The phonology of Shaoxing Chinese
Zhang, J.

Citation

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/4279

4   The Syllable Structure of Shaoxing

4.1 Introduction

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

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

---

[1] Hyman (1983, 1990, 2003) argues that the syllable is not universal, based on examples of extreme vowel collocation in Gokana, e.g. [mɛ́ɛ́ ɛ̀ kɔmm̀ kɛ̃̂ɛ̃̀ɛ̃̀ɛ̃̀ɛ̃́] 'who said I woke him up?'. Even so, this does not necessarily mean that there is no evidence for the syllable in Gokana.
[2] The context {C, #} means that some phonological rules take effect only when the relevant segment occurs before a consonant or at the end of the word. In both cases the target occurs at the end of the syllable. In SX, a monosyllabic language, every word orthographically presented in a Chinese character is a syllable.

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

$$[-son] \;\rightarrow\; [-voice] \;/\; \underline{\hspace{1em}} \left\{ \begin{matrix} C \\ \# \end{matrix} \right\}$$

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

$$[+lateral] \;\rightarrow\; \begin{bmatrix} +high \\ +back \end{bmatrix} \;/\; \underline{\hspace{1em}} \left\{ \begin{matrix} C \\ \# \end{matrix} \right\}$$

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

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

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

$$\emptyset \;\rightarrow\; V \;/\; \begin{bmatrix} C \\ \alpha Place \end{bmatrix} \underline{\hspace{1em}} \begin{bmatrix} C \\ \beta Place \end{bmatrix} \#$$

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

$$\begin{bmatrix} C \\ \alpha Feature \end{bmatrix} \;\rightarrow\; \emptyset \;/\; \begin{bmatrix} C \\ \beta Feature \end{bmatrix} \underline{\hspace{1em}} \#$$

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

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

---

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

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

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

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

---

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

4.2 The Syllable Types of Shaoxing

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

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

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

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

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

---

[5] Spoken in Northern Côte d'Ivoire and Mali.
[6] A Khoisan language of Botswana that is most closely related to !Xóõ in Africa.

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

---

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

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

        Mandarin
        d. [tjan # **an** # məŋ]   'Heaven **Peace** Gate'
        e. [tjan # **nan** # məŋ]   'Heaven **South** Gate'

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

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

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

(5)   [tsʰoŋ^{52}a]   →   [tsʰoŋ^{52}ŋa]   'dash ahead'
      [ɲɪŋ^{31}a]   →   [ɲɪŋ^{31}ŋa]   '(how many) people!'

The examples in SX in (5) both form a disyllabic unit, in which the second (toneless) syllable does not have real lexical meaning and only serves to emphasize the action expressed by the first syllable. Thus, the toneless vowel-initial syllable does not require a phonetic onset for the consonant-tone correlation, allowing liaison to take place. Yip (1980: 191) refers to such cases of liaison as 'resyllabification' and argues that resyllabification takes place when the first syllable ends in a consonant whilst the second

starts in a vowel. She cites an example in Mandarin from Chao (1968: 803), which demonstrates the absence of a phonetic onset of toneless syllables of the prosodic word in the same way, as shown in (6):

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

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

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

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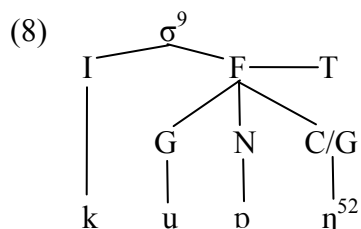
---

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

(7) a. C   /ŋ̩³¹/   'fish'
     b. V   /a³⁵/   'short'
     c. GV   /jɤ⁵²/   'low (voice)'
     d. VC   /ɿʔ⁵/   'one'
     e. GVC   /jaŋ⁵²/   'seedling'
     f. GVV   /jɑɒ³³/   'want'

However, a (large) minority of speakers use a glottal stop for all words with a zero initial (Chao 1968). The syllables in (7b-f) are also phonetically transcribed as [ʔa³⁵], [ʔjɤ⁵²], [ʔɿʔ⁵], [ʔjaŋ⁵²], and [ʔjɑɒ³³], respectively, following the glottal stop insertion (GSI) rule in (9) in chapter 2. The initial glottal stop [ʔ] in this case is phonetically present, but does not have a phonological position in the underlying syllable structure, since it is fully predictable. This was also discussed in chapter 2. In summary, there are altogether 11 syllable types in the syllable inventory of SX, as shown in (2) and (7).
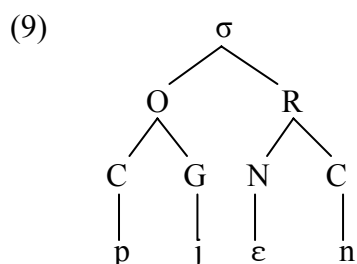
4.3 The Internal Syllable Structure

While all Chinese languages and dialects have different inventories and arrangements of phonological units, these languages nevertheless show a remarkable similarity with respect to syllable organization. SX, as one of the Chinese dialects, has much in common in its internal syllable structure with other Chinese dialects. It was claimed in traditional Chinese phonology that all Chinese dialects have a syllable structure maximally consisting of an Initial and a Final with a prenuclear glide, a nucleus vowel, a final which can be an off-glide or a consonant, and a tone (Chao 1968; Wang 1985; Norman 1988; among others), as shown in (8) with the example of [kuɒŋ⁵²] 'light' in SX:

(8)

```
                σ⁹
         ┌──────┴───┐
         I          F────────T
         │       ┌──┼──┐
         │       G  N  C/G
         │       │  │   │
         k       u  ɒ   ŋ⁵²
```

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

4.3.1   Onset clusters

Some scholars (e.g. Tung 1983; Yin 1989; Bao 1990) argue in favor of an OR analysis of Chinese syllable structure and assume that the prenuclear glide in Chinese syllable structure is the second member of the onset cluster, claiming, therefore, that there is a possible onset complex in Chinese. Taking [pjɛn⁵⁵] 'edge' (Mandarin) as an example, Bao (1990) represents the syllable structure as in (9) (in the following analysis and throughout this chapter, 'C' is used for consonant in both syllable-initial and syllable-final position):

(9)

```
                σ
          ┌─────┴─────┐
          O           R
        ┌─┴─┐       ┌─┴─┐
        C   G       N   C
        │   │       │   │
        p   j       ɛ   n
```

---

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

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

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

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

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

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

(11)   May-ku:

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

---

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

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

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

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

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

(12) La-mi:
  a. hen → len hin 'to turn pages'
  b. keʔ → le kiʔ 'to separate'
  c. kya → lya ki 'slope'
  d. kway → lway ki 'strange'

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

---

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

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

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

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

(14)   [ɑʊ] game in SX:
       a. pjẽ^{31} →   pɑʊ^{31} pjẽ^{31} 'edge'
       b. djɑʊ^{22} →   dɑʊ^{22} djɑʊ^{22} 'exchange'
       c. kʰwoʔ^{5} →   kʰɑʊ^{5} kʰwoʔ^{5} 'wide'

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

---

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

with the host syllable. For example, the negative forms [vəʔ³vɛ̃³¹] 'don't be naughty' and [vəʔ³ʔjɑɒ³³] 'do not want to' become [fɛ̃⁵²] and [fjɑɒ³³], respectively, after a series of phonological changes (see §5.6 in chapter 5 and Zhang (2005: 69-79) for more information). The phonological changes in these two forms of SX clitics are illustrated in (15):

(15)   a   b   c
 [vəʔ³vɛ̃³¹]   →   [v(əʔ³)+(v)ɛ̃ ³¹]   →   [vɛ̃⁵²]   →   [fɛ̃⁵²]
 [vəʔ³ʔjɑɒ³³]   →   [v(əʔ³)+(ʔ)jɑɒ³³]   →   [vjɑɒ³³]   →   [fjɑɒ³³]

The changes in (15) show that (a) the rhyme of the clitic syllables and the onset of the host syllables are both deleted; (b) the onset of the clitic syllables becomes the onset of the host syllables and the low-pitch tone becomes a high-pitch tone when fused into one syllable (which is made possible only by a change of the register feature. The change or spreading of tonal features will be discussed in detail in chapter 5); (c) the voiced initial obstruent becomes voiceless because of the high-pitched tone (according to the onset-condition constraints in (26) in ch. 2).[13] The original host verb [ʔjɑɒ³³] 'want' in (15) contains a prenuclear glide [j]. When the onset is deleted during the process of the syllable fusion, [j] remains in the rhyme. The syllable fusion of clitics in SX is a window on syllabic constituency in SX, and strongly suggests that the prenuclear glide is not in the Initial but in the Final.

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

---

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

[14] In modern SX, the poetic rhyming system does not include the prenuclear glide, e.g. [dzjɑɒ³¹] 'bridge' and [dɑɒ³¹] 'flea' rhyme perfectly.

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

(16)

Input /brændi/	*COMPLEX(ONS)	MAX-IO	DEP-IO
a. blẽdi	*!		
b. baʔdi		*!	
c. lẽdi		*!	
d. ☞ baʔlẽdi			**

Another example, which provides evidence with regard to the status of glides, is the loanword *Twain* from English as in *Mark Twain*, in which [tw-] is a complex onset in English. Both in SX and in Mandarin, [tw] is acceptable in principle, e.g. Mandarin [tʰwan³⁵] 'unite' and SX [tʰwə̃⁵²] 'swallow'. However, the status of [w] in SX and Mandarin is different from the status of [w] in English. In the former case, [w], as the pre-nuclear glide, is not in the onset while in the latter case, [w] is in an onset cluster. Thus, the optimal realization of the loanword *Twain* in Mandarin is [tʰu.wen] rather than [tʰwan] or [tʰwen] because of the *COMPLEX(ONS) constraint in Chinese. To capture this behaviour, I propose the following constraint, which preserves onset status:

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

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

(18)

Input /twein/	*COMPLEX (ONS)	IDENT-IO(ONS)	MAX-IO	DEP-IO
a. tw·en¹⁵	*!		*	
b. t·wen		*!	*	
c. t·wi.en		*!		
d. t·u.en			**!	*
e. ☞ t·u.w·en			*	*

---

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

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

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

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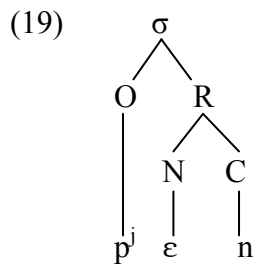
---

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

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

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

4.3.2 Secondary articulation

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

(19)

$$
\begin{array}{c}
\sigma \\
\diagup\diagdown \\
\text{O} \quad \text{R} \\
| \quad \diagup\diagdown \\
\quad \text{N} \quad \text{C} \\
| \quad | \quad | \\
\text{p}^{\text{j}} \quad \varepsilon \quad \text{n}
\end{array}
$$

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

(a) Howie's (1976) phonetic experiment shows that the prenuclear glide is not a TBU.[19]

(b) The poetic rhyming unit does not include the prenuclear glide.

(c) The Mandarin syllable maximally has CVX; i.e. only three segment slots.

(d) Syllables in Chinese have an obligatory onset. When there is no onset consonant before the nucleus vowel there is an obligatory zero onset; however, prenuclear glides are never preceded by a zero onset.

(e) Phonetically, in syllables such as [swei] 'age' in Mandarin, [s] sounds like [s$^\text{w}$], different from [s] in [swei] 'sway' in English.

---

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

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

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

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

---

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

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

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

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

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

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

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

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

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

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

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

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

The examples in (21) show that the phonetic onset (if there is one) does not occur before [w] or [j] whilst it does occur before a vowel-initial syllable. Duanmu first made the claim that the prenuclear G is in the onset. Then he assumes that CG forms one single sound, for G is only a secondary articulation. On such an assumption, he explains that the reason why the zero onset cannot be used with a prenuclear glide is that the zero onset is not an independent phoneme and cannot have a secondary articulation. In Duanmu's view, the zero onset should also occur before a prenuclear glide if this is not a secondary articulation. On Duanmu's secondary

articulation assumption, a CGV(C) syllable is possible only because G is a secondary articulation depending on the preceding C and *ʔG, *ŋG, *ɣG, or *ɦG is not possible, as shown in (21), because G, as a secondary articulation, cannot be realized in a zero onset. Duanmu may not have realized the fact that [w] and [j] in (21b) and (21c) are the onset by themselves so that it is clear enough that the zero onset never occurs in the case of (21b) and (21c) because the onset is not empty and as it is, zero onset never occurs before C (e.g. *ʔC). We should not confuse the prenuclear glides and the onset glides when talking about the Chinese syllable structure. The former always refers to G between C and V in CGV, and the latter is G in GV unless G cannot be in the onset of a language in question. However, the glides in Mandarin can be in the onset (Wang 1963, 2003), which is strongly supported by the existence of such syllables as [wu] and [ji] in Mandarin like many other languages. But the glides in SX cannot be in the onset and there is no such syllable as *[wu] or *[ji] in SX, because [(C)wu] and [(C)ji] badly violate OCP(H) (see (83), ch.2) when both segments are in the Final constituent. It is also true that the glides [w] and [j] can be preceded by the phonetic onset [ʔ] in SX when there is no other consonant in the onset, as exemplified in (22):

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

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

Finally, Duanmu (1990, 1999) contrasted the secondary articulation in syllables such as [swan^{55}] 'sour' and [swei^{51}] 'age' in Mandarin, in which [s] sounds like labialised [s^w], with words like [swɒn] 'swan' and [swei] 'sway' in English, in which [s] does sound as labialized as [s] in the Mandarin [swan^{55}] and [swei^{51}], and he argues that the difference in pronunciation corresponds to a difference in syllabic affiliation: secondary articulation in Mandarin, cluster in English. It is true that some lip rounding is expected to assimilate in words like [s^{w}wan^{55}] and [s^{w}wei^{51}] mentioned above, but we prefer to put this down to some form of fast-

speech assimilation; at any rate it does not tell us anything about the underlying syllabic status of the prenuclear glide.

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

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

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

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

---

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

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

(24)   a. Velarization in Russian:
           [dal]   'distance'   [dalˠ]   'gave'
           [bi:]   'be(imp.)'   [bˠi:]   'yellow'

        b. Laryngealization in Danish:
           [vɛn]   'friend'   [vɛn̰]   'turn'
           ['hɛnʌ]   'happens'   ['hɛn̰ʌ]   'hands'

        c. Pharyngealization in Qatari Arabic:
           [χal]   'to leave'   [χalˤ]   'vinegar'
           [sad]   'to prevail'   [sˤad]   'name of letter'

        d. Labialization in Tabassaran:
           [naq]   'yesterday'   [naqʷ]   'grave'

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

(25)   [nʲjou³⁵]   'cow'
        [nʷwo³⁵]   'move'
        [lan³⁵]   'blue'
        but
        *[_nʲ] #
        *[_nʷ] #

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

4.3.3   The rhyme constituent
The previous analysis shows that the prenuclear glide in SX is not in the onset – neither in the onset cluster, nor as a secondary articulation –,

suggesting that it may well be in the rhyme. Some authors (e.g. Wang & Chang 2001, among others) argue that the prenuclear glide is in the rhyme constituent in the Chinese syllable, as illustrated in the following syllable structure:

(26)
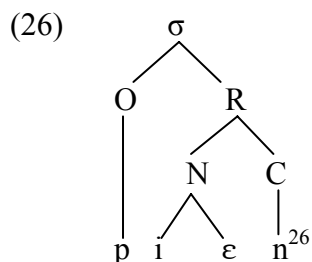
This argument is mainly based on the OR syllable model and the fact that the prenuclear glide does not belong to the onset. Since the prosodic hierarchy is usually assumed to demand exhaustive parsing of entities into the immediately higher constituent, this would mean that segments should all be parsed into Onset or Rhyme (Yip 2003). Thus, if a segment does not belong to the onset, it must be in the rhyme. Based on the traditional Chinese phonological system which treated the prenuclear glide as the rhyme part of the syllable (as *fanqie* did, which was discussed in the previous subsection), Wang & Chang (2001) performed an experiment in which Chinese subjects without any phonological knowledge were asked to fuse two syllables into one, to see which one is preferred, as shown in (27):

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

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

(28) pjao →   pjan + tjao / pan + tjao

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

---

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

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

In one sense, I follow Duanmu's (1990, 1999) and Wang's (1999) argument that the prenuclear glides are not counted in the Chinese poetic rhyming system. This is uncontroversial in the literature of Chinese, including SX. For example, [tɕjan⁵⁵] 'wicked' and [ʂan⁵⁵] 'mountain' rhyme, and [tɕʰjau⁵⁵] 'knock' and [tau⁵⁵] 'knife' rhyme in Mandarin. As a matter of fact, these syllables rhyme with each other not because the prenuclear glides are excluded from the rhyming system but because in none of these syllables was there a prenuclear glide in Middle Chinese. According to the *fanqie* system in *Guangyun* (1008), the syllables for *wicked* and *mountain* were [kan] and [san], respectively; and those for *knock* and *knife* were [kau] and [tɑu], respectively (Wang 2003), just as in modern SX, the syllables for *wicked* and *mountain* are [kɛ̃⁵²] and [sɛ̃⁵²]²⁷, respectively, and the syllables for *knock* and *knife* are [kɑɒ⁵²] and [tɑɒ⁵²], respectively, both pairs forming perfect rhymes.

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

---

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

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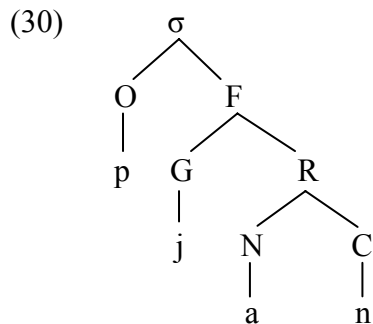
4.3.4   Head of the Final

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

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

(29)

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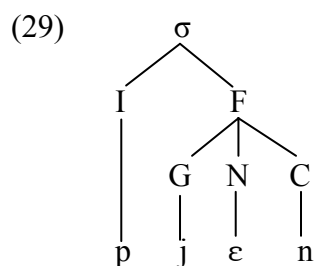
The syllable structure in (29) reflects the traditional Chinese phonology view on syllable structure in Middle Chinese. However, this traditional syllable structure becomes problematic in the face of the modern Chinese dialects in that (a) the Final is not the rhyming unit in poetic writing; (b) the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

(30)

```
                    σ
                 ╱     ╲
              O           F
              │         ╱   ╲
              p       G       R
                      │     ╱   ╲
                      j    N     C
                           │     │
                           a     n
```

    Cheng (1966) claims that in the Chinese syllable structure, as shown in (30), G is the head of the Final and outside of R, which is a constituent composed of N and C. The notion that G is head of F remains problematic because heads are always obligatory, while G is optional in Chinese. However, this syllable structure explains how segments are parsed into syllabic constituents under one syllable node in Chinese phonology. In this structure, R is the weight unit of the syllable and also the rhyming unit in literary writing; F is also a constituent structure higher than R, which is a remarkable sub-constituent only postulated, as far as I am aware, for Chinese. The existence of the Final is supported in traditional Chinese phonology, especially on the basis of the *fanqie* system and different language games. It is true that the prenuclear glide is treated differently in different language games or from dialect to dialect, as was discussed above. I assume that this is because the F structure is available in the syllable. In some language games, the onset is replaced so that the

prenuclear glide is left with the rhyme, while in other games the rhyme is replaced so that the glide is left with the onset. On this view, the prenuclear glide is neither in the rhyme nor in the onset. However, structurally speaking, if one were to make a break in the syllable in terms of paradigmatic substitution, it is obvious that the most likely break would be between the Initial and the rest of the syllable, the Final (Walton 1983). The evidence for the F constituent in SX syllables, then, can be summarized as follows:

(a) The prenuclear glides are not in the onset, as suggested by loanword incorporation and language games in SX.

(b) The prenuclear glides are not in the rhyme, which is suggested by the poetic rhyming system in SX and also proven by phonetic experiments (Howie 1976; Lin 1995) which show that medial glides are not part of the TBU in Chinese.

(c) The prenuclear glides and the rhyme are treated as one constituent in the SX language game.

(d) The prenuclear glides are usually treated as the rhyme part in the *fanqie* system of Middle Chinese, many phonetic and phonological characteristics of which are still retained in SX.

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

4.3.5   Other options within OR theory

Although the syllable structure in (30) can express the organization of the Chinese syllable in some ways, there are some problems recognizing this as a general syllable template, as was mentioned above. The complexity of the status of the prenuclear glides is mainly caused by the ambiguous

role they play in the syllable structure, e.g. in language games of different dialects, or speech errors. Taking the OR syllable theory as the framework to work with and based on the data of language games in different Chinese dialects, Yip (2003) proposes two other options without further discussion, both illustrated below: in (31a), prenuclear glides could be in both Onset and Rime, just as 'ambisyllabic' segments have been argued to be part of two syllables simultaneously; or in (31b), prenuclear glides might be in neither Onset nor Rime, associated directly to the syllable node. These two options are formalized in the following tree diagrams, respectively:

(31) a.   σ   b.   σ

           O   R   O   R

                  N   C   N   C

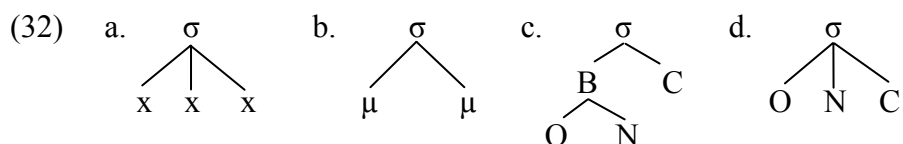   p   j   a   n   p   j   a   n
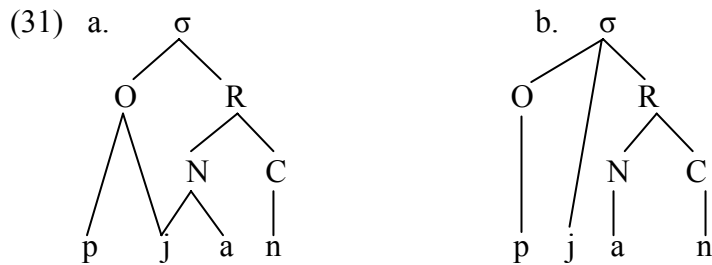
The structure in (31a) shows a representation in which the prenuclear glide is in both Onset and Rime. The structure in (31b) shows a representation in which the prenuclear glide is in neither Onset nor Rime, associated directly to the syllable node. However, both proposals are problematic in one way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

4.3.6   A syntactic approach

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

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

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

(32)   a.   σ   b.   σ   c.   σ   d.   σ
              / | \   /   \   /   \   / | \
             x x x   μ   μ   B   C   O N C
                                                      / \
                                                     O   N

   Apart from these attempts, there are even more approaches to the syllable. Hall (1992) claims that the syllable in German dominates the subsyllabic constituents of onset, nucleus, and coda, without a rhyme. Geudens & Sandra (2001) present an experiment on acquisition of syl-

lable structure and argue that Dutch-speaking pre-reading children between six and seven show no evidence for Onset and Rhyme units in their explicit phonological awareness. Based on the status of the prenuclear glide of Chinese, Yip (2003) argues that evidence for the existence of Onset and Rhyme constituents is scanty and inconsistent, and that the simpler models are sufficient to capture the facts.

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

(33) a. X″

specifier X′

X complement

b. N″ (syllable)

N′ (rhyme)

N⁰ (nucleus)

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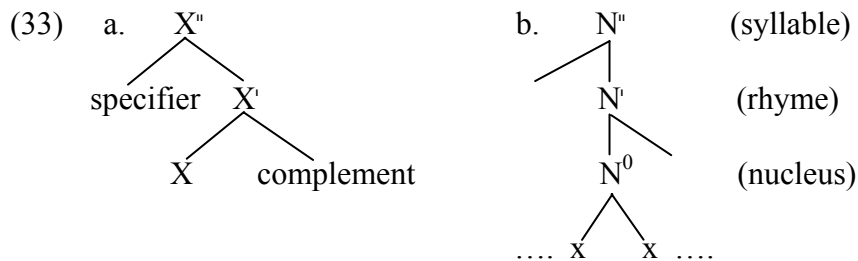
The structure in (33a) is syntactic X-bar structure. Syntactically, specifier means that the grammatical function is fulfilled by certain types of constituent which (in English) precede the head of their containing phrase. For example, in a sentence such as *What did John do*? *what* is the specifier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

The structure in (33b) is an X-bar syllable structure which builds on the idea that the syllable is a headed constituent in which the coda and onset can be treated as the complement and specifier, respectively. In syntactic terms, both complement and specifier are in principle optional (Hornstein 2004: 191) and the head is obligatory, which captures the facts of the syllabic organization in SX. Like the OR model, this model (referred to as the X-bar model) must also stipulate that only material ad-

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

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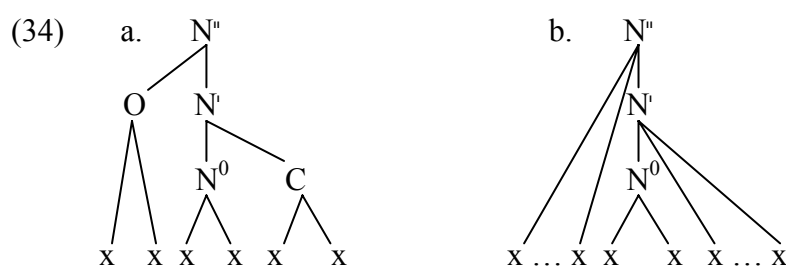
The difference between the two views is related to the complexity of the syllable margins. I will not discuss the difference between these two views. Since SX has no complex onsets or complex codas, the syllabic margins are rather simple, with only one onset consonant and one coda consonant maximally.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

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

---

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

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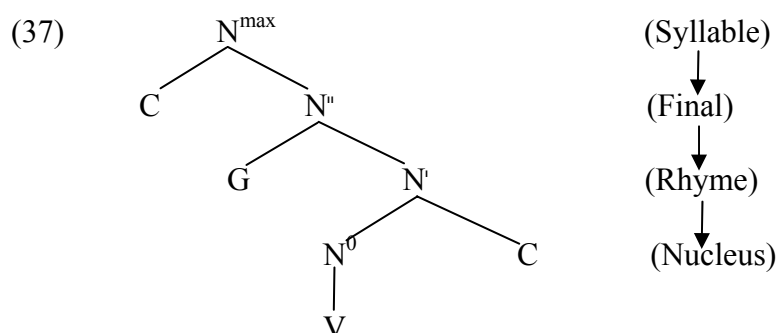
(35) a. N''   b. N''   c. N''   d.   N''
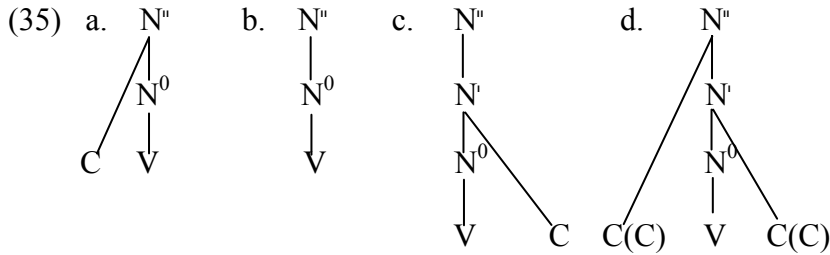
In the X-bar schema, X'' is X^{max} and X^0 is X^{min}. However, both in syntactic structure and in syllabic structure, only X^0 and X^{max} are universally required, as shown in (35). The existence (or non-existence) of an intermediate projection N' is a language-specific matter. Syllabically speaking,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

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

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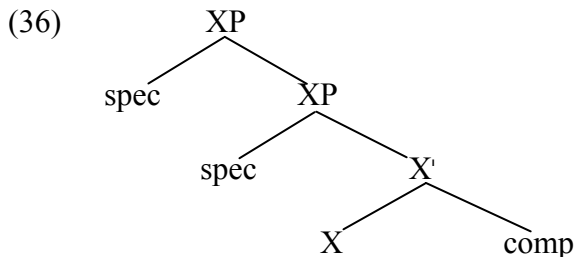
In the multiple-Spec[29] construction in (36), there are two XPs, viz. X^{max} and X", each having a specifier, and only X is the head.[30] From this syntactic X-bar structure, I just borrow its structural framework for syllable organization without exploiting any parameterization of syntax. However,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

(37)

$$
\begin{array}{c}
N^{max} \quad\quad\quad\quad (Syllable) \\
C \quad\quad N'' \quad\quad (Final) \\
G \quad\quad N' \quad\quad (Rhyme) \\
N^0 \quad\quad C \quad\quad (Nucleus) \\
V
\end{array}
$$

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

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

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

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

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

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

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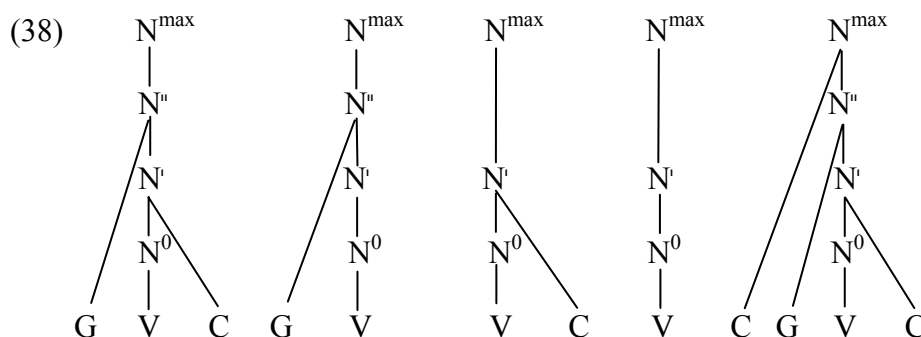
The X-bar syllable structures in (38), all derived from (35a), have the same N^{max} realization as those in (35), which Levin calls N-placement. The structures in (38) show that G cannot link directly to N^{max} even if there is no onset consonant in the syllable (spec of N^{max}) in underlying form,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

4.4 The Coda in Shaoxing

According to the SX syllable structure in (37), there is a coda position, which seems well motivated by the SX data in such examples as in (2) and (7) above. However, there is a strict constraint on the syllable coda in SX in that only a glottal stop or velar nasal is acceptable in the coda (see CODA-COND (18), ch.2). This leads to an extremely asymmetrical picture that in the onset all the 29 consonants in SX can occur while in the coda only [ʔ] or [ŋ] are acceptable. This phenomenon gives rise to the natural question whether it is really necessary to postulate a full coda position in

the SX syllable. In this section, I will discuss the phonetic and phonological properties of the coda consonants in SX. I claim that the glottal stop [ʔ] is the result of debuccalization of syllable final stops and the velar nasal [ŋ] is what has remained as a coda when the other two nasals (bilabial nasal [m] and alveolar nasal [n]) were debuccalized, resulting in the nasalization of the preceding vowels. The debuccalization of syllable final consonants is caused by diachronic attrition, which will be discussed in this subsection.

4.4.1 Previous argumentations for different dialects

Because of the extremely limited number of the coda consonants, some researchers have claimed that there is no real coda in the underlying syllable-final position in Chinese dialects such as Beijing (on which Mandarin is mainly based) (Wang 1993, 1999; Duanmu 1999) and Shanghai (which is one of the Wu dialects, just like SX) (Duanmu 1999). Wang argues that there is no coda in Beijing because the five postnuclear segments in the syllables are [j], [w], [n], [ŋ] and [r], all of which are sonorant. Wang (1999: 226) assumes that the Beijing syllable finals [n], [ŋ] and [r] are actually all glides and vocalic, just like [j] and [w], so that all Beijing syllables are open, without a final C. Wang's syllable structure of Beijing, taking [fan⁵⁵] 'turn over' as an example, is represented as follows in the OR model:

(39)

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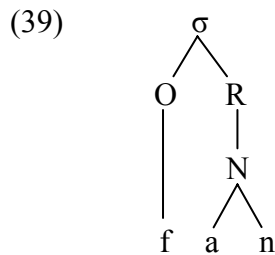
Duanmu (1999) argues that Shanghai has no coda consonants underlyingly and all syllables have a CV structure. Duanmu (1999: 494) assumes that the VN combinations such as [ən in yn oŋ] can be phonemically analyzed as nasalized vowels [ə̃ ĩ ỹ õ] and all [Vʔ] combinations are underlyingly laryngealized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

Chan (1997) proposes that in Fuzhou (one of the Min dialects) the syllable-final glottal stop [ʔ] is placed in the nucleus with a plain vowel preceding, e.g. [pɑʔ^5] 'white', whose syllable structure can be represented in the OR model as in (40):
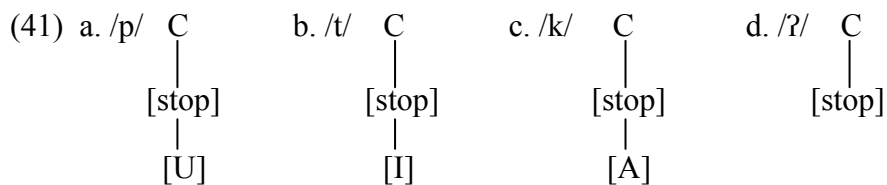
(40)
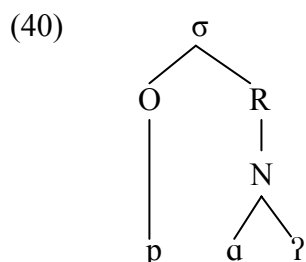
Chan (1997: 279) assumes that [ʔ] in Fuzhou [pɑʔ^5] is disappearing glottal stop and that it occupies the same position as an off-glide in the nucleus, forming an open syllable so that it has the same length as an open syllable ending in a vowel. What Chen means by "disappearing glottal stop" is, as I understand, that the glottal stop is never like a common syllable-final consonantal stop with specific place articulation and it has no robust phonetic realization. In Shanghai and Fuzhou, the syllable-final [ʔ] is assumed to be dropped, for the sake of the syllable length, either resulting in shorter duration, as in Shanghai (Duanmu 1999), or in longer duration, as in Fuzhou (Chan 1997). The motivation of the two proposals is to assume that there is no coda consonant in Mandarin or in Shanghai or in Fuzhou. However, I assume that both the final nasal and the final glottal stop are in coda position in SX, which will be discussed in the following subsections.

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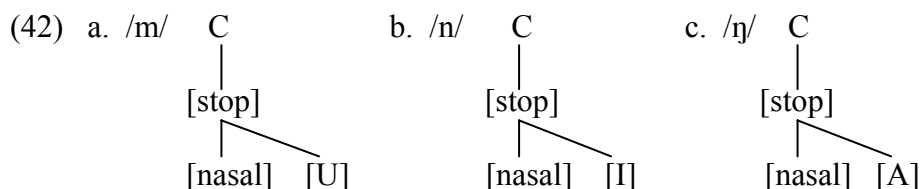
and SX has only /ŋ/ and /ʔ/ in syllable-final position. What happened to the SX syllable-final consonants? I assume that /m/ and /n/ and /p/, /t/ and /k/ have not really disappeared but that they have debuccalized.

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

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

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

---

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

(42) a. /m/   C   b. /n/   C   c. /ŋ/   C
              | | |
           [stop]   [stop]   [stop]
            /\   /\   /\
      [nasal]   [U]   [nasal]   [I]   [nasal]   [A]

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

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

4.4.3   Nasal debuccalization

Debuccalization can be either a diachronic process or a synchronic rule. For example, in some Spanish dialects, [s] debuccalizes into [h] unless it precedes a vowel, e.g. *mismos* [mihmoh] 'similar'. Such a phenomenon of debuccalization is essentially the loss of the supraglottal articulation with retention only of the open glottis gesture (McCarthy 1988: 88). Nasal debuccalization in SX is caused by historical attrition in the same way as

debuccalization of syllable-final stops. Through debuccalization, nasals lost their place compenent, with nasality remaining. Since nasality cannot be realized in a form of segment without articulation component, debuccalized nasals get deleted, resulting in the nasalization of the preceding vowel.

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

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

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

---

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

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

(44)

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

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

(45)

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

The tableau in (45) shows that [ʔ] is the optimal output as the syllable coda in SX for it violates no constraint against place of articulation. In short, in nasal debuccalization in SX, the bilabial nasal may have been lost first and the velar nasal remained, perhaps because velar is the least marked place of articulation among coda consonants (cf. van der Torre 2003) and because the coda position was still required for the sake of the weight of the rhyme in SX. As a result of debuccalization, SX has only two consonant segments, /ŋ/ and /ʔ/, as potential syllable codas. SX underwent debuccalization of all three coda stops, which became realised as /ʔ/. As for the nasals, only /m/ and /n/ debuccalized, resulting in

nasalization of the preceding vowel, while /ŋ/ remained in the coda, because the underlying nasal(s) in the vowel nasalization in SX cannot be /ŋ/, but usually /n/, which can be supported by the nasal restoration triggered by liaison (which was discussed in chapter 3), in which the nasal is restored from the nasal vowel when followed by a vowel and the restored nasal becomes the onset of the following syllable. Since the nasality in nasal vowels receives the default Place specification, namely, coronal, the restored nasal is usually [n] in surface representation, as shown in (46):

(46)   Monosyllables: Disyllables:

    b. [nɛ̃^{31}] 'difficult'   [nɛ̃^{31}a]   → [nɛ̃^{31}na]   'Difficult!'

    c. [t^{h}jɛ̃^{52}] 'heaven'   [t^{h}jɛ̃^{52}a] → [t^{h}jɛ̃^{52}na] 'God!'

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

(47) [ts^{h}oŋ^{52}a]   → [ts^{h}oŋ^{52}ŋa] 'dash ahead'

    [ɲɪŋ^{31}a]   → [ɲɪŋ^{31}ŋa]   '(how many) people!'

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

---

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

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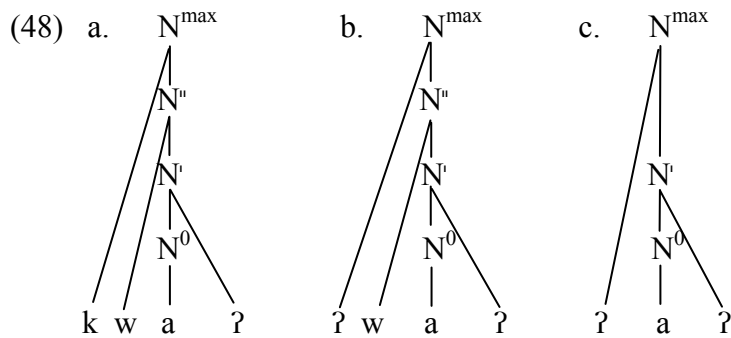

4.5 Syllable Weight

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

4.5.1   Weight-irrelevance of CG

As was argued in chapter 1, the advantage of the mora model in analyzing syllable structure is to offer a clear-cut picture of the weight of syllables. However, the X-bar syllable structures as shown in (38) also explicitly explains the weight relevance or weight irrelevance of syllabic constituents or non-constituent elements. As was discussed previously in this chapter, the SX syllable structure has three internal constituents: Final, Rhyme and Nucleus, represented as $[N^{max}[N''[N'[N^0]]]]$ in X-bar framework, within which the onset is specifier of N^{max} (Syllable), the prenuclear glide is specifier of N'' (Final), and the coda is complement of N' (Rhyme). The syllable is an N-centered structure in the X-bar framework so that N is an obligatory weight contributor to the syllable. Most phonologists hold that syllables are maximally bimoraic (Hyman 1985; Davis 1999, among others), although there are some arguments for trimoraic structure over superheavy rhymes which, however, do not occur in SX. Weight-by-stress ((47), ch.2) demands that stressed syllables are heavy and bimoraic and unstressed syllables are light and monomoraic for all the syllable types in SX. Some researchers (e.g. Duanmu 1999) claim that heavy syllables must be CVV or CVC and light syllables are CV,

assuming that the syllable weight determines the syllable types or the reverse. Moreover, it is well known that the onset is weight-irrelevant in the syllables of the world's languages (McCarthy & Prince 1986; Hayes 1989; Goedemans 1998, among others). Some phonetic experiments (Howie 1976; Lin 1995) show that the prenuclear glides in Chinese are also weight-irrelevant. Thus, as for a CGVC syllable in SX, both C and G before V are non-moraic, i.e. a specifier is weight-irrelevant in syllable structure. Accordingly, such syllables as [kwaʔ⁵] 'scrub', [ʔwaʔ⁵] 'dig' and [ʔaʔ⁵] 'duck' in SX are of the same weight. Their syllable structures can be presented as in (48):

(48) a.

N^{max}
N''
N'
N^0
k w a ʔ

b.

N^{max}
N''
N'
N^0
ʔ w a ʔ

c.

N^{max}
N'
N^0
ʔ a ʔ

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

4.5.2   The weight domain in SX

The weight domain is the constituent that directly dominates the segments contributing to weight. The elements of the weight domain can be different from language to language. For example, a glide before the peak vowel can be in the weight domain if it is vocalic, like [j] in English [tjuːn] 'tune'; whereas [j] in SX [tjɤ⁵²] 'lose' is not in the weight domain, as discussed above. Consonants can also contribute to weight according to

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

(49)

$$
\begin{array}{ccc}
N' & \quad & N' \\
\mid & & \mid \\
N^0 & & N^0 \\
\mid \quad \diagdown & & \mid \quad \diagdown \\
V \quad C & & \mu \quad \mu
\end{array}
$$
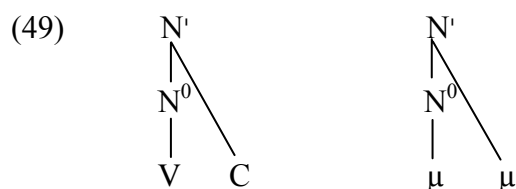
The structure in (49) shows that the syllable final C is moraic in SX. Duanmu (1999) assumes that all syllables of Shanghai Chinese[36] are weak syllables because (i) Shanghai has no diphthongs and vowels are not distinct for length, like SX. (ii) if Shanghai has heavy syllables, there are no

---

[34] Some assume that in some languages like Pirahã, the onset can also be moraic (see Davis 1985; Everett 1988; Topintzi 2004). Cf. Goedemans (1998) for discussion.

[35] Hayes (1989) assumes that an obstruent in the coda is usually not moraic, but if it is in the position of ambisyllabicity it is moraic.

[36] Shanghai is also one of the Wu dialects like SX and its syllable structure is very similar to that of SX, with no diphthongs. However, it has an alveolar nasal and a velar nasal in coda position as well as the glottal stop [ʔ] (Duanmu 1999).

weak syllables, because all the syllables in Shanghai have the same structure; (iii) there is no phonological evidence that all rhymes in Shanghai are heavy (Duanmu 1999: 493). Weak syllables, however, are all monomoraic (Duanmu 2000a). However, I assume that all syllables in SX are heavy when stressed and can be weak when unstressed. The most important phonological evidence that the rhymes in SX are basically heavy is that contour tones should be realized on two moras or by bimoraic rhymes, because a contour has two tone features, either [hl] or [lh], and one mora can bear at most one tone feature. The bimoraic status is realized by tense vowels or VC combinations if the nucleus vowels are [−tense]. Thus, the two syllable-final consonants in SX, [ŋ] and [ʔ], also contribute to weight. For example, the syllables like [be³¹] 'compensate' and [pɛʔ⁵] 'eight' are well-formed because the glottal stop is also moraic. But *[Cɛ][37] is ill-formed in SX for its lack of enough weight when stressed. This can be represented in terms of the weight domain as in (50):

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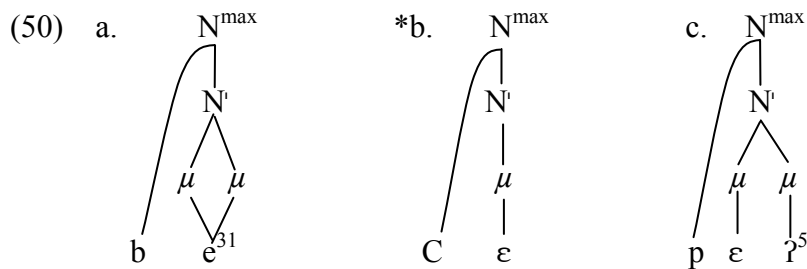
Syllables like (50b) are not acceptable in SX, as stipulated by the segment filter (see (45) in ch.3). Although syllables ending with the glottal stop is phonetically short and realized in one-digit tone pitch, phonologically, the tones [5] on [pɛʔ] and [31] on [be] are both born by bimoraic syllables,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

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

---

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

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

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

(51)   WSP ≫ MAX-μ[38]

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

(52)

$$N^{max}$$
$$N^{\mu}$$
$$N^{'}$$
$$\mu$$
$$C \quad G \quad V \quad C$$

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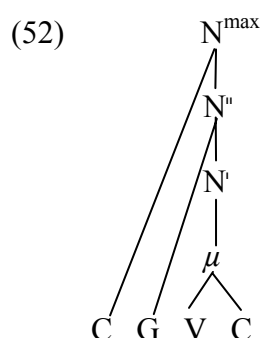
[38] In (51), WSP (Prince 1990) requires a stressed syllable to be bimoraic; MAX-μ (after Kager 1999) requires an unstressed syllable to be monomoraic.

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

4.6 Phonotactics

4.6.1   Simplicity of segment sequences in SX

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

---

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

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

As was discussed above, the X-bar structure in (37) very well captures the segmental organization of the syllable structure in SX, which tells us that the maximal segment sequences of a SX syllable is CGVC, each segment belonging to a different hierarchical constituent. Except for [ɑɒ] as in [hɑɒ³⁵] 'good', which is a combination of vowels (VV) in the nucleus, there is no two-segment sequence under a minimal sub-constituent in a syllable. In other words, the segment sequence is rather simple in terms of phonotactics. The simplicity of SX segment sequences can be briefly summarized as follows:

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

4.6.2   Distribution of the Initials and the Finals

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

As was discussed previously in this chapter, syllables in SX are typically parsed into two parts: Initial (onset consonant) and Final (all that is left after the onset). As a sub-constituent, the Final always appears as one integral unit in the forms of VC, GV and GVC, as was discussed in chapter 2. There are no segment sequence constraints involved between V and C (except for the constraint that any [−tense] nucleus vowel has to be followed by a coda), or between G and V, or G and V and C, each combination functioning as a sort of closed element, which means their number is

limited so that an increase of the number of lexical syllables is only made possible by combinations of Initials, Finals and tones. As was discussed in chapter 2, there are 29 consonants, 48 finals (in five different types) and eight tones. Roughly speaking, there should be logically 620 syllables (excluding the shaded cells for non-existing syllables and '?' cells for undecided status, as shown in the tables of (54), (64), (70) and (73) below, and disregarding the different tones) and when different tones are added, the number of syllables could be about three or four times bigger (considering four tones are in complementary distribution with the other four in obstruent-initial syllables, and some undecided possibilities and accidental gaps. However, no stock has been taken of the existing SX syllables, most of which have more than one lexical meaning for each, symbolized by different Chinese characters.

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

(53) 48 Finals in SX:

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

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

such combinations (syllables), focusing on the distribution of the 29 consonants, when combined with the different Finals.

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

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

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

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

segment distribution in the syllable combinations between the onset consonants and the Final V, VV or Ṽ can be formalized as follows:

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

(55)   /i/ → [ɻ] / $\begin{bmatrix} +\text{cons} \\ +\text{apical} \end{bmatrix}$ ___

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

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

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

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

(57) *{+ant, –lateral/+dors}-[y]

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

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

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

(59) $\begin{bmatrix} +\text{nas} \\ +\text{cor} \end{bmatrix}$ → [ɲ] / ___ $\begin{bmatrix} -\text{back} \\ +\text{high} \end{bmatrix}$

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

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

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

(61) *[ŋ][+high]

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

(62) *[h][+high, −back]

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

(63) ?{+back, +con}-[ə̃]⁴⁰

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

---

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

(64)   Cross-tabulation of C against VC

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

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

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

(65) ?{+labial, +cons}-[ə][41]

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

(66) *{+ant, +cont}-[ɪŋ]

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

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

(67) [tsɪʔ⁵]   'fold'   [tɕɪʔ⁵]   'urgent'
      [tsʰɪʔ⁵] 'exit'   [tɕʰɪʔ⁵]   'eat'
      [dzɪʔ³]   'nephew'   [dʑɪʔ³]   'fetch (water)'
      [sɪʔ⁵]   'brush'   [ɕɪʔ⁵]   'snow'
      [zɪʔ³]   'enter'   [ʑɪʔ³]   'medicine'

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

---

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

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

(68)   CV → Ci li CV lV
    a.   pʰi li pʰa la   'noise of fire crackers'
         ti li ta la   'sound of rain drop'
         tʰi li tʰu lu   'slurping'
    b.   tɕi li k(ʷ)a la   'chattering noise'
         tɕʰi li kʰ(ʷ)a   'noise of falling objects'
         ɕi li xu lu   'eating fast'

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

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

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

---

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

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

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

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

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

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

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

　　　　The cross-tabulation of C against GV and GṼ in (70) shows quite a different picture, with only about 31% of combinations which are attested

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

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

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

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

(72) *[+apical]-GV

(iii) [n] cannot occur before [jV], due to the rule in (59).

(iv) The alveolo-palatal affricates and fricatives cannot occur before [wV], which also follows the constraint AGREECV/G[+H/−B] in (60).

(v) [ŋ] cannot occur before any GV combination, due to the constraint *[ŋ][+high] in (61).

(vi) [h] cannot occur before a [jV], because of the constraint *[h][+high, −back].

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

---

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

(73) Cross-tabulation of C against GVC

	jaŋ	jaʔ	joŋ	joʔ	waŋ	wɒŋ	woŋ	waʔ	woʔ
p									
pʰ									
b									
m									
f									
v									
ts									
tsʰ									
dz									
s									
z									
n									
l			?	?					
t									
tʰ									
d									
tɕ									
tɕʰ									
dʑ									
ɕ									
ʑ									
ɲ									
k									
kʰ									
g									
ŋ									
h									
(ɦ)									
(ʔ)									

The cross-tabulation in (73) shows an even smaller proportion of the possible syllables combining the 29 consonants and the nine GVC Finals in SX, with only 24.5% of all the combinations being existing syllables. The distribution shows that all GVC combinations seem only to yield possible syllables when preceded by alveolo-palatal, dorsal and glottal

consonants, and that all [ant] consonants are irrelevant to the syllable formation with GVC, except the lateral [l], which can occur before [jaŋ] and [jaʔ], as shown in (74):

(74)   [ljaŋ²²]   'two'
        [ljaʔ³]   'omit'

The combinations between the onset consonants and GVC can be formalized as follows:

(i)   No [+ant, –lateral] consonant can combine with GVC, which suggests a constraint in (75):

(75)   *{+ant, –lateral}GVC

(ii)   Alveolo-palatal affricates and fricatives cannot occur before [wVG], which still follows the constraint of AGREECV/G[+H, −B] in (60).

(iii) Velar stops cannot occur before [jGC], due to the constraint *[+dors]-[−back, +high] in (58).

(iv)   [ŋ] cannot occur before any GVC combination, due to the constraint *[ŋ][+high] in (61).

(v)   [h] cannot occur before any [jVC] combination due to the constraint *[h][+high, −back].

In brief summary, all the four tables with the 29 consonants and the 48 Finals in SX present an overall distribution of all these segments in SX, from which we have found that there is still a lot of room for more syllables (even more so when combined with different tones, though the Finals are fixed and limited in number. The distribution of all the segments in SX as shown in (54), (64), (70) and (73) also proves that the pre-nuclear glides are not in the onset, and that GV and GVC occur as a sub-syllabic constituent in the formation of syllables in SX. The distribution of the segment sequences as shown in the above four tables involve different phonotactic constraints which were discussed above. Some consonants are in complementary distribution with others; some can occur before more different vowels or Finals than others. The four cross-tabulations also show the phenomenon of segment distribution, viz. different onset consonants are decided by the following glides (if there is one) or by the nucleus vowel (when there is no medial glide) while the allophonic

vowels in the Finals are decided by either the preceding onset consonants or the following coda. However, the phonotactic constraints of SX, as a tonal language, also involve other constrains in terms of the tonal system. This will be the topic of the next chapter.