak’	*[lɤŋ]	*[lɤʔ]
	[tsɤ <sup>35</sup> ]	‘walk’	*[tsɤŋ]	*[tsɤʔ]
	[k <sup>h</sup> ɤ <sup>33</sup> ]	‘button’	*[k <sup>h</sup> ɤŋ]	*[k <sup>h</sup> ɤʔ]

However, there is no reason to doubt that [ɤ] is a vowel. Neither Wiese’s filter in (59) nor his argument for \*[sɤn] and \*[sɤŋ] can support his denying [ɤ] the status of a vowel. If, as Wiese assumes, [ɤ] is a syllabic fricative identical in place and continuancy, it must be an allophone. Then what is the underlying segment— a phonemic /z/ or /ʒ/? If a syllabic fricative itself is a distinctive phoneme, what is its phonological property as an underlying phoneme? The facts suggest the contrary, viz. [ɤ] is a vowel and phonologically is in complementary distribution with /i/ in both SX and Mandarin. Phonetically, [ɤ] 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),<sup>32</sup> among which the formants of [i] and [ɤ] pronounced by male and female are shown in the following table:

---

<sup>32</sup> 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)	
	BJ	SH	BJ	SH	BJ	SH <sup>33</sup>
[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<sup>35</sup>] ‘paper’  
 [dzɿ<sup>22</sup>] ‘late’  
 [ts<sup>h</sup>ɿ<sup>33</sup>] ‘wing’  
 [zɿ<sup>22</sup>] ‘word’  
 [sɿ<sup>33</sup>] ‘try’

Following Bright (1978), Ladefoged and Maddieson (1996: 163) regard “the dental sibilant as being apical”.<sup>34</sup> Presumably, the reason that [ɿ] only occurs after [ts ts<sup>h</sup> dz s z] in SX is that the apical vowel [ɿ] shares with the apical sibilants the [+apical] feature,<sup>35</sup> 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

<sup>33</sup> BJ refers to Beijing speakers who speak SC; SH refers to Shanghai (Wu) speakers who speak ASH.

<sup>34</sup> 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<sup>h</sup> dz s z].

<sup>35</sup> 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):

$$(65) \quad /i/ \rightarrow [ɿ] / \left[ \begin{array}{l} +\text{cons} \\ +\text{apical} \end{array} \right] \_$$

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 <sup>22</sup> ]	‘earth’
	[du <sup>22</sup> ]	‘ferry’
	[de <sup>22</sup> ]	‘pocket’
	[dɤ <sup>22</sup> ]	‘bean’
	[do <sup>22</sup> ]	‘big’
	[da <sup>22</sup> ]	‘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 [ʃ] and [ʒ], leaving many systematic impossibilities. Some examples are presented in (67):

(67)	[ly <sup>31</sup> ]	‘donkey’
	[tɕy <sup>33</sup> ]	‘expensive’
	[tɕ <sup>h</sup> y <sup>35</sup> ]	‘take’
	[dzy <sup>22</sup> ]	‘live’
	[ɕy <sup>52</sup> ]	‘book’
	[ʒy <sup>22</sup> ]	‘tree’
	[ʒy <sup>33</sup> ]	‘silted’
	[fiy <sup>13</sup> ]	‘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. [ẽ̃], [ẽ̃] and [õ̃]. Consider the following examples:

- (68) [k<sup>h</sup>ẽ̃<sup>33</sup>] ‘look at/watch’  
       [pẽ̃<sup>33</sup>] ‘half’  
       [pẽ̃<sup>52</sup>] ‘class’

Of the three examples in (68), only [ẽ̃] has an oral counterpart [e], a phonemic vowel that can appear alone as the rhyme. The other two vowels, [ẽ̃] and [õ̃], 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 <sup>35</sup> ]	[nẽ <sup>31</sup> ]	‘difficult’
	[hwan <sup>35</sup> ]	[vwẽ <sup>31</sup> ]	‘return’
	[t <sup>h</sup> wan <sup>55</sup> ]	[t <sup>h</sup> ẽ <sup>33</sup> ]	‘swallow’
	[k <sup>h</sup> un <sup>51</sup> ]	[k <sup>h</sup> wẽ <sup>33</sup> ]	‘sleepy’
	[pen <sup>35</sup> ]	[bẽ <sup>22</sup> ]	‘pail’
	[kan <sup>35</sup> ]	[kẽ <sup>35</sup> ]	‘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):

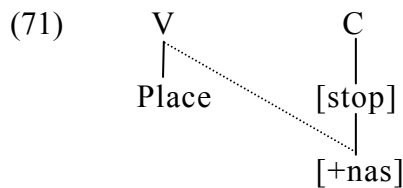
(70)	Monosyllable		Disyllable		
	a. [ts <sup>h</sup> oŋ <sup>52</sup> ]	‘dash’	[ts <sup>h</sup> oŋ.a] <sup>36</sup>	→	[ts <sup>h</sup> oŋ <sup>52</sup> .ŋa] ‘Dash!’
	b. [nẽ <sup>31</sup> ]	‘difficult’	[nẽ.a]	→	[nẽ <sup>31</sup> .na] ‘Difficult!’
	c. [t <sup>h</sup> jẽ <sup>52</sup> ]	‘heaven’	[t <sup>h</sup> jẽ.a]	→	[t <sup>h</sup> jẽ <sup>52</sup> .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<sup>h</sup>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

<sup>36</sup> 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<sup>h</sup>oŋ.a] is uttered with a geminated velar nasal as [ts<sup>h</sup>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 [nẽ] in (70b) and [t<sup>h</sup>jẽ] 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):



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

(72)  $V \rightarrow \tilde{V} / \_ \_ N/$

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  $[k^h \tilde{e}^{33}]$  ‘watch/look at’ is  $/k^h V n^{33}/$  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], [jê], [jễ], [wễ], [wễ̃], and [ɥễ], 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 [ễ], 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 [ễ], with the exception of back rounded [o]. This phenomenon suggests a rule or an equivalent constraint, as in (73) and (74), respectively:

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

or

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

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 \*[jə], permitting the existence of [jo] in SX. This roundedness rule is well supported by the existence of [ɥẽ] 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 [ɥ ẽ] exists in SX. Undoubtedly, [ɥ] is an allophone of the underlying glide /j/ in [ɥẽ]. 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 <sup>33</sup> ]	‘blood’	[fɥwe <sup>31</sup> ]	‘return’
	[ɕja <sup>33</sup> ]	‘write’	[fɥwo <sup>22</sup> ]	‘speech’
	[ɲɥẽ <sup>13</sup> ]	‘soft’	[gwẽ <sup>22</sup> ]	‘circle’
	[tɕ <sup>h</sup> jɿ <sup>33</sup> ]	‘autumn’	[hwẽ <sup>33</sup> ]	‘happy’
	[pjẽ <sup>52</sup> ]	‘edge’	[fɥwa <sup>22</sup> ]	‘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 pre-nuclear glides has been the topic of many Chinese linguistic studies. I claim that the pre-nuclear 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],<sup>37</sup> 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 pre-nuclear 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 <sup>55</sup> ]	'family'	[ma <sup>55</sup> ]	'mother'
		[kwa <sup>55</sup> ]	'melon'	[la <sup>55</sup> ]	'pull'

---

<sup>37</sup> 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 <sup>35</sup> ]	‘chest’	[ts <sup>h</sup> ai <sup>35</sup> ]	‘finance’
	[kwai <sup>55</sup> ]	‘well-behaved’	[k <sup>h</sup> ai <sup>55</sup> ]	‘open’
c.	[t <sup>h</sup> jau <sup>55</sup> ]	‘choose’	[t <sup>h</sup> au <sup>55</sup> ]	‘hollow out’
	[p <sup>h</sup> jau <sup>55</sup> ]	‘flow’	[p <sup>h</sup> au <sup>55</sup> ]	‘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.<sup>38</sup> 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.<sup>39</sup> 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 [ɔ̃]. It is very frequently found in syllables with an initial consonant, such as the following:

(77)	[sɔ̃ <sup>35</sup> ]	‘little/few’
	[pɔ̃ <sup>52</sup> ]	‘wrap’
	[ɕjɔ̃ <sup>35</sup> ]	‘small’
	[dzjɔ̃ <sup>31</sup> ]	‘bridge’
	[ŋɔ̃ <sup>31</sup> ]	‘tolerate’

<sup>38</sup> This is the Chinese poetic rhyming pattern.

<sup>39</sup> The position of prenuclear glides in the syllable structure will be discussed in more detail in chapter 4.

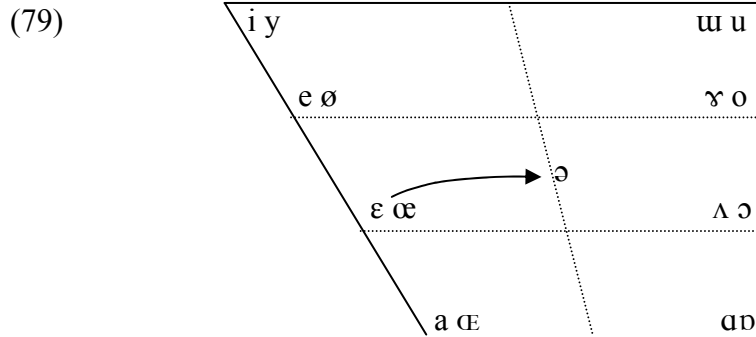


VV combinations as the nucleus of a syllable can be found in many other languages such as [ɛə] in English (Heffner 1949) and [ɔə] and [eə] in German (Rohler 1999, from International Phonetic Association, Corporate Author International Phonetic 1999). In terms of diphthongs, one element of the combination 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], [ɔə] and [eə] 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], [ɔə] and [eə] as diphthongs challenges Trask's diphthong definition (ii). Moreover, whether [ə] in SX is a diphthong remains a question.

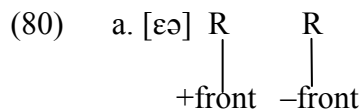
Before answering this question, I first introduce a well-attested principle: the Obligatory Contour Principle (Leben 1973; Goldsmith 1976; McCarthy 1986; Yip 1988; Wiese 1997):

- (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):



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, [æɔ] in SX is composed of [a] and [ɔ], which are identical in height and position and only differ in [ $\pm$ 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 /æ/ 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, /æ/ 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 /æ/ and /ao/ as diphthongs will challenge Trask’s diphthong criterion (iii). With the [æɔ] 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:



- b. [ai] R      R  
           |      |  
       -high +high
- c. \*[iy] R      R  
           |      |  
       +front +front  
       (or +high +high)
- d. \*[ɔɐ] 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 [ɔɐ] is not a diphthong in SX, but a true VV combination. What matters whether [ɔɐ] is a diphthong or not is a matter of a vowel system of SX. The exclusion of [ɔɐ] 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], [ɔŋ], [oŋ], [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 <sup>31</sup> ]  | ‘able’   | [fəʔ <sup>3</sup> ] | ‘join’  |
|      | [tɕɪŋ <sup>52</sup> ] | ‘gold’   | [zɪʔ <sup>3</sup> ] | ‘enter’ |
|      | [ŋaŋ <sup>22</sup> ]  | ‘hard’   | [baʔ <sup>3</sup> ] | ‘white’ |
|      | [zɔŋ <sup>31</sup> ]  | ‘taste’  | [pɛʔ <sup>5</sup> ] | ‘eight’ |
|      | [loŋ <sup>31</sup> ]  | ‘dragan’ | [moʔ <sup>3</sup> ] | ‘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 [ɔ] in [ɔŋ] 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], [wɔŋ], [joŋ], [woŋ], [woʔ], [joʔ], [jaʔ], [weʔ]. 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 <sup>52</sup> ]	‘scent’	[k <sup>h</sup> woʔ <sup>5</sup> ]	‘wide’
	[vwaŋ <sup>31</sup> ]	‘horizontal’	[tɕ <sup>h</sup> joʔ <sup>5</sup> ]	‘lack’
	[kwɔŋ <sup>52</sup> ]	‘light’	[tɕjaʔ <sup>5</sup> ]	‘foot’
	[dʒjoŋ <sup>31</sup> ]	‘poor’	[vwaʔ <sup>3</sup> ]	‘slide’
	[fiwoŋ <sup>31</sup> ]	‘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], \*[qi], \*[qu], \*[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], [qu] 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 <sup>35</sup> ]	[tɕjɤ <sup>33</sup> ]	‘nine’
[ji <sup>35</sup> ]	[hi <sup>31</sup> ]	‘move’
[qu <sup>214</sup> ]	[hiy <sup>13</sup> ]	‘rain’
[k <sup>h</sup> wi <sup>55</sup> ]	[k <sup>h</sup> we <sup>52</sup> ]	‘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, [jɔɐ], which has remained undiscussed so far. In fact, [ɔɐ] in SX behaves like a long vowel phonologically, though it is a combination of [ɑ] and [ɔ] phonetically, as discussed above, so that [jɔɐ] 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<sup>5</sup>] ‘duck’  
       [ʔɪʔ<sup>5</sup>] ‘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<sup>5</sup>] ‘supervise’                      c. [baʔ<sup>3</sup>] ‘white’  
       b. [to<sup>52</sup>] ‘many’                            d. [ba<sup>31</sup>] ‘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,<sup>40</sup> 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<sup>41</sup> 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.

---

<sup>40</sup> 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.

<sup>41</sup> 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			ə	ɤ	o	Mid
			ɛ				
	ɛ						
Open							Low
	a				ɑ	ɒ	
	⊗round	+round	⊗roun d	+round	⊗roun d	+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.